4 Qualitative theory of differential equations

4.1 Introduction

In this chapter we consider the differential equation

$$\dot{\mathbf{x}} = \mathbf{f}(t,\mathbf{x}) \tag{1}$$

where

$$\mathbf{x} = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix},$$

and

$$\mathbf{f}(t,\mathbf{x}) = \begin{bmatrix} f_1(t,x_1,\dots,x_n) \\ \vdots \\ f_n(t,x_1,\dots,x_n) \end{bmatrix}$$

is a nonlinear function of $x_1,\dots,x_n$. Unfortunately, there are no known methods of solving Equation (1). This, of course, is very disappointing. However, it is not necessary, in most applications, to find the solutions of (1) explicitly. For example, let $x_1(t)$ and $x_2(t)$ denote the populations, at time t, of two species competing amongst themselves for the limited food and living space in their microcosm. Suppose, moreover, that the rates of growth of $x_1(t)$ and $x_2(t)$ are governed by the differential equation (1). In this case, we are not really interested in the values of $x_1(t)$ and $x_2(t)$ at every time t. Rather, we are interested in the qualitative properties of $x_1(t)$ and $x_2(t)$. Specically, we wish to answer the following questions.

1. Do there exist values ξ_1 and ξ_2 at which the two species coexist together in a steady state? That is to say, are there numbers ξ_1, ξ_2 such that $x_1(t)\equiv\xi_1$, $x_2(t)\equiv\xi_2$ is a solution of (1)? Such values ξ_1,ξ_2, if they exist, are called *equilibrium points* of (1).

2. Suppose that the two species are coexisting in equilibrium. Suddenly, we add a few members of species 1 to the microcosm. Will $x_1(t)$ and $x_2(t)$ remain close to their equilibrium values for all future time? Or perhaps the extra few members give species 1 a large advantage and it will proceed to annihilate species 2.

3. Suppose that x_1 and x_2 have arbitrary values at $t=0$. What happens as t approaches infinity? Will one species ultimately emerge victorious, or will the struggle for existence end in a draw?

More generally, we are interested in determining the following properties of solutions of (1).

1. Do there exist equilibrium values

$$\mathbf{x}^0=\begin{bmatrix} x_1^0 \\ \vdots \\ x_n^0 \end{bmatrix}$$

for which $\mathbf{x}(t)\equiv\mathbf{x}^0$ is a solution of (1)?

2. Let $\boldsymbol{\phi}(t)$ be a solution of (1). Suppose that $\boldsymbol{\psi}(t)$ is a second solution with $\boldsymbol{\psi}(0)$ very close to $\boldsymbol{\phi}(0)$; that is, $\psi_j(0)$ is very close to $\phi_j(0), j=1,\dots,n$. Will $\boldsymbol{\psi}(t)$ remain close to $\boldsymbol{\phi}(t)$ for all future time, or will $\boldsymbol{\psi}(t)$ diverge from $\boldsymbol{\phi}(t)$ as t approaches infinity? This question is often referred to as the problem of *stability*. It is the most fundamental problem in the qualitative theory of differential equations, and has occupied the attention of many mathematicians for the past hundred years.

3. What happens to solutions $\mathbf{x}(t)$ of (1) as t approaches infinity? Do all solutions approach equilibrium values? If they don't approach equilibrium values, do they at least approach a periodic solution?

This chapter is devoted to answering these three questions. Remarkably, we can often give satisfactory answers to these questions, even though we cannot solve Equation (1) explicitly. Indeed, the first question can be answered immediately. Observe that $\dot{\mathbf{x}}(t)$ is identically zero if $\mathbf{x}(t)\equiv\mathbf{x}^0$. Hence, $\mathbf{x}^0$ is an equilibrium value of (1), if, and only if,

$$\mathbf{f}(t,\mathbf{x}^0)\equiv\mathbf{0}. \tag{2}$$

Example 1. Find all equilibrium values of the system of differential equations

$$\frac{dx_1}{dt}=1-x_2,\qquad \frac{dx_2}{dt}=x_1^3+x_2.$$

Solution.

$$\mathbf{x}^0=\begin{pmatrix} x_1^0 \\ x_2^0 \end{pmatrix}$$

is an equilibrium value if, and only if, $1-x_2^0=0$ and $(x_1^0)^3+x_2^0=0$. This implies that $x_2^0=1$ and $x_1^0=-1$. Hence $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ is the only equilibrium value of this system.

Example 2. Find all equilibrium solutions of the system

$$\frac{dx}{dt}=(x-1)(y-1), \qquad \frac{dy}{dt}=(x+1)(y+1).$$

Solution.

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

is an equilibrium value of this system if, and only if, $(x_0-1)(y_0-1)=0$ and $(x_0+1)(y_0+1)=0$. The first equation is satisfied if either x_0 or y_0 is 1, while the second equation is satisfied if either x_0 or y_0 is -1. Hence, $x=1$, $y=-1$ and $x=-1$, $y=1$ are the equilibrium solutions of this system.

The question of stability is of paramount importance in all physical applications, since we can never measure initial conditions exactly. For example, consider the case of a particle of mass one kgm attached to an elastic spring of force constant 1 N/m which is moving in a frictionless medium. In addition, an external force $F(t)=\cos 2t$ N is acting on the particle. Let $y(t)$ denote the position of the particle relative to its equilibrium position. Then $(d^2y/dt^2)+y=\cos 2t$. We convert this second-order equation into a system of two first-order equations by setting $x_1=y$, $x_2=y'$. Then,

$$\frac{dx_1}{dt}=x_2, \qquad \frac{dx_2}{dt}=-x_1+\cos 2t. \tag{3}$$

The functions $y_1(t)=\sin t$ and $y_2(t)=\cos t$ are two independent solutions of the homogeneous equation $y''+y=0$. Moreover, $y=-\frac{1}{3}\cos 2t$ is a particular solution of the nonhomogeneous equation. Therefore, every solution

$$\mathbf{x}(t)=\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

of (3) is of the form

$$\mathbf{x}(t)=c_1\begin{pmatrix} \sin t \\ \cos t \end{pmatrix}+c_2\begin{pmatrix} \cos t \\ -\sin t \end{pmatrix}+\begin{bmatrix} -\frac{1}{3}\cos 2t \\ \frac{2}{3}\sin 2t \end{bmatrix}. \tag{4}$$

At time $t=0$ we measure the position and velocity of the particle and obtain $y(0)=1, y'(0)=0$. This implies that $c_1=0$ and $c_2=\frac{4}{3}$. Consequently, the position and velocity of the particle for all future time are given by the equation

$$\begin{pmatrix} y(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{bmatrix} \frac{4}{3}\cos t - \frac{1}{3}\cos 2t \\ -\frac{4}{3}\sin t + \frac{2}{3}\sin 2t \end{bmatrix}. \tag{5}$$

However, suppose that our measurements permit an error of magnitude 10^{-4}. Will the position and velocity of the particle remain close to the values predicted by (5)? The answer to this question had better be yes, for otherwise, Newtonian mechanics would be of no practical value to us. Fortunately, it is quite easy to show, in this case, that the position and velocity of the particle remain very close to the values predicted by (5). Let $\hat{y}(t)$ and $\hat{y}'(t)$ denote the true values of $y(t)$ and $y'(t)$ respectively. Clearly,

$$y(t)-\hat{y}(t)=\left(\tfrac{4}{3}-c_2\right)\cos t - c_1 \sin t$$

$$y'(t)-\hat{y}'(t)=-c_1\cos t-\left(\tfrac{4}{3}-c_2\right)\sin t$$

where c_1 and c_2 are two constants satisfying

$$-10^{-4}\leqslant c_1 \leqslant 10^{-4}, \qquad \tfrac{4}{3}-10^{-4}\leqslant c_2 \leqslant \tfrac{4}{3}+10^{-4}.$$

We can rewrite these equations in the form

$$y(t)-\hat{y}(t)=\left[c_1^2+\left(\tfrac{4}{3}-c_2\right)^2\right]^{1/2}\cos(t-\delta_1), \qquad \tan\delta_1=\frac{c_1}{c_2-\frac{4}{3}}$$

$$y'(t)-\hat{y}'(t)=\left[c_1^2+\left(\tfrac{4}{3}-c_2\right)^2\right]^{1/2}\cos(t-\delta_2), \qquad \tan\delta_2=\frac{\frac{4}{3}-c_2}{c_1}.$$

Hence, both $y(t)-\hat{y}(t)$ and $y'(t)-\hat{y}'(t)$ are bounded in absolute value by $[c_1^2+(\frac{4}{3}-c_2)^2]^{1/2}$. This quantity is at most $\sqrt{2}\ 10^{-4}$. Therefore, the true values of $y(t)$ and $y'(t)$ are indeed close to the values predicted by (5).

As a second example of the concept of stability, consider the case of a particle of mass m which is supported by a wire, or inelastic string, of length l and of negligible mass. The wire is always straight, and the system is free to vibrate in a vertical plane. This configuration is usually referred to as a simple pendulum. The equation of motion of the pendulum is

$$\frac{d^2y}{dt^2}+\frac{g}{l}\sin y=0$$

where y is the angle which the wire makes with the vertical line A0 (see

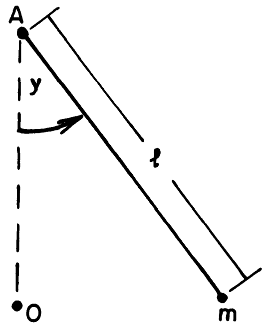

Figure 1

Figure 1). Setting $x_1 = y$ and $x_2 = dy/dt$ we see that

$$\frac{dx_1}{dt} = x_2, \qquad \frac{dx_2}{dt} = -\frac{g}{l}\sin x_1. \tag{6}$$

The system of equations (6) has equilibrium solutions $x_1 = 0$, $x_2 = 0$, and $x_1 = \pi$, $x_2 = 0$. (If the pendulum is suspended in the upright position $y = \pi$ with zero velocity, then it will remain in this upright position for all future time.) These two equilibrium solutions have very different properties. If we disturb the pendulum slightly from the equilibrium position $x_1 = 0$, $x_2 = 0$, by either displacing it slightly, or giving it a small velocity, then it will execute small oscillations about $x_1 = 0$. On the other hand, if we disturb the pendulum slightly from the equilibrium position $x_1 = \pi$, $x_2 = 0$, then it will either execute very large oscillations about $x_1 = 0$, or it will rotate around and around ad infinitum. Thus, the slightest disturbance causes the pendulum to deviate drastically from its equilibrium position $x_1 = \pi$, $x_2 = 0$. Intuitively, we would say that the equilibrium value $x_1 = 0$, $x_2 = 0$ of (6) is stable, while the equilibrium value $x_1 = \pi$, $x_2 = 0$ of (6) is unstable. This concept will be made precise in Section 4.2.

The question of stability is usually very difficult to resolve, because we cannot solve (1) explicitly. The only case which is manageable is when $\mathbf{f}(t, \mathbf{x})$ does not depend explicitly on t; that is, $\mathbf{f}$ is a function of $\mathbf{x}$ alone. Such differential equations are called *autonomous*. And even for autonomous differential equations, there are only two instances, generally, where we can completely resolve the stability question. The first case is when $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ and it will be treated in the next section. The second case is when we are only interested in the stability of an equilibrium solution of $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$. This case will be treated in Section 4.3.

Question 3 is extremely important in many applications since an answer to this question is a prediction concerning the long time evolution of the system under consideration. We answer this question, when possible, in Sections 4.6–4.8 and apply our results to some extremely important applications in Sections 4.9–4.12.

EXERCISES

In each of Problems 1–8, find all equilibrium values of the given system of differential equations.

1. $\frac{dx}{dt}=x-x^2-2xy$
$\frac{dy}{dt}=2y-2y^2-3xy$

2. $\frac{dx}{dt}=-\beta xy+\mu$
$\frac{dy}{dt}=\beta xy-\gamma y$

3. $\frac{dx}{dt}=ax-bxy$
$\frac{dy}{dt}=-cy+dxy$
$\frac{dz}{dt}=z+x^2+y^2$

4. $\frac{dx}{dt}=-x-xy^2$
$\frac{dy}{dt}=-y-yx^2$
$\frac{dz}{dt}=1-z+x^2$

5. $\frac{dx}{dt}=xy^2-x$
$\frac{dy}{dt}=x\sin\pi y$

6. $\frac{dx}{dt}=\cos y$
$\frac{dy}{dt}=\sin x-1$

7. $\frac{dx}{dt}=-1-y-e^x$
$\frac{dy}{dt}=x^2+y(e^x-1)$
$\frac{dz}{dt}=x+\sin z$

8. $\frac{dx}{dt}=x-y^2$
$\frac{dy}{dt}=x^2-y$
$\frac{dz}{dt}=e^z-x$

9. Consider the system of differential equations

$$\frac{dx}{dt}=ax+by,\qquad \frac{dy}{dt}=cx+dy. \tag{*}$$

(i) Show that $x=0, y=0$ is the only equilibrium point of (*) if $ad-bc\neq 0$.
(ii) Show that (*) has a line of equilibrium points if $ad-bc=0$.

10. Let $x=x(t), y=y(t)$ be the solution of the initial-value problem

$$\frac{dx}{dt}=-x-y,\qquad \frac{dy}{dt}=2x-y,\qquad x(0)=y(0)=1.$$

Suppose that we make an error of magnitude 10^{-4} in measuring $x(0)$ and $y(0)$. What is the largest error we make in evaluating $x(t), y(t)$ for $0\leqslant t<\infty$?

11. (a) Verify that

$$\mathbf{x}=\begin{pmatrix}1\\0\\0\end{pmatrix}e^{-t}$$

is the solution of the initial-value problem

$$\dot{\mathbf{x}}=\begin{pmatrix}1&1&1\\2&1&-1\\-3&2&4\end{pmatrix}\mathbf{x}-\begin{pmatrix}2\\2\\-3\end{pmatrix}e^{-t},\qquad \mathbf{x}(0)=\begin{pmatrix}1\\0\\0\end{pmatrix}.$$

(b) Let $\mathbf{x}=\psi(t)$ be the solution of the above differential equation which satisfies the initial condition

$$\mathbf{x}(0)=\mathbf{x}^0\neq\begin{pmatrix}1\\0\\0\end{pmatrix}.$$

Show that each component of $\psi(t)$ approaches infinity, in absolute value, as $t\to\infty$.

4.2 Stability of linear systems

In this section we consider the stability question for solutions of autonomous differential equations. Let $\mathbf{x}=\phi(t)$ be a solution of the differential equation

$$\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x}). \tag{1}$$

We are interested in determining whether $\phi(t)$ is stable or unstable. That is to say, we seek to determine whether every solution $\psi(t)$ of (1) which starts sufficiently close to $\phi(t)$ at $t=0$ must remain close to $\phi(t)$ for all future time $t\geqslant 0$. We begin with the following formal definition of stability.

Definition. The solution $\mathbf{x}=\phi(t)$ of (1) is stable if every solution $\psi(t)$ of (1) which starts sufficiently close to $\phi(t)$ at $t=0$ must remain close to $\phi(t)$ for all future time t. The solution $\phi(t)$ is unstable if there exists at least one solution $\psi(t)$ of (1) which starts near $\phi(t)$ at $t=0$ but which does not remain close to $\phi(t)$ for all future time. More precisely, the solution $\phi(t)$ is stable if for every $\varepsilon>0$ there exists $\delta=\delta(\varepsilon)$ such that

$$|\psi_j(t)-\phi_j(t)|<\varepsilon \quad \text{if} \quad |\psi_j(0)-\phi_j(0)|<\delta(\varepsilon), \qquad j=1,\dots,n$$

for every solution $\psi(t)$ of (1).

The stability question can be completely resolved for each solution of the linear differential equation

$$\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}. \tag{2}$$

This is not surprising, of course, since we can solve Equation (2) exactly. We have the following important theorem.

Theorem 1. (a) *Every solution* $\mathbf{x}=\phi(t)$ *of* (2) *is stable if all the eigenvalues of* $\mathbf{A}$ *have negative real part.*

(b) *Every solution* $\mathbf{x}=\phi(t)$ *of* (2) *is unstable if at least one eigenvalue of* $\mathbf{A}$ *has positive real part.*

(c) *Suppose that all the eigenvalues of* $\mathbf{A}$ *have real part* $\leqslant 0$ *and* $\lambda_1=i\sigma_1,\dots,\lambda_l=i\sigma_l$ *have zero real part. Let* $\lambda_j=i\sigma_j$ *have multiplicity* k_j. *This means that the characteristic polynomial of* $\mathbf{A}$ *can be factored into the form*

$$p(\lambda)=(\lambda-i\sigma_1)^{k_1}\dots(\lambda-i\sigma_l)^{k_l}q(\lambda)$$

where all the roots of $q(\lambda)$ have negative real part. Then, every solution $\mathbf{x}=\boldsymbol{\phi}(t)$ of (1) *is stable if* $\mathbf{A}$ *has k_j linearly independent eigenvectors for each eigenvalue $\lambda_j = i\sigma_j$. Otherwise, every solution $\boldsymbol{\phi}(t)$ is unstable.*

Our first step in proving Theorem 1 is to show that every solution $\boldsymbol{\phi}(t)$ is stable if the equilibrium solution $\mathbf{x}(t)\equiv\mathbf{0}$ is stable, and every solution $\boldsymbol{\phi}(t)$ is unstable if $\mathbf{x}(t)\equiv\mathbf{0}$ is unstable. To this end, let $\boldsymbol{\psi}(t)$ be any solution of (2). Observe that $\mathbf{z}(t)=\boldsymbol{\phi}(t)-\boldsymbol{\psi}(t)$ is again a solution of (2). Therefore, if the equilibrium solution $\mathbf{x}(t)\equiv\mathbf{0}$ is stable, then $\mathbf{z}(t)=\boldsymbol{\phi}(t)-\boldsymbol{\psi}(t)$ will always remain small if $\mathbf{z}(0)=\boldsymbol{\phi}(0)-\boldsymbol{\psi}(0)$ is sufficiently small. Consequently, every solution $\boldsymbol{\phi}(t)$ of (2) is stable. On the other hand suppose that $\mathbf{x}(t)\equiv\mathbf{0}$ is unstable. Then, there exists a solution $\mathbf{x}=\mathbf{h}(t)$ which is very small initially, but which becomes large as t approaches infinity. The function $\boldsymbol{\psi}(t)=\boldsymbol{\phi}(t)+\mathbf{h}(t)$ is clearly a solution of (2). Moreover, $\boldsymbol{\psi}(t)$ is close to $\boldsymbol{\phi}(t)$ initially, but diverges from $\boldsymbol{\phi}(t)$ as t increases. Therefore, every solution $\mathbf{x}=\boldsymbol{\phi}(t)$ of (2) is unstable.

Our next step in proving Theorem 1 is to reduce the problem of showing that n quantities $\psi_j(t), j=1,\dots,n$ are small to the much simpler problem of showing that only one quantity is small. This is accomplished by introducing the concept of length, or magnitude, of a vector.

Definition. Let

$$\mathbf{x}=\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

be a vector with n components. The numbers $x_1,\dots,x_n$ may be real or complex. We define the length of $\mathbf{x}$, denoted by $\|\mathbf{x}\|$ as

$$\|\mathbf{x}\|\equiv\max\{|x_1|,|x_2|,\dots,|x_n|\}.$$

For example, if

$$\mathbf{x}=\begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix},$$

then $\|\mathbf{x}\|=3$ and if

$$\mathbf{x}=\begin{bmatrix} 1+2i \\ 2 \\ -1 \end{bmatrix}$$

then $\|\mathbf{x}\|=\sqrt{5}$.

The concept of the length, or magnitude of a vector corresponds to the concept of the length, or magnitude of a number. Observe that $\|\mathbf{x}\|\geqslant 0$ for any vector $\mathbf{x}$ and $\|\mathbf{x}\|=0$ only if $\mathbf{x}=\mathbf{0}$. Second, observe that

$$\|\lambda\mathbf{x}\|=\max\{|\lambda x_1|,\dots,|\lambda x_n|\}=|\lambda|\max\{|x_1|,\dots,|x_n|\}=|\lambda|\|\mathbf{x}\|.$$

Finally, observe that

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|&=\max\{|x_1+y_1|,\dots,|x_n+y_n|\}\leqslant\max\{|x_1|+|y_1|,\dots,|x_n|+|y_n|\}\\&\leqslant\max\{|x_1|,\dots,|x_n|\}+\max\{|y_1|,\dots,|y_n|\}=\|\mathbf{x}\|+\|\mathbf{y}\|.\end{aligned}$$

Thus, our definition really captures the meaning of length.

In Section 4.7 we give a simple geometric proof of Theorem 1 for the case $n=2$. The following proof is valid for arbitrary n.

PROOF OF THEOREM 1. (a) Every solution $\mathbf{x}=\psi(t)$ of $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$ is of the form $\psi(t)=e^{\mathbf{A}t}\psi(0)$. Let $\phi_{ij}(t)$ be the ij element of the matrix $e^{\mathbf{A}t}$, and let $\psi_1^0,\dots,\psi_n^0$ be the components of $\psi(0)$. Then, the ith component of $\psi(t)$ is

$$\psi_i(t)=\phi_{i1}(t)\psi_1^0+\dots+\phi_{in}(t)\psi_n^0\equiv\sum_{j=1}^{n}\phi_{ij}(t)\psi_j^0.$$

Suppose that all the eigenvalues of $\mathbf{A}$ have negative real part. Let $-\alpha_1$ be the largest of the real parts of the eigenvalues of $\mathbf{A}$. It is a simple matter to show (see Exercise 17) that for every number $-\alpha$, with $-\alpha_1<-\alpha<0$, we can find a number K such that $|\phi_{ij}(t)|\leqslant Ke^{-\alpha t}$, $t\geqslant 0$. Consequently,

$$|\psi_i(t)|\leqslant\sum_{j=1}^{n}Ke^{-\alpha t}|\psi_j^0|=Ke^{-\alpha t}\sum_{j=1}^{n}|\psi_j^0|$$

for some positive constants K and α. Now, $|\psi_j^0|\leqslant\|\psi(0)\|$. Hence,

$$\|\psi(t)\|=\max\{|\psi_1(t)|,\dots,|\psi_n(t)|\}\leqslant nKe^{-\alpha t}\|\psi(0)\|.$$

Let $\varepsilon>0$ be given. Choose $\delta(\varepsilon)=\varepsilon/nK$. Then, $\|\psi(t)\|<\varepsilon$ if $\|\psi(0)\|<\delta(\varepsilon)$ and $t\geqslant 0$, since

$$\|\psi(t)\|\leqslant nKe^{-\alpha t}\|\psi(0)\|<nK\varepsilon/nK=\varepsilon.$$

Consequently, the equilibrium solution $\mathbf{x}(t)\equiv\mathbf{0}$ is stable.

(b) Let λ be an eigenvalue of $\mathbf{A}$ with positive real part and let $\mathbf{v}$ be an eigenvector of $\mathbf{A}$ with eigenvalue λ. Then, $\psi(t)=ce^{\lambda t}\mathbf{v}$ is a solution of $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$ for any constant c. If λ is real then $\mathbf{v}$ is also real and $\|\psi(t)\|=|c|e^{\lambda t}\|\mathbf{v}\|$. Clearly, $\|\psi(t)\|$ approaches infinity as t approaches infinity, for any choice of $c\neq 0$, no matter how small. Therefore, $\mathbf{x}(t)\equiv\mathbf{0}$ is unstable. If $\lambda=\alpha+i\beta$ is complex, then $\mathbf{v}=\mathbf{v}^1+i\mathbf{v}^2$ is also complex. In this case

$$\begin{aligned}e^{(\alpha+i\beta)t}(\mathbf{v}^1+i\mathbf{v}^2)&=e^{\alpha t}(\cos\beta t+i\sin\beta t)(\mathbf{v}^1+i\mathbf{v}^2)\\&=e^{\alpha t}\left[(\mathbf{v}^1\cos\beta t-\mathbf{v}^2\sin\beta t)+i(\mathbf{v}^1\sin\beta t+\mathbf{v}^2\cos\beta t)\right]\end{aligned}$$

is a complex-valued solution of (2). Therefore

$$\psi^1(t)=ce^{\alpha t}(\mathbf{v}^1\cos\beta t-\mathbf{v}^2\sin\beta t)$$

is a real-valued solution of (2), for any choice of constant c. Clearly,

$\|\psi^1(t)\|$ is unbounded as t approaches infinity if c and either $\mathbf{v}^1$ or $\mathbf{v}^2$ is nonzero. Thus, $\mathbf{x}(t)\equiv\mathbf{0}$ is unstable.

(c) If $\mathbf{A}$ has k_j linearly independent eigenvectors for each eigenvalue $\lambda_j = i\sigma_j$ of multiplicity k_j, then we can find a constant K such that $|(e^{\mathbf{A}t})_{ij}| \leq K$ (see Exercise 18). There, $\|\psi(t)\| \leq nK\|\psi(0)\|$ for every solution $\psi(t)$ of (2). It now follows immediately from the proof of (a) that $\mathbf{x}(t)\equiv\mathbf{0}$ is stable.

On the other hand, if $\mathbf{A}$ has fewer than k_j linearly independent eigenvectors with eigenvalue $\lambda_j = i\sigma_j$, then $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ has solutions $\psi(t)$ of the form

$$\psi(t) = ce^{i\sigma_j t}\left[\mathbf{v} + t(\mathbf{A} - i\sigma_j\mathbf{I})\mathbf{v}\right]$$

where $(\mathbf{A} - i\sigma_j\mathbf{I})\mathbf{v} \neq \mathbf{0}$. If $\sigma_j = 0$, then $\psi(t) = c(\mathbf{v} + t\mathbf{A}\mathbf{v})$ is real-valued. Moreover, $\|\psi(t)\|$ is unbounded as t approaches infinity for any choice of $c \neq 0$. Similarly, both the real and imaginary parts of $\psi(t)$ are unbounded in magnitude for arbitrarily small $\psi(0) \neq \mathbf{0}$, if $\sigma_j \neq 0$. Therefore, the equilibrium solution $\mathbf{x}(t)\equiv\mathbf{0}$ is unstable. □

If all the eigenvalues of $\mathbf{A}$ have negative real part, then every solution $\mathbf{x}(t)$ of $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ approaches zero as t approaches infinity. This follows immediately from the estimate $\|\mathbf{x}(t)\| \leq Ke^{-\alpha t}\|\mathbf{x}(0)\|$ which we derived in the proof of part (a) of Theorem 1. Thus, not only is the equilibrium solution $\mathbf{x}(t)\equiv\mathbf{0}$ stable, but every solution $\psi(t)$ of (2) approaches it as t approaches infinity. This very strong type of stability is known as *asymptotic stability*.

Definition. A solution $\mathbf{x} = \phi(t)$ of (1) is asymptotically stable if it is stable, and if every solution $\psi(t)$ which starts sufficiently close to $\phi(t)$ must approach $\phi(t)$ as t approaches infinity. In particular, an equilibrium solution $\mathbf{x}(t) = \mathbf{x}^0$ of (1) is asymptotically stable if every solution $\mathbf{x} = \psi(t)$ of (1) which starts sufficiently close to $\mathbf{x}^0$ at time $t = 0$ not only remains close to $\mathbf{x}^0$ for all future time, but ultimately approaches $\mathbf{x}^0$ as t approaches infinity.

Remark. The asymptotic stability of any solution $\mathbf{x} = \phi(t)$ of (2) is clearly equivalent to the asymptotic stability of the equilibrium solution $\mathbf{x}(t)\equiv\mathbf{0}$.

Example 1. Determine whether each solution $\mathbf{x}(t)$ of the differential equation

$$\dot{\mathbf{x}} = \begin{pmatrix} -1 & 0 & 0 \\ -2 & -1 & 2 \\ -3 & -2 & -1 \end{pmatrix}\mathbf{x}$$

is stable, asymptotically stable, or unstable.

Solution. The characteristic polynomial of the matrix

$$\mathbf{A} = \begin{pmatrix} -1 & 0 & 0 \\ -2 & -1 & 2 \\ -3 & -2 & -1 \end{pmatrix}$$

is

$$p(\lambda)=(\det\mathbf{A}-\lambda\mathbf{I})=\det\begin{bmatrix}-1-\lambda & 0 & 0\\ -2 & -1-\lambda & 2\\ -3 & -2 & -1-\lambda\end{bmatrix}$$
$$=-(1+\lambda)^3-4(1+\lambda)=-(1+\lambda)(\lambda^2+2\lambda+5).$$

Hence, $\lambda=-1$ and $\lambda=-1\pm 2i$ are the eigenvalues of $\mathbf{A}$. Since all three eigenvalues have negative real part, we conclude that every solution of the differential equation $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$ is asymptotically stable.

Example 2. Prove that every solution of the differential equation

$$\dot{\mathbf{x}}=\begin{pmatrix}1 & 5\\ 5 & 1\end{pmatrix}\mathbf{x}$$

is unstable.
Solution. The characteristic polynomial of the matrix

$$\mathbf{A}=\begin{pmatrix}1 & 5\\ 5 & 1\end{pmatrix}$$

is

$$p(\lambda)=\det(\mathbf{A}-\lambda\mathbf{I})=\det\begin{pmatrix}1-\lambda & 5\\ 5 & 1-\lambda\end{pmatrix}=(1-\lambda)^2-25.$$

Hence $\lambda=6$ and $\lambda=-4$ are the eigenvalues of $\mathbf{A}$. Since one eigenvalue of $\mathbf{A}$ is positive, we conclude that every solution $\mathbf{x}=\boldsymbol{\phi}(t)$ of $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$ is unstable.

Example 3. Show that every solution of the differential equation

$$\dot{\mathbf{x}}=\begin{pmatrix}0 & -3\\ 2 & 0\end{pmatrix}\mathbf{x}$$

is stable, but not asymptotically stable.
Solution. The characteristic polynomial of the matrix

$$\mathbf{A}=\begin{pmatrix}0 & -3\\ 2 & 0\end{pmatrix}$$

is

$$p(\lambda)=\det(\mathbf{A}-\lambda\mathbf{I})=\det\begin{pmatrix}-\lambda & -3\\ 2 & -\lambda\end{pmatrix}=\lambda^2+6.$$

Thus, the eigenvalues of $\mathbf{A}$ are $\lambda=\pm\sqrt{6}\,i$. Therefore, by part (c) of Theorem 1, every solution $\mathbf{x}=\boldsymbol{\phi}(t)$ of $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$ is stable. However, no solution is asymptotically stable. This follows immediately from the fact that the general solution of $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$ is

$$\mathbf{x}(t)=c_1\begin{pmatrix}-\sqrt{6}\sin\sqrt{6}\,t\\ 2\cos\sqrt{6}\,t\end{pmatrix}+c_2\begin{pmatrix}\sqrt{6}\cos\sqrt{6}\,t\\ 2\sin\sqrt{6}\,t\end{pmatrix}.$$

Hence, every solution $\mathbf{x}(t)$ is periodic, with period $2\pi/\sqrt{6}$, and no solution $\mathbf{x}(t)$ (except $\mathbf{x}(t)\equiv\mathbf{0}$) approaches 0 as t approaches infinity.

Example 4. Show that every solution of the differential equation

$$\dot{\mathbf{x}}=\begin{bmatrix} 2 & -3 & 0 \\ 0 & -6 & -2 \\ -6 & 0 & -3 \end{bmatrix}\mathbf{x}$$

is unstable.

Solution. The characteristic polynomial of the matrix

$$\mathbf{A}=\begin{bmatrix} 2 & -3 & 0 \\ 0 & -6 & -2 \\ -6 & 0 & -3 \end{bmatrix}$$

is

$$p(\lambda)=\det(\mathbf{A}-\lambda\mathbf{I})=\det\begin{bmatrix} 2-\lambda & -3 & 0 \\ 0 & -6-\lambda & -2 \\ -6 & 0 & -3-\lambda \end{bmatrix}=-\lambda^2(\lambda+7).$$

Hence, the eigenvalues of $\mathbf{A}$ are $\lambda=-7$ and $\lambda=0$. Every eigenvector $\mathbf{v}$ of $\mathbf{A}$ with eigenvalue 0 must satisfy the equation

$$\mathbf{A}\mathbf{v}=\begin{bmatrix} 2 & -3 & 0 \\ 0 & -6 & -2 \\ -6 & 0 & -3 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

This implies that $v_1=3v_2/2$ and $v_3=-3v_2$, so that every eigenvector $\mathbf{v}$ of $\mathbf{A}$ with eigenvalue 0 must be of the form

$$\mathbf{v}=c\begin{bmatrix} 3 \\ 2 \\ -6 \end{bmatrix}.$$

Consequently, every solution $\mathbf{x}=\boldsymbol{\phi}(t)$ of $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$ is unstable, since $\lambda=0$ is an eigenvalue of multiplicity two and $\mathbf{A}$ has only one linearly independent eigenvector with eigenvalue 0.

Exercises

Determine the stability or instability of all solutions of the following systems of differential equations.

1. $\dot{\mathbf{x}}=\begin{pmatrix} 1 & 1 \\ -2 & -2 \end{pmatrix}\mathbf{x}$

2. $\dot{\mathbf{x}}=\begin{pmatrix} -3 & -4 \\ 2 & 1 \end{pmatrix}\mathbf{x}$

3. $\dot{\mathbf{x}}=\begin{pmatrix} -5 & 3 \\ -1 & 1 \end{pmatrix}\mathbf{x}$

4. $\dot{\mathbf{x}}=\begin{pmatrix} 1 & -4 \\ 4 & -7 \end{pmatrix}\mathbf{x}$

5. $\dot{\mathbf{x}}=\begin{pmatrix} -7 & 1 & -6 \\ 10 & -4 & 12 \\ 2 & -1 & 1 \end{pmatrix}\mathbf{x}$

6. $\dot{\mathbf{x}}=\begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}\mathbf{x}$

7. $\dot{\mathbf{x}}=\begin{pmatrix} 0 & 2 & 1 \\ -1 & -3 & -1 \\ 1 & 1 & -1 \end{pmatrix}\mathbf{x}$ **8.** $\dot{\mathbf{x}}=\begin{pmatrix} -2 & 1 & 1 \\ -3 & 2 & 3 \\ 1 & -1 & -2 \end{pmatrix}\mathbf{x}$

9. $\dot{\mathbf{x}}=\begin{bmatrix} 0 & 2 & 0 & 0 \\ -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & -2 & 0 \end{bmatrix}\mathbf{x}$ **10.** $\dot{\mathbf{x}}=\begin{bmatrix} 0 & 2 & 1 & 0 \\ -2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & -2 & 0 \end{bmatrix}\mathbf{x}$

11. Determine whether the solutions $x(t)\equiv 0$ and $x(t)\equiv 1$ of the single scalar equation $\dot{x}=x(1-x)$ are stable or unstable.

12. Determine whether the solutions $x(t)\equiv 0$ and $x(t)\equiv 1$ of the single scalar equation $\dot{x}=-x(1-x)$ are stable or unstable.

13. Consider the differential equation $\dot{x}=x^2$. Show that all solutions $x(t)$ with $x(0)\geqslant 0$ are unstable while all solutions $x(t)$ with $x(0)<0$ are asymptotically stable.

14. Consider the system of differential equations

$$\begin{aligned}\dot{x}_1&=x_2\\ \dot{x}_2&=-\tfrac{1}{2}\left[x_1^2+(x_1^2+4x_2^2)^{1/2}\right]x_1.\end{aligned}\qquad (*)$$

(a) Show that

$$\mathbf{x}(t)=\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}=\begin{pmatrix} c\sin(ct+d) \\ c^2\cos(ct+d) \end{pmatrix}$$

is a solution of (*) for any choice of constants c and d.

(b) Assume that a solution $\mathbf{x}(t)$ of (*) is uniquely determined once $x_1(0)$ and $x_2(0)$ are prescribed. Prove that (a) represents the general solution of (*).

(c) Show that the solution $\mathbf{x}=\mathbf{0}$ of (*) is stable, but not asymptotically stable.

(d) Show that every solution $\mathbf{x}(t)\not\equiv\mathbf{0}$ of (*) is unstable.

15. Show that the stability of any solution $\mathbf{x}(t)$ of the nonhomogeneous equation $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}+\mathbf{f}(t)$ is equivalent to the stability of the equilibrium solution $\mathbf{x}\equiv\mathbf{0}$ of the homogeneous equation $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$.

16. Determine the stability or instability of all solutions $\mathbf{x}(t)$ of the differential equation

$$\dot{\mathbf{x}}=\begin{pmatrix} -1 & -1 \\ 2 & -1 \end{pmatrix}\mathbf{x}+\begin{pmatrix} 1 \\ 5 \end{pmatrix}.$$

17. (a) Let $f(t)=t^a e^{-bt}$, for some positive constants a and b, and let c be a positive number smaller than b. Show that we can find a positive constant K such that $|f(t)|\leqslant Ke^{-ct}$, $0\leqslant t<\infty$. *Hint*: Show that $f(t)/e^{-ct}$ approaches zero as t approaches infinity.

(b) Suppose that all the eigenvalues of $\mathbf{A}$ have negative real part. Show that we can find positive constants K and α such that $|(e^{\mathbf{A}t})_{ij}|\leqslant Ke^{-\alpha t}$ for $1\leqslant i$, $j\leqslant n$. *Hint*: Each component of $e^{\mathbf{A}t}$ is a finite linear combination of functions of the form $q(t)e^{\lambda t}$, where $q(t)$ is a polynomial in t (of degree $\leqslant n-1$) and λ is an eigenvalue of $\mathbf{A}$.

18. (a) Let $\mathbf{x}(t)=e^{i\sigma t}\mathbf{v}$, σ real, be a complex-valued solution of $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$. Show that both the real and imaginary parts of $\mathbf{x}(t)$ are bounded solutions of $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$.

(b) Suppose that all the eigenvalues of $\mathbf{A}$ have real part $\leqslant 0$ and $\lambda_1=i\sigma_1,\dots,\lambda_l=i\sigma_l$ have zero real part. Let $\lambda_j=i\sigma_j$ have multiplicity k_j, and suppose that $\mathbf{A}$ has k_j linearly independent eigenvectors for each eigenvalue λ_j, $j=1,\dots,l$. Prove that we can find a constant K such that $|(e^{\mathbf{A}t})_{ij}|\leqslant K$.

19. Let

$$\mathbf{x}=\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

and define $\|x\|_1=|x_1|+\dots+|x_n|$. Show that

(i) $\|\mathbf{x}\|_1\geqslant 0$ and $\|\mathbf{x}\|_1=0$ only if $\mathbf{x}=\mathbf{0}$

(ii) $\|\lambda\mathbf{x}\|_1=|\lambda|\,\|\mathbf{x}\|_1$

(iii) $\|\mathbf{x}+\mathbf{y}\|_1\leqslant\|\mathbf{x}\|_1+\|\mathbf{y}\|_1$

20. Let

$$\mathbf{x}=\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

and define $\|\mathbf{x}\|_2=[|x_1|^2+\dots+|x_n|^2]^{1/2}$. Show that

(i) $\|\mathbf{x}\|_2\geqslant 0$ and $\|\mathbf{x}\|_2=0$ only if $\mathbf{x}=\mathbf{0}$

(ii) $\|\lambda\mathbf{x}\|_2=|\lambda|\,\|\mathbf{x}\|_2$

(iii) $\|\mathbf{x}+\mathbf{y}\|_2\leqslant\|\mathbf{x}\|_2+\|\mathbf{y}\|_2$

21. Show that there exist constants M and N such that

$$M\|\mathbf{x}\|_1\leqslant\|\mathbf{x}\|_2\leqslant N\|\mathbf{x}\|_1.$$

4.3 Stability of equilibrium solutions

In Section 4.2 we treated the simple equation $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$. The next simplest equation is

$$\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}+\mathbf{g}(\mathbf{x}) \tag{1}$$

where

$$\mathbf{g}(\mathbf{x})=\begin{bmatrix} g_1(\mathbf{x}) \\ \vdots \\ g_n(\mathbf{x}) \end{bmatrix}$$

is very small compared to $\mathbf{x}$. Specifically we assume that

$$\frac{g_1(\mathbf{x})}{\max\{|x_1|,\dots,|x_n|\}},\dots,\frac{g_n(\mathbf{x})}{\max\{|x_1|,\dots,|x_n|\}}$$

are continuous functions of $x_1,\dots,x_n$ which vanish for $x_1=\dots=x_n=0$. This is always the case if each component of $\mathbf{g}(\mathbf{x})$ is a polynomial in $x_1,\dots,x_n$ which begins with terms of order 2 or higher. For example, if

$$\mathbf{g}(\mathbf{x})=\begin{pmatrix} x_1x_2^2 \\ x_1x_2 \end{pmatrix},$$

then both $x_1x_2^2/\max\{|x_1|,|x_2|\}$ and $x_1x_2/\max\{|x_1|,|x_2|\}$ are continuous functions of x_1,x_2 which vanish for $x_1=x_2=0$.

If $\mathbf{g}(\mathbf{0})=\mathbf{0}$ then $\mathbf{x}(t)\equiv\mathbf{0}$ is an equilibrium solution of (1). We would like to determine whether it is stable or unstable. At first glance this would seem impossible to do, since we cannot solve Equation (1) explicitly. However, if $\mathbf{x}$ is very small, then $\mathbf{g}(\mathbf{x})$ is very small compared to $\mathbf{A}\mathbf{x}$. Therefore, it seems plausible that the stability of the equilibrium solution $\mathbf{x}(t)\equiv\mathbf{0}$ of (1) should be determined by the stability of the "approximate" equation $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$. This is almost the case as the following theorem indicates.

Theorem 2. *Suppose that the vector-valued function*

$$\mathbf{g}(\mathbf{x})/\|\mathbf{x}\|\equiv\mathbf{g}(\mathbf{x})/\max\{|x_1|,\dots,|x_n|\}$$

is a continuous function of $x_1,\dots,x_n$ *which vanishes for* $\mathbf{x}=\mathbf{0}$. *Then,*

(a) *The equilibrium solution* $\mathbf{x}(t)\equiv\mathbf{0}$ *of* (1) *is asymptotically stable if the equilibrium solution* $\mathbf{x}(t)\equiv\mathbf{0}$ *of the "linearized" equation* $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$ *is asymptotically stable. Equivalently, the solution* $\mathbf{x}(t)\equiv\mathbf{0}$ *of* (1) *is asymptotically stable if all the eigenvalues of* $\mathbf{A}$ *have negative real part.*

(b) *The equilibrium solution* $\mathbf{x}(t)\equiv\mathbf{0}$ *of* (1) *is unstable if at least one eigenvalue of* $\mathbf{A}$ *has positive real part.*

(c) *The stability of the equilibrium solution* $\mathbf{x}(t)\equiv\mathbf{0}$ *of* (1) *cannot be determined from the stability of the equilibrium solution* $\mathbf{x}(t)\equiv\mathbf{0}$ *of* $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$ *if all the eigenvalues of* $\mathbf{A}$ *have real part* $\leqslant 0$ *but at least one eigenvalue of* $\mathbf{A}$ *has zero real part.*

PROOF. (a) The key step in many stability proofs is to use the variation of parameters formula of Section 3.12. This formula implies that any solution $\mathbf{x}(t)$ of (1) can be written in the form

$$\mathbf{x}(t)=e^{\mathbf{A}t}\mathbf{x}(0)+\int_0^t e^{\mathbf{A}(t-s)}\mathbf{g}(\mathbf{x}(s))\,ds. \tag{2}$$

We wish to show that $\|\mathbf{x}(t)\|$ approaches zero as t approaches infinity. To this end recall that if all the eigenvalues of $\mathbf{A}$ have negative real part, then we can find positive constants K and α such that (see Exercise 17, Section 4.2).

$$\|e^{\mathbf{A}t}\mathbf{x}(0)\|\leqslant Ke^{-\alpha t}\|\mathbf{x}(0)\|$$

and

$$\|e^{\mathbf{A}(t-s)}\mathbf{g}(\mathbf{x}(s))\|\leqslant Ke^{-\alpha(t-s)}\|\mathbf{g}(\mathbf{x}(s))\|.$$

Moreover, we can find a positive constant σ such that

$$\|\mathbf{g}(\mathbf{x})\| \leqslant \frac{\alpha}{2K}\|\mathbf{x}\| \quad \text{if} \quad \|\mathbf{x}\| \leqslant \sigma.$$

This follows immediately from our assumption that $g(\mathbf{x})/\|\mathbf{x}\|$ is continuous and vanishes at $\mathbf{x}=\mathbf{0}$. Consequently, Equation (2) implies that

$$\begin{aligned}\|\mathbf{x}(t)\| &\leqslant \|e^{\mathbf{A}t}\mathbf{x}(0)\| + \int_0^t \|e^{\mathbf{A}(t-s)}\mathbf{g}(\mathbf{x}(s))\|\,ds \\ &\leqslant Ke^{-\alpha t}\|\mathbf{x}(0)\| + \frac{\alpha}{2}\int_0^t e^{-\alpha(t-s)}\|\mathbf{x}(s)\|\,ds\end{aligned}$$

as long as $\|\mathbf{x}(s)\| \leqslant \sigma$, $0 \leqslant s \leqslant t$. Multiplying both sides of this inequality by $e^{\alpha t}$ gives

$$e^{\alpha t}\|\mathbf{x}(t)\| \leqslant K\|\mathbf{x}(0)\| + \frac{\alpha}{2}\int_0^t e^{\alpha s}\|\mathbf{x}(s)\|\,ds. \tag{3}$$

The inequality (3) can be simplified by setting $z(t)=e^{\alpha t}\|\mathbf{x}(t)\|$, for then

$$z(t) \leqslant K\|\mathbf{x}(0)\| + \frac{\alpha}{2}\int_0^t z(s)\,ds. \tag{4}$$

We would like to differentiate both sides of (4) with respect to t. However, we cannot, in general, differentiate both sides of an inequality and still preserve the sense of the inequality. We circumvent this difficulty by the clever trick of setting

$$U(t) = \frac{\alpha}{2}\int_0^t z(s)\,ds.$$

Then

$$\frac{dU(t)}{dt} = \frac{\alpha}{2}z(t) \leqslant \frac{\alpha}{2}K\|\mathbf{x}(0)\| + \frac{\alpha}{2}U(t)$$

or

$$\frac{dU(t)}{dt} - \frac{\alpha}{2}U(t) \leqslant \frac{\alpha K}{2}\|\mathbf{x}(0)\|.$$

Multiplying both sides of this inequality by the integrating factor $e^{-\alpha t/2}$ gives

$$\frac{d}{dt}e^{-\alpha t/2}U \leqslant \frac{\alpha K}{2}\|\mathbf{x}(0)\|e^{-\alpha t/2},$$

or

$$\frac{d}{dt}e^{-\alpha t/2}\left[U(t) + K\|\mathbf{x}(0)\|\right] \leqslant 0.$$

Consequently,

$$e^{-\alpha t/2}\left[U(t) + K\|\mathbf{x}(0)\|\right] \leqslant U(0) + K\|\mathbf{x}(0)\| = K\|\mathbf{x}(0)\|,$$

so that $U(t) \leqslant -K\|\mathbf{x}(0)\| + K\|\mathbf{x}(0)\|e^{\alpha t/2}$. Returning to the inequality (4), we see that

$$\|\mathbf{x}(t)\| = e^{-\alpha t}z(t) \leqslant e^{-\alpha t}\left[K\|\mathbf{x}(0)\| + U(t)\right]$$
$$\leqslant K\|\mathbf{x}(0)\|e^{-\alpha t/2} \tag{5}$$

as long as $\|\mathbf{x}(s)\| \leqslant \sigma$, $0 \leqslant s \leqslant t$. Now, if $\|\mathbf{x}(0)\| \leqslant \sigma/K$, then the inequality (5) guarantees that $\|\mathbf{x}(t)\| \leqslant \sigma$ for all future time t. Consequently, the inequality (5) is true for all $t \geqslant 0$ if $\|\mathbf{x}(0)\| \leqslant \sigma/K$. Finally, observe from (5) that $\|\mathbf{x}(t)\| \leqslant K\|\mathbf{x}(0)\|$ and $\|\mathbf{x}(t)\|$ approaches zero as t approaches infinity. Therefore, the equilibrium solution $\mathbf{x}(t) \equiv \mathbf{0}$ of (1) is asymptotically stable.
(b) The proof of (b) is too difficult to present here.
(c) We will present two differential equations of the form (1) where the nonlinear term $\mathbf{g}(\mathbf{x})$ determines the stability of the equilibrium solution $\mathbf{x}(t) \equiv \mathbf{0}$. Consider first the system of differential equations

$$\frac{dx_1}{dt} = x_2 - x_1(x_1^2 + x_2^2), \qquad \frac{dx_2}{dt} = -x_1 - x_2(x_1^2 + x_2^2). \tag{6}$$

The linearized equation is

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

and the eigenvalues of the matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

are $\pm i$. To analyze the behavior of the nonlinear system (6) we multiply the first equation by x_1, the second equation by x_2 and add; this gives

$$x_1\frac{dx_1}{dt} + x_2\frac{dx_2}{dt} = -x_1^2(x_1^2 + x_2^2) - x_2^2(x_1^2 + x_2^2)$$
$$= -(x_1^2 + x_2^2)^2.$$

But

$$x_1\frac{dx_1}{dt} + x_2\frac{dx_2}{dt} = \frac{1}{2}\frac{d}{dt}(x_1^2 + x_2^2).$$

Hence,

$$\frac{d}{dt}(x_1^2 + x_2^2) = -2(x_1^2 + x_2^2)^2.$$

This implies that

$$x_1^2(t) + x_2^2(t) = \frac{c}{1 + 2ct},$$

where

$$c = x_1^2(0) + x_2^2(0).$$

Thus, $x_1^2(t)+x_2^2(t)$ approaches zero as t approaches infinity for any solution $x_1(t)$, $x_2(t)$ of (6). Moreover, the value of $x_1^2+x_2^2$ at any time t is always less than its value at $t=0$. We conclude, therefore, that $x_1(t)\equiv 0$, $x_2(t)\equiv 0$ is asymptotically stable.

On the other hand, consider the system of equations

$$\frac{dx_1}{dt}=x_2+x_1\left(x_1^2+x_2^2\right), \qquad \frac{dx_2}{dt}=-x_1-x_2\left(x_1^2+x_2^2\right). \tag{7}$$

Here too, the linearized system is

$$\dot{\mathbf{x}}=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\mathbf{x}.$$

In this case, though, $(d/dt)(x_1^2+x_2^2)=2(x_1^2+x_2^2)^2$. This implies that

$$x_1^2(t)+x_2^2(t)=\frac{c}{1-2ct}, \qquad c=x_1^2(0)+x_2^2(0).$$

Notice that every solution $x_1(t)$, $x_2(t)$ of (7) with $x_1^2(0)+x_2^2(0)\neq 0$ approaches infinity in finite time. We conclude, therefore, that the equilibrium solution $x_1(t)\equiv 0$, $x_2(t)\equiv 0$ is unstable. □

Example 1. Consider the system of differential equations

$$\frac{dx_1}{dt}=-2x_1+x_2+3x_3+9x_2^3$$
$$\frac{dx_2}{dt}=-6x_2-5x_3+7x_3^5$$
$$\frac{dx_3}{dt}=-x_3+x_1^2+x_2^2.$$

Determine, if possible, whether the equilibrium solution $x_1(t)\equiv 0$, $x_2(t)\equiv 0$, $x_3(t)\equiv 0$ is stable or unstable.

Solution. We rewrite this system in the form $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}+\mathbf{g}(\mathbf{x})$ where

$$\mathbf{x}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \qquad \mathbf{A}=\begin{bmatrix} -2 & 1 & 3 \\ 0 & -6 & -5 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{g}(\mathbf{x})=\begin{bmatrix} 9x_2^3 \\ 7x_3^5 \\ x_1^2+x_2^2 \end{bmatrix}.$$

The function $\mathbf{g}(\mathbf{x})$ satisfies the hypotheses of Theorem 2, and the eigenvalues of $\mathbf{A}$ are -2, -6 and -1. Hence, the equilibrium solution $\mathbf{x}(t)\equiv\mathbf{0}$ is asymptotically stable.

Theorem 2 can also be used to determine the stability of equilibrium solutions of arbitrary autonomous differential equations. Let $\mathbf{x}^0$ be an equilibrium value of the differential equation

$$\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x}) \tag{8}$$

and set $\mathbf{z}(t)=\mathbf{x}(t)-\mathbf{x}^0$. Then

$$\dot{\mathbf{z}}=\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x}^0+\mathbf{z}). \tag{9}$$

Clearly, $\mathbf{z}(t)\equiv\mathbf{0}$ is an equilibrium solution of (9) and the stability of $\mathbf{x}(t)\equiv\mathbf{x}^0$ is equivalent to the stability of $\mathbf{z}(t)\equiv\mathbf{0}$.

Next, we show that $\mathbf{f}(\mathbf{x}^0+\mathbf{z})$ can be written in the form $\mathbf{f}(\mathbf{x}^0+\mathbf{z})=\mathbf{A}\mathbf{z}+\mathbf{g}(\mathbf{z})$ where $\mathbf{g}(\mathbf{z})$ is small compared to $\mathbf{z}$.

Lemma 1. *Let* $\mathbf{f}(\mathbf{x})$ *have two continuous partial derivatives with respect to each of its variables* $x_1,\dots,x_n$. *Then,* $\mathbf{f}(\mathbf{x}^0+\mathbf{z})$ *can be written in the form*

$$\mathbf{f}(\mathbf{x}^0+\mathbf{z})=\mathbf{f}(\mathbf{x}^0)+\mathbf{A}\mathbf{z}+\mathbf{g}(\mathbf{z}) \tag{10}$$

where $\mathbf{g}(\mathbf{z})/\max\{|z_1|,\dots,|z_n|\}$ *is a continuous function of* $\mathbf{z}$ *which vanishes for* $\mathbf{z}=\mathbf{0}$.

PROOF #1. Equation (10) is an immediate consequence of Taylor's Theorem which states that each component $f_j(\mathbf{x}^0+\mathbf{z})$ of $\mathbf{f}(\mathbf{x}^0+\mathbf{z})$ can be written in the form

$$f_j(\mathbf{x}^0+\mathbf{z})=f_j(\mathbf{x}^0)+\frac{\partial f_j(\mathbf{x}^0)}{\partial x_1}z_1+\dots+\frac{\partial f_j(\mathbf{x}^0)}{\partial x_n}z_n+g_j(\mathbf{z})$$

where $g_j(\mathbf{z})/\max\{|z_1|,\dots,|z_n|\}$ is a continuous function of $\mathbf{z}$ which vanishes for $\mathbf{z}=\mathbf{0}$. Hence,

$$\mathbf{f}(\mathbf{x}^0+\mathbf{z})=\mathbf{f}(\mathbf{x}^0)+\mathbf{A}\mathbf{z}+\mathbf{g}(\mathbf{z})$$

where

$$\mathbf{A}=\begin{pmatrix}\dfrac{\partial f_1(\mathbf{x}^0)}{\partial x_1} & \cdots & \dfrac{\partial f_1(\mathbf{x}^0)}{\partial x_n}\\ \vdots & & \vdots\\ \dfrac{\partial f_n(\mathbf{x}^0)}{\partial x_1} & \cdots & \dfrac{\partial f_n(\mathbf{x}^0)}{\partial x_n}\end{pmatrix}. \qquad \square$$

PROOF #2. If each component of $\mathbf{f}(\mathbf{x})$ is a polynomial (possibly infinite) in $x_1,\dots,x_n$, then each component of $\mathbf{f}(\mathbf{x}^0+\mathbf{z})$ is a polynomial in $z_1,\dots,z_n$. Thus,

$$f_j(\mathbf{x}^0+\mathbf{z})=a_{j0}+a_{j1}z_1+\dots+a_{jn}z_n+g_j(\mathbf{z}) \tag{11}$$

where $g_j(\mathbf{z})$ is a polynomial in $z_1,\dots,z_n$ beginning with terms of order two. Setting $\mathbf{z}=\mathbf{0}$ in (11) gives $f_j(\mathbf{x}^0)=a_{j0}$. Hence,

$$\mathbf{f}(\mathbf{x}^0+\mathbf{z})=\mathbf{f}(\mathbf{x}^0)+\mathbf{A}\mathbf{z}+\mathbf{g}(\mathbf{z}),\qquad \mathbf{A}=\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}$$

and each component of $\mathbf{g}(\mathbf{z})$ is a polynomial in $z_1,\dots,z_n$ beginning with terms of order two. □

Theorem 2 and Lemma 1 provide us with the following algorithm for determining whether an equilibrium solution $\mathbf{x}(t)\equiv\mathbf{x}^0$ of $\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x})$ is stable or unstable:

1. Set $\mathbf{z}=\mathbf{x}-\mathbf{x}^0$.
2. Write $\mathbf{f}(\mathbf{x}^0+\mathbf{z})$ in the form $\mathbf{A}\mathbf{z}+\mathbf{g}(\mathbf{z})$ where $\mathbf{g}(\mathbf{z})$ is a vector-valued polynomial in $z_1,\dots,z_n$ beginning with terms of order two or more.
3. Compute the eigenvalues of $\mathbf{A}$. If all the eigenvalues of $\mathbf{A}$ have negative real part, then $\mathbf{x}(t)\equiv\mathbf{x}^0$ is asymptotically stable. If one eigenvalue of $\mathbf{A}$ has positive real part, then $\mathbf{x}(t)\equiv\mathbf{x}^0$ is unstable.

Example 2. Find all equilibrium solutions of the system of differential equations

$$\frac{dx}{dt}=1-xy,\qquad \frac{dy}{dt}=x-y^3 \tag{12}$$

and determine (if possible) whether they are stable or unstable.

Solution. The equations $1-xy=0$ and $x-y^3=0$ imply that $x=1, y=1$ or $x=-1, y=-1$. Hence, $x(t)\equiv 1, y(t)\equiv 1$, and $x(t)\equiv -1, y(t)\equiv -1$ are the only equilibrium solutions of (12).

(i) $x(t)=1, y(t)=1$: Set $u=x-1$, $v=y-1$. Then,

$$\frac{du}{dt}=\frac{dx}{dt}=1-(1+u)(1+v)=-u-v-uv$$

$$\frac{dv}{dt}=\frac{dy}{dt}=(1+u)-(1+v)^3=u-3v-3v^2-v^3.$$

We rewrite this system in the form

$$\frac{d}{dt}\begin{pmatrix}u\\v\end{pmatrix}=\begin{pmatrix}-1&-1\\1&-3\end{pmatrix}\begin{pmatrix}u\\v\end{pmatrix}-\begin{pmatrix}uv\\3v^2+v^3\end{pmatrix}.$$

The matrix

$$\begin{pmatrix}-1&-1\\1&-3\end{pmatrix}$$

has a single eigenvalue $\lambda=-2$ since

$$\det\begin{pmatrix}-1-\lambda&-1\\1&-3-\lambda\end{pmatrix}=(1+\lambda)(3+\lambda)+1=(\lambda+2)^2.$$

Hence, the equilibrium solution $x(t)\equiv 1, y(t)\equiv 1$ of (12) is asymptotically stable.

(ii) $x(t)\equiv -1, y(t)\equiv -1$: Set $u=x+1$, $v=y+1$. Then,

$$\frac{du}{dt}=\frac{dx}{dt}=1-(u-1)(v-1)=u+v-uv$$

$$\frac{dv}{dt}=\frac{dy}{dt}=(u-1)-(v-1)^3=u-3v+3v^2-v^3.$$

We rewrite this system in the form

$$\frac{d}{dt}\begin{pmatrix}u\\v\end{pmatrix}=\begin{pmatrix}1&1\\1&-3\end{pmatrix}\begin{pmatrix}u\\v\end{pmatrix}+\begin{pmatrix}-uv\\3v^2-v^3\end{pmatrix}.$$

The eigenvalues of the matrix

$$\begin{pmatrix}1&1\\1&-3\end{pmatrix}$$

are $\lambda_1=-1-\sqrt{5}$, which is negative, and $\lambda_2=-1+\sqrt{5}$, which is positive. Therefore, the equilibrium solution $x(t)\equiv-1$, $y(t)\equiv-1$ of (12) is unstable.

Example 3. Find all equilibrium solutions of the system of differential equations

$$\frac{dx}{dt}=\sin(x+y),\qquad \frac{dy}{dt}=e^x-1 \tag{13}$$

and determine whether they are stable or unstable.

Solution. The equilibrium points of (13) are determined by the two equations $\sin(x+y)=0$ and $e^x-1=0$. The second equation implies that $x=0$, while the first equation implies that $x+y=n\pi$, n an integer. Consequently, $x(t)\equiv0$, $y(t)\equiv n\pi$, $n=0,\pm1,\pm2,\ldots$, are the equilibrium solutions of (13). Setting $u=x$, $v=y-n\pi$, gives

$$\frac{du}{dt}=\sin(u+v+n\pi),\qquad \frac{dv}{dt}=e^u-1.$$

Now, $\sin(u+v+n\pi)=\cos n\pi\sin(u+v)=(-1)^n\sin(u+v)$. Therefore,

$$\frac{du}{dt}=(-1)^n\sin(u+v),\qquad \frac{dv}{dt}=e^u-1.$$

Next, observe that

$$\sin(u+v)=u+v-\frac{(u+v)^3}{3!}+\ldots,\qquad e^u-1=u+\frac{u^2}{2!}+\ldots.$$

Hence,

$$\frac{du}{dt}=(-1)^n\left[(u+v)-\frac{(u+v)^3}{3!}+\ldots\right],\qquad \frac{dv}{dt}=u+\frac{u^2}{2!}+\ldots.$$

We rewrite this system in the form

$$\frac{d}{dt}\begin{pmatrix}u\\v\end{pmatrix}=\begin{pmatrix}(-1)^n&(-1)^n\\1&0\end{pmatrix}\begin{pmatrix}u\\v\end{pmatrix}+\begin{matrix}\text{terms of order 2 or}\\\text{higher in } u \text{ and } v.\end{matrix}$$

The eigenvalues of the matrix

$$\begin{pmatrix}(-1)^n&(-1)^n\\1&0\end{pmatrix}$$

are

$$\lambda_1=\frac{(-1)^n-\sqrt{1+4(-1)^n}}{2},\qquad \lambda_2=\frac{(-1)^n+\sqrt{1+4(-1)^n}}{2}.$$

When n is even, $\lambda_1=(1-\sqrt{5})/2$ is negative and $\lambda_2=(1+\sqrt{5})/2$ is positive. Hence, $x(t)\equiv 0, y(t)\equiv n\pi$ is unstable if n is even. When n is odd, both $\lambda_1=(-1-\sqrt{3}\,i)/2$ and $\lambda_2=(-1+\sqrt{3}\,i)/2$ have negative real part. Therefore, the equilibrium solution $x(t)\equiv 0, y(t)\equiv n\pi$ is asymptotically stable if n is odd.

Exercises

Find all equilibrium solutions of each of the following systems of equations and determine, if possible, whether they are stable or unstable.

1. $\dot{x}=x-x^3-xy^2$
$\dot{y}=2y-y^5-yx^4$

2. $\dot{x}=x^2+y^2-1$
$\dot{y}=x^2-y^2$

3. $\dot{x}=x^2+y^2-1$
$\dot{y}=2xy$

4. $\dot{x}=6x-6x^2-2xy$
$\dot{y}=4y-4y^2-2xy$

5. $\dot{x}=\tan(x+y)$
$\dot{y}=x+x^3$

6. $\dot{x}=e^y-x$
$\dot{y}=e^x+y$

Verify that the origin is an equilibrium point of each of the following systems of equations and determine, if possible, whether it is stable or unstable.

7. $\dot{x}=y+3x^2$
$\dot{y}=x-3y^2$

8. $\dot{x}=y+\cos y-1$
$\dot{y}=-\sin x+x^3$

9. $\dot{x}=e^{x+y}-1$
$\dot{y}=\sin(x+y)$

10. $\dot{x}=\ln(1+x+y^2)$
$\dot{y}=-y+x^3$

11. $\dot{x}=\cos y-\sin x-1$
$\dot{y}=x-y-y^2$

12. $\dot{x}=8x-3y+e^y-1$
$\dot{y}=\sin x^2-\ln(1-x-y)$

13. $\dot{x}=-x-y-(x^2+y^2)^{3/2}$
$\dot{y}=x-y+(x^2+y^2)^{3/2}$

14. $\dot{x}=x-y+z^2$
$\dot{y}=y+z-x^2$
$\dot{z}=z-x+y^2$

15. $\dot{x}=e^{x+y+z}-1$
$\dot{y}=\sin(x+y+z)$
$\dot{z}=x-y-z^2$

16. $\dot{x}=\ln(1-z)$
$\dot{y}=\ln(1-x)$
$\dot{z}=\ln(1-y)$

17. $\dot{x}=x-\cos y-z+1$
$\dot{y}=y-\cos z-x+1$
$\dot{z}=z-\cos x-y+1$

18 (a) Find all equilibrium solutions of the system of differential equations

$$\frac{dx}{dt}=gz-hx,\qquad \frac{dy}{dt}=\frac{c}{a+bx}-ky,\qquad \frac{dz}{dt}=ey-fz.$$

(This system is a model for the control of protein synthesis.)

(b) Determine the stability or instability of these solutions if either g, e, or c is zero.

4.4 The phase-plane

In this section we begin our study of the "geometric" theory of differential equations. For simplicity, we will restrict ourselves, for the most part, to the case $n=2$. Our aim is to obtain as complete a description as possible of all solutions of the system of differential equations

$$\frac{dx}{dt}=f(x,y), \qquad \frac{dy}{dt}=g(x,y). \tag{1}$$

To this end, observe that every solution $x=x(t)$, $y=y(t)$ of (1) defines a curve in the three-dimensional space t,x,y. That is to say, the set of all points $(t,x(t),y(t))$ describe a curve in the three-dimensional space t,x,y. For example, the solution $x=\cos t$, $y=\sin t$ of the system of differential equations

$$\frac{dx}{dt}=-y, \qquad \frac{dy}{dt}=x$$

describes a helix (see Figure 1) in (t,x,y) space.

The geometric theory of differential equations begins with the important observation that every solution $x=x(t)$, $y=y(t)$, $t_0\leqslant t\leqslant t_1$, of (1) also defines a curve in the $x-y$ plane. To wit, as t runs from t_0 to t_1, the set of points $(x(t),y(t))$ trace out a curve C in the $x-y$ plane. This curve is called the *orbit*, or *trajectory*, of the solution $x=x(t)$, $y=y(t)$, and the $x-y$ plane is called the *phase-plane* of the solutions of (1). Equivalently, we can think of the orbit of $x(t)$, $y(t)$ as the path that the solution traverses in the $x-y$ plane.

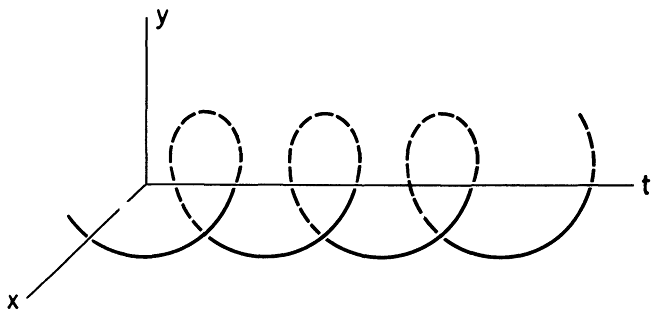

Figure 1. Graph of the solution $x=\cos t$, $y=\sin t$

Example 1. It is easily verified that $x=\cos t$, $y=\sin t$ is a solution of the system of differential equations $\dot{x}=-y$, $\dot{y}=x$. As t runs from 0 to 2π, the set of points $(\cos t,\sin t)$ trace out the unit circle $x^2+y^2=1$ in the $x-y$ plane. Hence, the unit circle $x^2+y^2=1$ is the orbit of the solution $x=\cos t$, $y=\sin t$, $0\leqslant t\leqslant 2\pi$. As t runs from 0 to ∞, the set of points $(\cos t,\sin t)$ trace out this circle infinitely often.

Example 2. It is easily verified that $x=e^{-t}\cos t, y=e^{-t}\sin t$, $-\infty<t<\infty$, is a solution of the system of differential equations $dx/dt=-x-y$, $dy/dt=x-y$. As t runs from $-\infty$ to ∞, the set of points $(e^{-t}\cos t, e^{-t}\sin t)$ trace out a spiral in the $x-y$ plane. Hence, the orbit of the solution $x=e^{-t}\cos t, y=e^{-t}\sin t$ is the spiral shown in Figure 2.

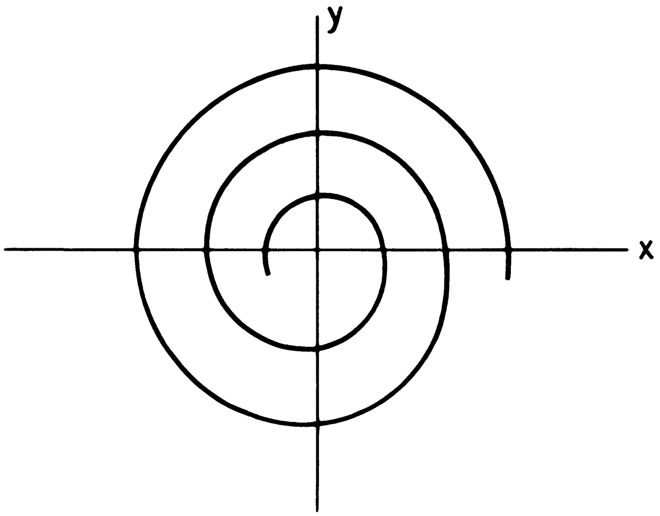

Figure 2. Orbit of $x=e^{-t}\cos t, y=e^{-t}\sin t$

Example 3. It is easily verified that $x=3t+2, y=5t+7$, $-\infty<t<\infty$ is a solution of the system of differential equations $dx/dt=3$, $dy/dt=5$. As t runs from $-\infty$ to ∞, the set of points $(3t+2,5t+7)$ trace out the straight line through the point $(2,7)$ with slope $\frac{5}{3}$. Hence, the orbit of the solution $x=3t+2, y=5t+7$ is the straight line $y=\frac{5}{3}(x-2)+7$, $-\infty<x<\infty$.

Example 4. It is easily verified that $x=3t^2+2, y=5t^2+7$, $0\leqslant t<\infty$ is a solution of the system of differential equations

$$\frac{dx}{dt}=6[(y-7)/5]^{1/2}, \qquad \frac{dy}{dt}=10[(x-2)/3]^{1/2}.$$

All of the points $(3t^2+2,5t^2+7)$ lie on the line through $(2,7)$ with slope $\frac{5}{3}$. However, x is always greater than or equal to 2, and y is always greater than or equal to 7. Hence, the orbit of the solution $x=3t^2+2, y=5t^2+7$, $0\leqslant t<\infty$, is the straight line $y=\frac{5}{3}(x-2)+7$, $2\leqslant x<\infty$.

Example 5. It is easily verified that $x=3t+2, y=5t^2+7$, $-\infty<t<\infty$, is a solution of the system of differential equations

$$\frac{dx}{dt}=y-\frac{5}{9}(x-2)^2-4, \qquad \frac{dy}{dt}=\frac{10}{3}(x-2).$$

The orbit of this solution is the set of all points $(x,y)=(3t+2,5t^2+7)$. Solving for $t=\frac{1}{3}(x-2)$, we see that $y=\frac{5}{9}(x-2)^2+7$. Hence, the orbit of the solution $x=3t+2, y=5t^2+7$ is the parabola $y=\frac{5}{9}(x-2)^2+7$, $|x|<\infty$.

One of the advantages of considering the orbit of the solution rather than the solution itself is that is is often possible to obtain the orbit of a solution without prior knowledge of the solution. Let $x=x(t), y=y(t)$ be a solution of (1). If $x'(t)$ is unequal to zero at $t=t_1$, then we can solve for $t=t(x)$ in a neighborhood of the point $x_1=x(t_1)$ (see Exercise 4). Thus, for t near t_1. the orbit of the solution $x(t), y(t)$ is the curve $y=y(t(x))$. Next, observe that

$$\frac{dy}{dx}=\frac{dy}{dt}\frac{dt}{dx}=\frac{dy/dt}{dx/dt}=\frac{g(x,y)}{f(x,y)}.$$

Thus, the orbits of the solutions $x=x(t)$, $y=y(t)$ of (1) are the solution curves of the first-order scalar equation

$$\frac{dy}{dx}=\frac{g(x,y)}{f(x,y)}. \tag{2}$$

Therefore, it is not necessary to find a solution $x(t), y(t)$ of (1) in order to compute its orbit; we need only solve the single first-order scalar differential equation (2).

Remark. From now on, we will use the phrase "the orbits of (1)" to denote the totality of orbits of solutions of (1).

Example 6. The orbits of the system of differential equations

$$\frac{dx}{dt}=y^2, \qquad \frac{dy}{dt}=x^2 \tag{3}$$

are the solution curves of the scalar equation $dy/dx=x^2/y^2$. This equation is separable, and it is easily seen that every solution is of the form $y(x)=(x^3-c)^{1/3}$, c constant. Thus, the orbits of (3) are the set of all curves $y=(x^3-c)^{1/3}$.

Example 7. The orbits of the system of differential equations

$$\frac{dx}{dt}=y(1+x^2+y^2), \qquad \frac{dy}{dt}=-2x(1+x^2+y^2) \tag{4}$$

are the solution curves of the scalar equation

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

This equation is separable, and all solutions are of the form $\frac{1}{2}y^2+x^2=c^2$. Hence, the orbits of (4) are the families of ellipses $\frac{1}{2}y^2+x^2=c^2$.

Warning. A solution curve of (2) is an orbit of (1) only if dx/dt and dy/dt are not zero simultaneously along the solution. If a solution curve of (2) passes through an equilibrium point of (1), then the entire solution curve is not an orbit. Rather, it is the union of several distinct orbits. For example, consider the system of differential equations

$$\frac{dx}{dt}=y(1-x^2-y^2), \qquad \frac{dy}{dt}=-x(1-x^2-y^2). \tag{5}$$

The solution curves of the scalar equation

$$\frac{dy}{dx}=\frac{dy/dt}{dx/dt}=-\frac{x}{y}$$

are the family of concentric circles $x^2+y^2=c^2$. Observe, however, that every point on the unit circle $x^2+y^2=1$ is an equilibrium point of (5). Thus, the orbits of this system are the circles $x^2+y^2=c^2$, for $c\neq 1$, and all points on the unit circle $x^2+y^2=1$. Similarly, the orbits of (3) are the curves $y=(x^3-c)^{1/3}$, $c\neq 0$; the half-lines $y=x$, $x>0$, and $y=x$, $x<0$; and the point $(0,0)$.

It is not possible, in general, to explicitly solve Equation (2). Hence, we cannot, in general, find the orbits of (1). Nevertheless, it is still possible to obtain an accurate description of all orbits of (1). This is because the system of differential equations (1) sets up a *direction field* in the $x-y$ plane. That is to say, the system of differential equations (1) tells us how fast a solution moves along its orbit, and in what direction it is moving. More precisely, let $x=x(t), y=y(t)$ be a solution of (1). As t increases, the point $(x(t),y(t))$ moves along the orbit of this solution. Its velocity in the x-direction is dx/dt; its velocity in the y-direction is dy/dt; and the magnitude of its velocity is $[(dx(t)/dt)^2+(dy(t)/dt)^2]^{1/2}$. But $dx(t)/dt=f(x(t),y(t))$, and $dy(t)/dt=g(x(t),y(t))$. Hence, at each point (x,y) in the phase plane of (1) we know (i), the tangent to the orbit through (x,y) (the line through (x,y) with direction numbers $f(x,y)$, $g(x,y)$ respectively) and (ii), the speed $[f^2(x,y)+g^2(x,y)]^{1/2}$ with which the solution is traversing its orbit. As we shall see in Sections 4.8–13, this information can often be used to deduce important properties of the orbits of (1).

The notion of orbit can easily be extended to the case $n>2$. Let $\mathbf{x}=\mathbf{x}(t)$ be a solution of the vector differential equation

$$\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x}), \qquad \mathbf{x}=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}, \qquad \mathbf{f}(\mathbf{x})=\begin{bmatrix}f_1(x_1,\dots,x_n)\\ \vdots\\ f_n(x_1,\dots,x_n)\end{bmatrix} \tag{6}$$

on the interval $t_0\leqslant t\leqslant t_1$. As t runs from t_0 to t_1, the set of points $(x_1(t),\dots,x_n(t))$ trace out a curve C in the n-dimensional space $x_1,x_2,\dots,x_n$. This curve is called the orbit of the solution $\mathbf{x}=\mathbf{x}(t)$, for $t_0\leqslant t\leqslant t_1$, and the n-dimensional space $x_1,\dots,x_n$ is called the phase-space of the solutions of (6).

EXERCISES

In each of Problems 1–3, verify that $x(t), y(t)$ is a solution of the given system of equations, and find its orbit.

1. $\dot{x}=1,\quad \dot{y}=2(1-x)\sin(1-x)^2$
   $x(t)=1+t,\quad y(t)=\cos t^2$

2. $\dot{x}=e^{-x},\quad \dot{y}=e^{e^x-1}$
   $x(t)=\ln(1+t),\quad y(t)=e^t$

3. $\dot{x}=1+x^2,\quad \dot{y}=(1+x^2)\sec^2 x$
   $x(t)=\tan t,\quad y(t)=\tan(\tan t)$

4. Suppose that $x'(t_1)\neq 0$. Show that we can solve the equation $x=x(t)$ for $t=t(x)$ in a neighborhood of the point $x_1=x(t_1)$. *Hint*: If $x'(t_1)\neq 0$, then $x(t)$ is a strictly monotonic function of t in a neighborhood of $t=t_1$.

Find the orbits of each of the following systems.

5. $\dot{x}=y,$
   $\dot{y}=-x$

6. $\dot{x}=y(1+x^2+y^2),$
   $\dot{y}=-x(1+x^2+y^2)$

7. $\dot{x}=y(1+x+y),$
   $\dot{y}=-x(1+x+y)$

8. $\dot{x}=y+x^2y,$
   $\dot{y}=3x+xy^2$

9. $\dot{x}=xye^{-3x},$
   $\dot{y}=-2xy^2$

10. $\dot{x}=4y,$
    $\dot{y}=x+xy^2$

11. $\dot{x}=ax-bxy,$
    $\dot{y}=cx-dxy$
    $(a,b,c,d$ positive)

12. $\dot{x}=x^2+\cos y,$
    $\dot{y}=-2xy$

13. $\dot{x}=2xy,$
    $\dot{y}=x^2-y^2$

14. $\dot{x}=y+\sin x,$
    $\dot{y}=x-y\cos x$

4.5 Mathematical theories of war

4.5.1. L. F. Richardson's theory of conflict

In this section we construct a mathematical model which describes the relation between two nations, each determined to defend itself against a possible attack by the other. Each nation considers the possibility of attack quite real, and reasonably enough, bases its apprehensions on the readiness of the other to wage war. Our model is based on the work of Lewis Fry Richardson. It is not an attempt to make scientific statements about foreign politics or to predict the date at which the next war will break out. This, of course, is clearly impossible. Rather, it is a description of what people would do if they did not stop to think. As Richardson writes: "Why are so many nations reluctantly but steadily increasing their armaments as if they were mechanically compelled to do so? Because, I say, they follow their traditions which are fixtures and their instincts which are mechanical;

and because they have not yet made a sufficiently strenuous intellectual and moral effort to control the situation. The process described by the ensuing equations is not to be thought of as inevitable. It is what would occur if instinct and tradition were allowed to act uncontrolled."

Let $x=x(t)$ denote the war potential, or armaments, of the first nation, which we will call Jedesland, and let $y(t)$ denote the war potential of the second nation, which we will call Andersland. The rate of change of $x(t)$ depends, obviously, on the war readiness $y(t)$ of Andersland, and on the grievances that Jedesland feels towards Andersland. In the most simplistic model we represent these terms by ky and g respectively, where k and g are positive constants. These two terms cause x to increase. On the other hand, the cost of armaments has a restraining effect on dx/dt. We represent this term by $-\alpha x$, where α is a positive constant. A similar analysis holds for dy/dt. Consequently, $x=x(t), y=y(t)$ is a solution of the linear system of differential equations

$$\frac{dx}{dt}=ky-\alpha x+g, \qquad \frac{dy}{dt}=lx-\beta y+h. \tag{1}$$

Remark. The model (1) is not limited to two nations; it can also represent the relation between two alliances. For example, Andersland and Jedesland can represent the alliances of France with Russia, and Germany with Austria–Hungary during the years immediately prior to World War I.

Throughout history, there has been a constant debate on the cause of war. Over two thousand years ago, Thucydides claimed that armaments cause war. In his account of the Peloponnesian war he writes: "The real though unavowed cause I believe to have been the growth of Athenian power, which terrified the Lacedaemonians and forced them into war." Sir Edward Grey, the British Foreign Secretary during World War I agrees. He writes: "The increase of armaments that is intended in each nation to produce consciousness of strength, and a sense of security, does not produce these effects. On the contrary, it produces a consciousness of the strength of other nations and a sense of fear. The enormous growth of armaments in Europe, the sense of insecurity and fear caused by them—it was these that made war inevitable. This is the real and final account of the origin of the Great War."

On the other hand, L. S. Amery, a member of Britain's parliament during the 1930's vehemently disagrees. When the opinion of Sir Edward Grey was quoted in the House of Commons, Amery replied: "With all due respect to the memory of an eminent statesman, I believe that statement to be entirely mistaken. The armaments were only the symptoms of the conflict of ambitions and ideals, of those nationalist forces which created the War. The War was brought about because Serbia, Italy and Rumania passionately desired the incorporation in their states of territories which at that time belonged to the Austrian Empire and which the Austrian government was not prepared to abandon without a struggle. France was prepared, if the opportunity ever came, to make an effort to recover Alsace–Lorraine. It was in those facts, in those insoluble conflicts of ambitions, and not in the armaments themselves, that the cause of the War lay."

The system of equations (1) takes both conflicting theories into account. Thucydides and Sir Edward Grey would take g and h small compared to k and l, while Mr. Amery would take k and l small compared to g and h.

The system of equations (1) has several important implications. Suppose that g and h are both zero. Then, $x(t)\equiv 0, y(t)\equiv 0$ is an equilibrium solution of (1). That is, if x, y, g, and h are all made zero simultaneously, then $x(t)$ and $y(t)$ will always remain zero. This ideal condition is permanent peace by disarmament and satisfaction. It has existed since 1817 on the border between Canada and the United States, and since 1905 on the border between Norway and Sweden.

These equations further imply that mutual disarmament without satisfaction is not permanent. Assume that x and y vanish simultaneously at some time $t=t_0$. At this time, $dx/dt=g$ and $dy/dt=h$. Thus, x and y will not remain zero if g and h are positive. Instead, both nations will rearm.

Unilateral disarmament corresponds to setting $y=0$ at a certain instant of time. At this time, $dy/dt=lx+h$. This implies that y will not remain zero if either h or x is positive. Thus, unilateral disarmament is never permanent. This accords with the historical fact that Germany, whose army was reduced by the Treaty of Versailles to 100,000 men, a level far below that of several of her neighbors, insisting on rearming during the years 1933–36.

A race in armaments occurs when the "defense" terms predominate in (1). In this case,

$$\frac{dx}{dt}=ky, \qquad \frac{dy}{dt}=lx. \tag{2}$$

Every solution of (2) is of the form

$$x(t)=Ae^{\sqrt{kl}\,t}+Be^{-\sqrt{kl}\,t}, \qquad y(t)=\sqrt{\frac{l}{k}}\left[Ae^{\sqrt{kl}\,t}-Be^{-\sqrt{kl}\,t}\right].$$

Therefore, both $x(t)$ and $y(t)$ approach infinity if A is positive. This infinity can be interpreted as war.

Now, the system of equations (1) is not quite correct, since it does not take into effect the cooperation, or trade, between Andersland and Jedesland. As we see today, mutual cooperation between nations tends to decrease their fears and suspicions. We correct our model by changing the meaning of $x(t)$ and $y(t)$; we let the variables $x(t)$ and $y(t)$ stand for "threats" minus "cooperation." Specifically, we set $x=U-U_0$ and $y=V-V_0$, where U is the defense budget of Jedesland, V is the defense budget of Andersland, U_0 is the amount of goods exported by Jedesland to

Andersland and V_0 is the amount of goods exported by Andersland to Jedesland. Observe that cooperation evokes reciprocal cooperation, just as armaments provoke more armaments. In addition, nations have a tendency to reduce cooperation on account of the expense which it involves. Thus, the system of equations (1) still describes this more general state of affairs.

The system of equations (1) has a single equilibrium solution

$$x = x_0 = \frac{kh + \beta g}{\alpha\beta - kl}, \qquad y = y_0 = \frac{lg + \alpha h}{\alpha\beta - kl} \tag{3}$$

if $\alpha\beta - kl \neq 0$. We are interested in determining whether this equilibrium solution is stable or unstable. To this end, we write (1) in the form $\dot{\mathbf{w}} = \mathbf{A}\mathbf{w} + \mathbf{f}$, where

$$\mathbf{w}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \qquad \mathbf{f} = \begin{pmatrix} g \\ h \end{pmatrix}, \quad \text{and} \quad \mathbf{A} = \begin{pmatrix} -\alpha & k \\ l & -\beta \end{pmatrix}.$$

The equilibrium solution is

$$\mathbf{w} = \mathbf{w}_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

where $\mathbf{A}\mathbf{w}_0 + \mathbf{f} = \mathbf{0}$. Setting $\mathbf{z} = \mathbf{w} - \mathbf{w}_0$, we obtain that

$$\dot{\mathbf{z}} = \dot{\mathbf{w}} = \mathbf{A}\mathbf{w} + \mathbf{f} = \mathbf{A}(\mathbf{z} + \mathbf{w}_0) + \mathbf{f} = \mathbf{A}\mathbf{z} + \mathbf{A}\mathbf{w}_0 + \mathbf{f} = \mathbf{A}\mathbf{z}.$$

Clearly, the equilibrium solution $\mathbf{w}(t) = \mathbf{w}_0$ of $\dot{\mathbf{w}} = \mathbf{A}\mathbf{w} + \mathbf{f}$ is stable if, and only if, $\mathbf{z} = \mathbf{0}$ is a stable solution of $\dot{\mathbf{z}} = \mathbf{A}\mathbf{z}$. To determine the stability of $\mathbf{z} = \mathbf{0}$ we compute

$$p(\lambda) = \det\begin{pmatrix} -\alpha - \lambda & k \\ l & -\beta - \lambda \end{pmatrix} = \lambda^2 + (\alpha + \beta)\lambda + \alpha\beta - kl.$$

The roots of $p(\lambda)$ are

$$\lambda = \frac{-(\alpha+\beta) \pm \left[(\alpha+\beta)^2 - 4(\alpha\beta - kl)\right]^{1/2}}{2}$$

$$= \frac{-(\alpha+\beta) \pm \left[(\alpha-\beta)^2 + 4kl\right]^{1/2}}{2}.$$

Notice that both roots are real and unequal to zero. Moreover, both roots are negative if $\alpha\beta - kl > 0$, and one root is positive if $\alpha\beta - kl < 0$. Thus, $\mathbf{z}(t) \equiv \mathbf{0}$, and consequently the equilibrium solution $x(t) \equiv x_0$, $y(t) \equiv y_0$ is stable if $\alpha\beta - kl > 0$ and unstable if $\alpha\beta - kl < 0$.

Let us now tackle the difficult problem of estimating the coefficients α, β, k, l, g, and h. There is no way, obviously, of measuring g and h. However, it is possible to obtain reasonable estimates for α, β, k, and l. Observe that the units of these coefficients are reciprocal times. Physicists and engineers would call α^{-1} and β^{-1} relaxation times, for if y and g were identically zero, then $x(t) = e^{-\alpha(t-t_0)}x(t_0)$. This implies that $x(t_0 + \alpha^{-1}) =$

$x(t_0)/e$. Hence, α^{-1} is the time required for Jedesland's armaments to be reduced in the ratio 2.718 if that nation has no grievances and no other nation has any armaments. Richardson estimates α^{-1} to be the lifetime of Jedesland's parliament. Thus, $\alpha=0.2$ for Great Britain, since the lifetime of Britain's parliament is five years.

To estimate k and l we take a hypothetical case in which $g=0$ and $y=y_1$, so that $dx/dt=ky_1-\alpha x$. When $x=0$, $1/k=y_1/(dx/dt)$. Thus, $1/k$ is the time required for Jedesland to catch up to Andersland provided that (i) Andersland's armaments remain constant, (ii) there are no grievances, and (iii) the cost of armaments doesn't slow Jedesland down. Consider now the German rearmament during 1933–36. Germany started with nearly zero armaments and caught up with her neighbors in about three years. Assuming that the slowing effect of α nearly balanced the Germans' very strong grievances g, we take $k=0.3$ (year)$^{-1}$ for Germany. Further, we observe that k is obviously proportional to the amount of industry that a nation has. Thus, $k=0.15$ for a nation which has only half the industrial capacity of Germany, and $k=0.9$ for a nation which has three times the industrial capacity of Germany.

Let us now check our model against the European arms race of 1909–1914. France was allied with Russia, and Germany was allied with Austria–Hungary. Neither Italy or Britain was in a definite alliance with either party. Thus, let Jedesland represent the alliance of France with Russia, and let Andersland represent the alliance of Germany with Austria–Hungary. Since these two alliances were roughly equal in size we take $k=l$, and since each alliance was roughly three times the size of Germany, we take $k=l=0.9$. We also assume that $\alpha=\beta=0.2$. Then,

$$\frac{dx}{dt}=-\alpha x+ky+g,\qquad \frac{dy}{dt}=kx-\alpha y+h. \tag{4}$$

Equation (4) has a unique equilibrium point

$$x_0=\frac{kh+\alpha g}{\alpha^2-k^2},\qquad y_0=\frac{kg+\alpha h}{\alpha^2-k^2}.$$

This equilibrium is unstable since

$$\alpha\beta-kl=\alpha^2-k^2=0.04-0.81=-0.77.$$

This, of course, is in agreement with the historical fact that these two alliances went to war with each other.

Now, the model we have constructed is very crude since it assumes that the grievances g and h are constant in time. This is obviously not true. The grievances g and h are not even continuous functions of time since they jump instantaneously by large amounts. (It's safe to assume, though, that g and h are relatively constant over long periods of time.) In spite of this, the system of equations (4) still provides a very accurate description of the arms race preceding World War I. To demonstrate this, we add the two equations of (4) together, to obtain that

$$\frac{d}{dt}(x+y)=(k-\alpha)(x+y)+g+h. \tag{5}$$

Recall that $x=U-U_0$ and $y=V-V_0$, where U and V are the defense budgets of the two alliances, and U_0 and V_0 are the amount of goods exported from each alliance to the other. Hence,

$$\frac{d}{dt}(U+V)=(k-\alpha)\left\{U+V-\left[U_0+V_0-\frac{g+h}{k-\alpha}-\frac{1}{k-\alpha}\frac{d}{dt}(U_0+V_0)\right]\right\}. \tag{6}$$

The defense budgets for the two alliances are set out in Table I.

Table 1. Defense budgets expressed in millions of £ sterling

	1909		1910		1911		1912		1913
France	48.6		50.9		57.1		63.2		74.7
Russia	66.7		68.5		70.7		81.8		92.0
Germany	63.1		62.0		62.5		68.2		95.4
Austria–Hungary	20.8		23.4		24.6		25.5		26.9
Total $U+V$	199.2		204.8		214.9		238.7		289.0
$\Delta(U+V)$		5.6		10.1		23.8		50.3	
$U+V$ at same date		202.0		209.8		226.8		263.8	

In Figure 1 we plot the annual increment of $U+V$ against the average of $U+V$ for the two years used in forming the increment. Notice how close these four points, denoted by $\circ$, are to the straight line

$$\Delta(U+V)=0.73(U+V-194). \tag{7}$$

Thus, foreign politics does indeed have a machine-like predictability. Equations (6) and (7) imply that

$$g+h=(k-\alpha)(U_0+V_0)-\Delta(U_0+V_0)-194$$

and $k-\alpha=0.73$. This is in excellent agreement with Richardson's estimates of 0.9 for k and 0.2 for α. Finally, observe from (7) that the total defense budgets of the two alliances will increase if $U+V$ is greater than 194 million, and will decrease otherwise. In actual fact, the defense expenditures of the two alliances was 199.2 million in 1909 while the trade between the two alliances amounted to only 171.8 million. Thus began an arms race which led eventually to World War I.

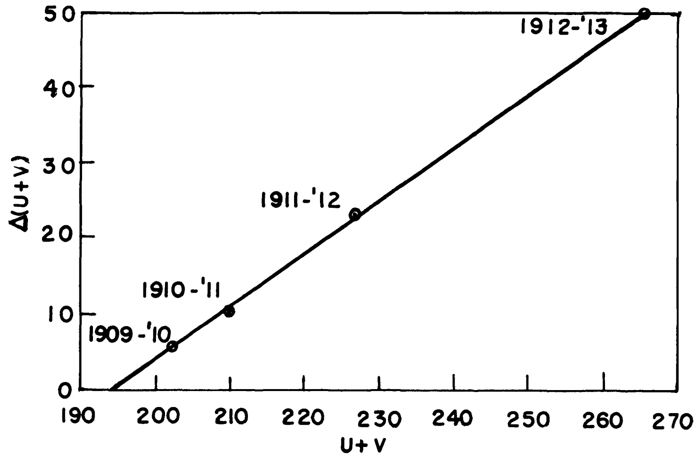

Figure 1. Graph of $\Delta(U+V)$ versus $U+V$

Reference

Richardson, L. F., "Generalized foreign politics," The British Journal of Psychology, monograph supplement #23, 1939.

Exercises

1. Suppose that what moves a government to arm is not the magnitude of other nations' armaments, but the difference between its own and theirs. Then,

$$\frac{dx}{dt}=k(y-x)-\alpha x+g, \qquad \frac{dy}{dt}=l(x-y)-\beta y+h.$$

Show that every solution of this system of equations is stable if $k_1l_1<(\alpha_1+k_1)(\beta_1+l_1)$ and unstable if $k_1l_1>(\alpha_1+k_1)(\beta_1+l_1)$.

2. Consider the case of three nations, each having the same defense coefficient k and the same restraint coefficient α. Then,

$$\frac{dx}{dt}=-\alpha x+ky+kz+g_1$$
$$\frac{dy}{dt}=kx-\alpha y+kz+g_2$$
$$\frac{dz}{dt}=kx+ky-\alpha z+g_3.$$

Setting

$$\mathbf{u}=\begin{pmatrix}x\\y\\z\end{pmatrix}, \quad \mathbf{A}=\begin{pmatrix}-\alpha & k & k\\ k & -\alpha & k\\ k & k & -\alpha\end{pmatrix}, \quad \text{and} \quad \mathbf{g}=\begin{pmatrix}g_1\\g_2\\g_3\end{pmatrix},$$

we see that $\dot{\mathbf{u}}=\mathbf{A}\mathbf{u}+\mathbf{g}$.

(a) Show that $p(\lambda)\equiv\det(\mathbf{A}-\lambda\mathbf{I})=-(\alpha+\lambda)^3+3k^2(\alpha+\lambda)+2k^3$.

(b) Show that $p(\lambda)=0$ when $\lambda=-\alpha-k$. Use this information to find the remaining two roots of $p(\lambda)$.

(c) Show that every solution $\mathbf{u}(t)$ is stable if $2k<\alpha$, and unstable if $2k>\alpha$.

3. Suppose in Problem 2 that the z nation is a pacifist nation, while x and y are pugnacious nations. Then,

$$\begin{aligned} \frac{dx}{dt} &= -\alpha x + ky + kz + g_1 \\ \frac{dy}{dt} &= kx - \alpha y + kz + g_2 \\ \frac{dz}{dt} &= 0\cdot x + 0\cdot y - \alpha z + g_3. \end{aligned} \tag{*}$$

Show that every solution $x(t), y(t), z(t)$ of (*) is stable if $k<\alpha$, and unstable if $k>\alpha$.

4.5.2 *Lanchester's combat models and the battle of Iwo Jima*

During the first World War, F. W. Lanchester [4] pointed out the importance of the concentration of troops in modern combat. He constructed mathematical models from which the expected results of an engagement could be obtained. In this section we will derive two of these models, that of a conventional force versus a conventional force, and that of a conventional force versus a guerilla force. We will then solve these models, or equations, and derive "Lanchester's square law," which states that the *strength* of a combat force is proportional to the square of the number of combatants entering the engagement. Finally, we will fit one of these models, with astonishing accuracy, to the battle of Iwo Jima in World War II.

(a) *Construction of the models*

Suppose that an "x-force" and a "y-force" are engaged in combat. For simplicity, we define the strengths of these two forces as their number of combatants. (See Howes and Thrall [3] for another definition of combat strength.) Thus let $x(t)$ and $y(t)$ denote the number of combatants of the x and y forces, where t is measured in days from the start of the combat. Clearly, the rate of change of each of these quantities equals its *reinforcement rate* minus its *operational loss rate* minus its *combat loss rate*.

The operational loss rate. The operational loss rate of a combat force is its loss rate due to non-combat mishaps; i.e., desertions, diseases, etc. Lanchester proposed that the operational loss rate of a combat force is proportional to its strength. However, this does not appear to be very realistic. For example, the desertion rate in a combat force depends on a host of psychological and other intangible factors which are difficult even to describe, let alone quantify. We will take the easy way out here and consider only those engagements in which the operational loss rates are negligible.

The combat loss rate. Suppose that the x-force is a conventional force which operates in the open, comparatively speaking, and that every member of this force is within "kill range" of the enemy y. We also assume that as soon as the conventional force suffers a loss, fire is concentrated on the remaining combatants. Under these "ideal" conditions, the combat loss rate of a conventional force x equals $ay(t)$, for some positive constant a. This constant is called the *combat effectiveness coefficient* of the y-force.

The situation is very different if x is a guerilla force, invisible to its opponent y and occupying a region R. The y-force fires into R but cannot know when a kill has been made. It is certainly plausible that the combat loss rate for a guerilla force x should be proportional to $x(t)$, for the larger $x(t)$, the greater the probability that an opponent's shot will kill. On the other hand, the combat loss rate for x is also proportional to $y(t)$, for the larger y, the greater the number of x-casualties. Thus, the combat loss rate for a guerilla force x equals $cx(t)y(t)$, where the constant c is called the *combat effectiveness coefficient* of the opponent y.

The reinforcement rate. The reinforcement rate of a combat force is the rate at which new combatants enter (or are withdrawn from) the battle. We denote the reinforcement rates of the x- and y-forces by $f(t)$ and $g(t)$ respectively.

Under the assumptions listed above, we can now write down the following two Lanchestrian models for conventional–guerilla combat.

Conventional combat:
$$\begin{cases} \dfrac{dx}{dt} = -ay+f(t) \\ \dfrac{dy}{dt} = -bx+g(t) \end{cases} \tag{1a}$$

Conventional–guerilla combat: (x = guerilla)
$$\begin{cases} \dfrac{dx}{dt} = -cxy+f(t) \\ \dfrac{dy}{dt} = -dx+g(t) \end{cases} \tag{1b}$$

The system of equations (1a) is a linear system and can be solved explicitly once a, b, $f(t)$, and $g(t)$ are known. On the other hand, the system of equations (1b) is nonlinear, and its solution is much more difficult. (Indeed, it can only be obtained with the aid of a digital computer.)

It is very instructive to consider the special case where the reinforcement rates are zero. This situation occurs when the two forces are "isolated." In this case (1a) and (1b) reduce to the simpler systems

$$\frac{dx}{dt} = -ay, \qquad \frac{dy}{dt} = -bx \tag{2a}$$

and

$$\frac{dx}{dt} = -cxy, \qquad \frac{dy}{dt} = -dx. \tag{2b}$$

Conventional combat: *The square law*. The orbits of system (2a) are the solution curves of the equation

$$\frac{dy}{dx}=\frac{bx}{ay} \quad \text{or} \quad ay\frac{dy}{dx}=bx.$$

Integrating this equation gives

$$ay^2-bx^2=ay_0^2-bx_0^2=K. \tag{3}$$

The curves (3) define a family of hyperbolas in the x–y plane and we have indicated their graphs in Figure 1. The arrowheads on the curves indicate the direction of changing strengths as time passes.

Let us adopt the criterion that one force wins the battle if the other force vanishes first. Then, y wins if $K>0$ since the x-force has been annihilated by the time $y(t)$ has decreased to $\sqrt{K/a}$. Similarly, x wins if $K<0$.

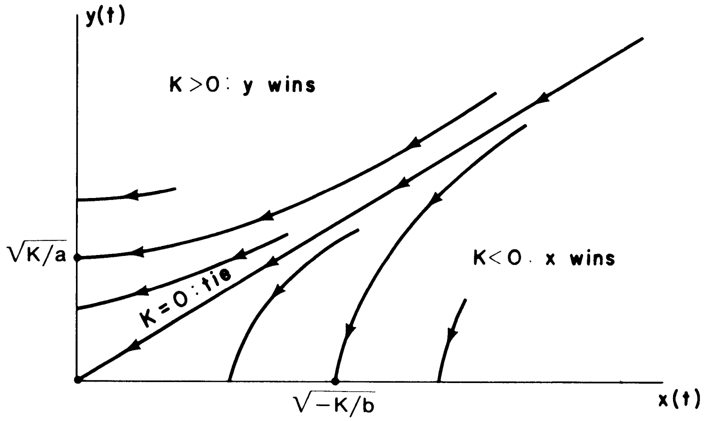

Figure 1. The hyperbolas defined by (3)

Remark 1. Equation (3) is often referred to as "Lanchester's square law," and the system (2a) is often called the square law model, since the strengths of the opposing forces appear *quadratically* in (3). This terminology is rather anomolous since the system (2a) is actually a linear system.

Remark 2. The y-force always seeks to establish a setting in which $K>0$. That is to say, the y-force wants the inequality

$$ay_0^2>bx_0^2$$

to hold. This can be accomplished by increasing a; i.e. by using stronger and more accurate weapons, or by increasing the initial force y_0. Notice though that a doubling of a results in a doubling of ay_0^2 while a doubling of y_0 results in a *four-fold* increase of ay_0^2. This is the essence of Lanchester's square law of conventional combat.

Conventional–guerilla combat. The orbits of system (2b) are the solution curves of the equation

$$\frac{dy}{dx}=\frac{dx}{cxy}=\frac{d}{cy}. \tag{4}$$

Multiplying both sides of (4) by cy and integrating gives

$$cy^2-2dx=cy_0^2-2dx_0=M. \tag{5}$$

The curves (5) define a family of parabolas in the x–y plane, and we have indicated their graphs in Figure 2. The y-force wins if $M>0$, since the x-force has been annihilated by the time $y(t)$ has decreased to $\sqrt{M/c}$. Similarly, x wins if $M<0$.

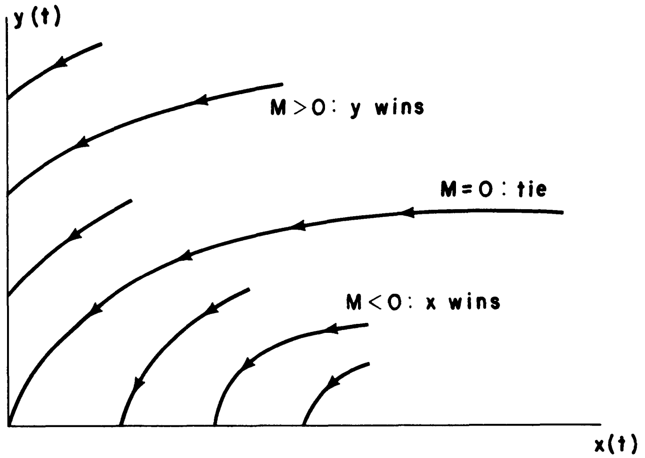

Figure 2. The parabolas defined by (5)

Remark. It is usually impossible to determine, a priori, the numerical value of the combat coefficients a, b, c, and d. Thus, it would appear that Lanchester's combat models have little or no applicability to real-life engagements. However, this is not so. As we shall soon see, it is often possible to determine suitable values of a and b (or c and d) using data from the battle itself. Once these values are established for one engagement, they are known for all other engagements which are fought under similar conditions.

(b) *The battle of Iwo Jima*

One of the fiercest battles of World War II was fought on the island of Iwo Jima, 660 miles south of Tokyo. Our forces coveted Iwo Jima as a bomber base close to the Japanese mainland, while the Japanese needed the island as a base for fighter planes attacking our aircraft on their way to bombing missions over Tokyo and other major Japanese cities. The American invasion of Iwo Jima began on February 19, 1945, and the fighting was intense throughout the month long combat. Both sides suffered heavy casualties (see Table 1). The Japanese had been ordered to fight to the last man, and this is exactly what they did. The island was declared "secure" by the American forces on the 28th day of the battle, and all active combat ceased on the 36th day. (The last two Japanese holdouts surrendered in 1951!)

Table 1. Casualties at Iwo Jima

Total United States casualties at Iwo Jima				
	Killed, missing or died of wounds	Wounded	Combat Fatigue	Total
Marines	5,931	17,272	2,648	25,851
Navy units:				
Ships and air units	633	1,158		1,791
Medical corpsmen	195	529		724
Seabees	51	218		269
Doctors and dentists	2	12		14
Army units in battle	9	28		37
Grand totals	6,821	19,217	2,648	28,686
Japanese casualties at Iwo Jima				
Defense forces (Estimated)		Prisoners		Killed
21,000		Marine 216		20,000
		Army 867		
		Total 1,083		

(Newcomb [6], page 296)

The following data is available to us from the battle of Iwo Jima.

1. *Reinforcement rates*. During the conflict Japanese troops were neither withdrawn nor reinforced. The Americans, on the other hand, landed 54,000 troops on the first day of the battle, none on the second, 6,000 on the third, none on the fourth and fifth, 13,000 on the sixth day, and none thereafter. There were no American troops on Iwo Jima prior to the start of the engagement.

2. *Combat losses*. Captain Clifford Morehouse of the United States Marine Corps (see Morehouse [5]) kept a daily count of all American combat losses. Unfortunately, no such records are available for the Japanese forces. Most probably, the casualty lists kept by General Kuribayashi (commander of the Japanese forces on Iwo Jima) were destroyed in the battle itself, while whatever records were kept in Tokyo were consumed in the fire bombings of the remaining five months of the war. However, we can infer from Table 1 that approximately 21,500 Japanese forces were on Iwo Jima at the start of the battle. (Actually, Newcomb arrived at the fig-

ure of 21,000 for the Japanese forces, but this is a little low since he apparently did not include some of the living and dead found in the caves in the final days.)

3. *Operational losses*. The operational losses on both sides were negligible.

Now, let $x(t)$ and $y(t)$ denote respectively, the active American and Japanese forces on Iwo Jima t days after the battle began. The data above suggests the following Lanchestrian model for the battle of Iwo Jima:

$$\begin{aligned}\frac{dx}{dt}&=-ay+f(t)\\ \frac{dy}{dt}&=-bx\end{aligned}\tag{6}$$

where a and b are the combat effectiveness coefficients of the Japanese and American forces, respectively, and

$$f(t)=\begin{cases}54{,}000 & 0\leqslant t<1\\ 0 & 1\leqslant t<2\\ 6{,}000 & 2\leqslant t<3\\ 0 & 3\leqslant t<5\\ 13{,}000 & 5\leqslant t<6\\ 0 & t\geqslant 6\end{cases}$$

Using the method of variation of parameters developed in Section 3.12 or the method of elimination in Section 2.14, it is easily seen that the solution of (6) which satisfies $x(0)=0$, $y(0)=y_0=21{,}500$ is given by

$$x(t)=-\sqrt{\frac{a}{b}}\;y_0\cosh\sqrt{ab}\;t+\int_0^t\cosh\sqrt{ab}\;(t-s)f(s)\,ds\tag{7a}$$

and

$$y(t)=y_0\cosh\sqrt{ab}\;t-\sqrt{\frac{b}{a}}\;\int_0^t\sinh\sqrt{ab}\;(t-s)f(s)\,ds\tag{7b}$$

where

$$\cosh x\equiv(e^x+e^{-x})/2\quad\text{and}\quad\sinh x\equiv(e^x-e^{-x})/2.$$

The problem before us now is this: Do there exist constants a and b so that (7a) yields a good fit to the data compiled by Morehouse? This is an extremely important question. An affirmative answer would indicate that Lanchestrian models do indeed describe real life battles, while a negative answer would shed a dubious light on much of Lanchester's work.

As we mentioned previously, it is extremely difficult to compute the combat effectiveness coefficients a and b of two opposing forces. However, it is often possible to determine suitable values of a and b once the data for the battle is known, and such is the case for the battle of Iwo Jima.

The calculation of a and b. Integrating the second equation of (6) between 0 and s gives

$$y(s)-y_0=-b\int_0^s x(t)\,dt$$

so that

$$b=\frac{y_0-y(s)}{\int_0^s x(t)\,dt}. \tag{8}$$

In particular, setting $s=36$ gives

$$b=\frac{y_0-y(36)}{\int_0^{36} x(t)\,dt}=\frac{21,500}{\int_0^{36} x(t)\,dt}. \tag{9}$$

Now the integral on the right-hand side of (9) can be approximated by the Riemann sum

$$\int_0^{36} x(t)\,dt \cong \sum_{i=1}^{36} x(i)$$

and for $x(i)$ we enter the number of effective American troops on the ith day of the battle. Using the data available from Morehouse, we compute for b the value

$$b=\frac{21,500}{2,037,000}=0.0106. \tag{10}$$

Remark. We would prefer to set $s=28$ in (8) since that was the day the island was declared secure, and the fighting was only sporadic after this day. However, we don't know $y(28)$. Thus, we are forced here to take $s=36$.

Next, we integrate the first equation of (6) between $t=0$ and $t=28$ and obtain that

$$\begin{aligned} x(28)&=-a\int_0^{28} y(t)\,dt+\int_0^{28} f(t)\,dt \\ &=-a\int_0^{28} y(t)\,dt+73,000. \end{aligned}$$

There were 52,735 effective American troops on the 28th day of the battle. Thus

$$a=\frac{73,000-52,735}{\int_0^{28} y(t)\,dt}=\frac{20,265}{\int_0^{28} y(t)\,dt}. \tag{11}$$

Finally, we approximate the integral on the right-hand side of (11) by the Riemann sum

$$\int_0^{28} y(t)\,dt \cong \sum_{j=1}^{28} y(j)$$

and we approximate $y(j)$ by

$$y(j) = y_0 - b\int_0^j x(t)\,dt$$

$$\cong 21{,}500 - b\sum_{i=i}^{j} x(i).$$

Again, we replace $x(i)$ by the number of effective American troops on the ith day of the battle. The result of this calculation is (see Engel [2])

$$a = \frac{20{,}265}{372{,}500} = 0.0544. \tag{12}$$

Figure 3 below compares the actual American troop strength with the values predicted by Equation (7a) (with $a = 0.0544$ and $b = 0.0106$). The fit is remarkably good. Thus, it appears that a Lanchestrian model does indeed describe real life engagements.

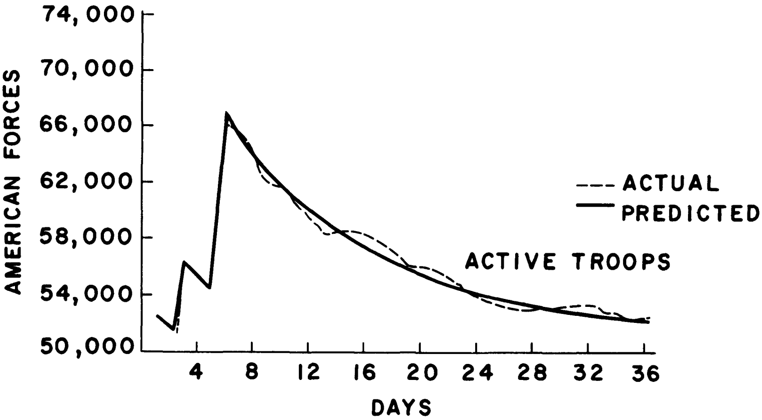

Figure 3. Comparison of actual troop strength with predicted troop strength

Remark. The figures we have used for American reinforcements include *all* the personnel put ashore, both combat troops and support troops. Thus the numbers a and b that we have computed should be interpreted as the *average* effectiveness per man ashore.

References

1. Coleman, C. S., Combat Models, MAA Workshop on Modules in Applied Math, Cornell University, Aug. 1976.
2. Engel, J. H., A verification of Lanchester's law, *Operations Research*, **2**, (1954), 163–171.
3. Howes, D. R., and Thrall, R. M., A theory of ideal linear weights for heterogeneous combat forces, *Naval Research Logistics Quarterly*, vol. 20, 1973, pp. 645–659.
4. Lanchester, F. W., *Aircraft in Warfare, the Dawn of the Fourth Arm*. Tiptree, Constable and Co., Ltd., 1916.
5. Morehouse, C. P., *The Iwo Jima Operation*, USMCR, Historical Division, Hdqr. USMC, 1946.
6. Newcomb, R. F., *Iwo Jima*. New York: Holt, Rinehart, and Winston, 1965.

Exercises

1. Derive Equations (7a) and (7b).

2. The system of equations

$$\begin{aligned} \dot{x} &= -ay \\ \dot{y} &= -by - cxy \end{aligned} \tag{13}$$

is a Lanchestrian model for conventional–guerilla combat, in which the operational loss rate of the guerilla force y is proportional to $y(t)$.
(a) Find the orbits of (13).
(b) Who wins the battle?

3. The system of equations

$$\begin{aligned} \dot{x} &= -ay \\ \dot{y} &= -bx - cxy \end{aligned} \tag{14}$$

is a Lanchestrian model for conventional–guerilla combat, in which the operational loss rate of the guerilla force y is proportional to the strength of the conventional force x. Find the orbits of (14).

4. The system of equations

$$\begin{aligned} \dot{x} &= -cxy \\ \dot{y} &= -dxy \end{aligned} \tag{15}$$

is a Lanchestrian model for guerilla–guerilla combat in which the operational loss rates are negligible.
(a) Find the orbits of (15).
(b) Who wins the battle?

5. The system of equations

$$\begin{aligned} \dot{x} &= -ay - cxy \\ \dot{y} &= -bx - dxy \end{aligned} \tag{16}$$

is a Lanchestrian model for guerilla–guerilla combat in which the operational loss rate of each force is proportional to the strength of its opponent. Find the orbits of (16).

6. The system of equations

$$\begin{aligned}\dot{x} &= -ax - cxy \\ \dot{y} &= -by - dxy\end{aligned} \tag{17}$$

is a Lanchestrian model for guerilla–guerilla combat in which the operational loss rate of each force is proportional to its strength.

(a) Find the orbits of (17).

(b) Show that the x and y axes are both orbits of (17).

(c) Using the fact (to be proved in Section 4.6) that two orbits of (17) cannot intersect, show that there is no clear-cut winner in this battle. *Hint*: Show that $x(t)$ and $y(t)$ can never become zero in finite time. (Using lemmas 1 and 2 of Section 4.8, it is easy to show that both $x(t)$ and $y(t)$ approach zero as $t \to \infty$.)

4.6 Qualitative properties of orbits

In this section we will derive two very important properties of the solutions and orbits of the system of differential equations

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(x_1,\dots,x_n) \\ \vdots \\ f_n(x_1,\dots,x_n) \end{bmatrix}. \tag{1}$$

The first property deals with the existence and uniqueness of orbits, and the second property deals with the existence of periodic solutions of (1). We begin with the following existence–uniqueness theorem for the solutions of (1).

Theorem 3. *Let each of the functions* $f_1(x_1,\dots,x_n),\dots,f_n(x_1,\dots,x_n)$ *have continuous partial derivatives with respect to* $x_1,\dots,x_n$. *Then, the initial-value problem* $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, $\mathbf{x}(t_0) = \mathbf{x}^0$ *has one, and only one solution* $\mathbf{x} = \mathbf{x}(t)$, *for every* $\mathbf{x}^0$ *in* R^n.

We prove Theorem 3 in exactly the same manner as we proved the existence–uniqueness theorem for the scalar differential equation $\dot{x} = f(t,x)$. Indeed, the proof given in Section 1.10 carries over here word for word. We need only interpret the quantity $|\mathbf{x}(t) - \mathbf{y}(t)|$, where $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are vector-valued functions, as the length of the vector $\mathbf{x}(t) - \mathbf{y}(t)$. That is to say, if we interpret $|\mathbf{x}(t) - \mathbf{y}(t)|$ as

$$|\mathbf{x}(t) - \mathbf{y}(t)| = \max\{|x_1(t) - y_1(t)|,\dots,|x_n(t) - y_n(t)|\},$$

then the proof of Theorem 2, Section 1.10 is valid even for vector-valued functions $\mathbf{f}(t,\mathbf{x})$ (see Exercises 13–14).

Next, we require the following simple but extremely useful lemma.

Lemma 1. *If* $\mathbf{x} = \boldsymbol{\phi}(t)$ *is a solution of* (1), *then* $\mathbf{x} = \boldsymbol{\phi}(t + c)$ *is again a solution of* (1).

The meaning of Lemma 1 is the following. Let $\mathbf{x}=\boldsymbol{\phi}(t)$ be a solution of (1) and let us replace every t in the formula for $\boldsymbol{\phi}(t)$ by $t+c$. In this manner we obtain a new function $\hat{\mathbf{x}}(t)=\boldsymbol{\phi}(t+c)$. Lemma 1 states that $\hat{\mathbf{x}}(t)$ is again a solution of (1). For example, $x_1=\tan t$, $x_2=\sec^2 t$ is a solution of the system of differential equations $dx_1/dt=x_2$, $dx_2/dt=2x_1x_2$. Hence, $x_1=\tan(t+c)$, $x_2=\sec^2(t+c)$ is again a solution, for any constant c.

PROOF OF LEMMA 1. If $\mathbf{x}=\boldsymbol{\phi}(t)$ is a solution of (1), then $d\boldsymbol{\phi}(t)/dt=\mathbf{f}(\boldsymbol{\phi}(t))$; that is, the two functions $d\boldsymbol{\phi}(t)/dt$ and $\mathbf{h}(t)\equiv\mathbf{f}(\boldsymbol{\phi}(t))$ agree at every single time. Fix a time t and a constant c. Since $d\boldsymbol{\phi}/dt$ and $\mathbf{h}$ agree at every time, they must agree at time $t+c$. Hence,

$$\frac{d\boldsymbol{\phi}}{dt}(t+c)=\mathbf{h}(t+c)=\mathbf{f}(\boldsymbol{\phi}(t+c)).$$

But, $d\boldsymbol{\phi}/dt$ evaluated at $t+c$ equals the derivative of $\hat{\mathbf{x}}(t)\equiv\boldsymbol{\phi}(t+c)$, evaluated at t. Therefore,

$$\frac{d}{dt}\boldsymbol{\phi}(t+c)=\mathbf{f}(\boldsymbol{\phi}(t+c)).$$ □

Remark 1. Lemma 1 can be verified explicitly for the linear equation $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$. Every solution $\mathbf{x}(t)$ of this equation is of the form $\mathbf{x}(t)=e^{\mathbf{A}t}\mathbf{v}$, for some constant vector $\mathbf{v}$. Hence,

$$\mathbf{x}(t+c)=e^{\mathbf{A}(t+c)}\mathbf{v}=e^{\mathbf{A}t}e^{\mathbf{A}c}\mathbf{v}$$

since $(\mathbf{A}t)\mathbf{A}c=\mathbf{A}c(\mathbf{A}t)$ for all values of t and c. Therefore, $\mathbf{x}(t+c)$ is again a solution of $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$ since it is of the form $e^{\mathbf{A}t}$ times the constant vector $e^{\mathbf{A}c}\mathbf{v}$.

Remark 2. Lemma 1 is not true if the function $\mathbf{f}$ in (1) depends explicitly on t. To see this, suppose that $\mathbf{x}=\boldsymbol{\phi}(t)$ is a solution of the nonautonomous differential equation $\dot{\mathbf{x}}=\mathbf{f}(t,\mathbf{x})$. Then, $\dot{\boldsymbol{\phi}}(t+c)=\mathbf{f}(t+c,\boldsymbol{\phi}(t+c))$. Consequently, the function $\mathbf{x}=\boldsymbol{\phi}(t+c)$ satisfies the differential equation

$$\dot{\mathbf{x}}=\mathbf{f}(t+c,\mathbf{x}),$$

and this equation is different from (1) if $\mathbf{f}$ depends explicitly on t.

We are now in a position to derive the following extremely important properties of the solutions and orbits of (1).

Property 1. (Existence and uniqueness of orbits.) Let each of the functions $f_1(x_1,\ldots,x_n),\ldots,f_n(x_1,\ldots,x_n)$ have continuous partial derivatives with respect to $x_1,\ldots,x_n$. Then, there exists one, and only one, orbit through every point $\mathbf{x}^0$ in R^n. In particular, if the orbits of two solutions $\mathbf{x}=\boldsymbol{\phi}(t)$ and $\mathbf{x}=\boldsymbol{\psi}(t)$ of (1) have one point in common, then they must be identical.

Property 2. Let $\mathbf{x}=\boldsymbol{\phi}(t)$ be a solution of (1). If $\boldsymbol{\phi}(t_0+T)=\boldsymbol{\phi}(t_0)$ for some t_0 and $T>0$, then $\boldsymbol{\phi}(t+T)$ is identically equal to $\boldsymbol{\phi}(t)$. In other words, if a solution $\mathbf{x}(t)$ of (1) returns to its starting value after a time $T>0$, then it must be periodic, with period T (i.e. it must repeat itself over every time interval of length T.)

PROOF OF PROPERTY 1. Let $\mathbf{x}^0$ be any point in the n-dimensional phase space $x_1,\ldots,x_n$, and let $\mathbf{x}=\boldsymbol{\phi}(t)$ be the solution of the initial-value problem $\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x})$, $\mathbf{x}(0)=\mathbf{x}^0$. The orbit of this solution obviously passes through $\mathbf{x}^0$. Hence, there exists at least one orbit through every point $\mathbf{x}^0$. Now, suppose that the orbit of another solution $\mathbf{x}=\boldsymbol{\psi}(t)$ also passes through $\mathbf{x}^0$. This means that there exists $t_0(\neq 0)$ such that $\boldsymbol{\psi}(t_0)=\mathbf{x}^0$. By Lemma 1,

$$\mathbf{x}=\boldsymbol{\psi}(t+t_0)$$

is also a solution of (1). Observe that $\boldsymbol{\psi}(t+t_0)$ and $\boldsymbol{\phi}(t)$ have the same value at $t=0$. Hence, by Theorem 3, $\boldsymbol{\psi}(t+t_0)$ equals $\boldsymbol{\phi}(t)$ for all time t. This implies that the orbits of $\boldsymbol{\phi}(t)$ and $\boldsymbol{\psi}(t)$ are identical. To wit, if $\boldsymbol{\xi}$ is a point on the orbit of $\boldsymbol{\phi}(t)$; that is, $\boldsymbol{\xi}=\boldsymbol{\phi}(t_1)$ for some t_1, then $\boldsymbol{\xi}$ is also on the orbit of $\boldsymbol{\psi}(t)$, since $\boldsymbol{\xi}=\boldsymbol{\phi}(t_1)=\boldsymbol{\psi}(t_1+t_0)$. Conversely, if $\boldsymbol{\xi}$ is a point on the orbit of $\boldsymbol{\psi}(t)$; that is, there exists t_2 such that $\boldsymbol{\psi}(t_2)=\boldsymbol{\xi}$, then $\boldsymbol{\xi}$ is also on the orbit of $\boldsymbol{\phi}(t)$ since $\boldsymbol{\xi}=\boldsymbol{\psi}(t_2)=\boldsymbol{\phi}(t_2-t_0)$. □

PROOF OF PROPERTY 2. Let $\mathbf{x}=\boldsymbol{\phi}(t)$ be a solution of (1) and suppose that $\boldsymbol{\phi}(t_0+T)=\boldsymbol{\phi}(t_0)$ for some numbers t_0 and T. Then, the function $\boldsymbol{\psi}(t)=\boldsymbol{\phi}(t+T)$ is also a solution of (1) which agrees with $\boldsymbol{\phi}(t)$ at time $t=t_0$. By Theorem 3, therefore, $\boldsymbol{\psi}(t)=\boldsymbol{\phi}(t+T)$ is identically equal to $\boldsymbol{\phi}(t)$. □

Property 2 is extremely useful in applications, especially when $n=2$. Let $x=x(t)$, $y=y(t)$ be a periodic solution of the system of differential equations

$$\frac{dx}{dt}=f(x,y), \qquad \frac{dy}{dt}=g(x,y). \tag{2}$$

If $x(t+T)=x(t)$ and $y(t+T)=y(t)$, then the orbit of this solution is a closed curve C in the x–y plane. In every time interval $t_0\leqslant t\leqslant t_0+T$, the solution moves once around C. Conversely, suppose that the orbit of a solution $x=x(t)$, $y=y(t)$ of (2) is a closed curve containing no equilibrium points of (2). Then, the solution $x=x(t)$, $y=y(t)$ is periodic. To prove this, recall that a solution $x=x(t)$, $y=y(t)$ of (2) moves along its orbit with velocity $[f^2(x,y)+g^2(x,y)]^{1/2}$. If its orbit C is a closed curve containing no equilibrium points of (2), then the function $[f^2(x,y)+g^2(x,y)]^{1/2}$ has a positive minimum for (x,y) on C. Hence, the orbit of $x=x(t)$, $y=y(t)$ must return to its starting point $x_0=x(t_0)$, $y_0=y(t_0)$ in some finite time T. But this implies that $x(t+T)=x(t)$ and $y(t+T)=y(t)$ for all t.

Example 1. Prove that every solution $z(t)$ of the second-order differential equation $(d^2z/dt^2)+z+z^3=0$ is periodic.

PROOF. We convert this second-order equation into a system of two first-order equations by setting $x=z$, $y=dz/dt$. Then,

$$\frac{dx}{dt}=y, \qquad \frac{dy}{dt}=-x-x^3. \tag{3}$$

The orbits of (3) are the solution curves

$$\frac{y^2}{2}+\frac{x^2}{2}+\frac{x^4}{4}=c^2 \tag{4}$$

of the scalar equation $dy/dx=-(x+x^3)/y$. Equation (4) defines a closed curve in the x–y plane (see Exercise 7). Moreover, the only equilibrium point of (3) is $x=0, y=0$. Consequently, every solution $x=z(t), y=z'(t)$ of (3) is a periodic function of time. Notice, however, that we cannot compute the period of any particular solution. □

Example 2. Prove that every solution of the system of differential equations

$$\frac{dx}{dt}=ye^{1+x^2+y^2}, \qquad \frac{dy}{dt}=-xe^{1+x^2+y^2} \tag{5}$$

is periodic.

Solution. The orbits of (5) are the solution curves $x^2+y^2=c^2$ of the first-order scalar equation $dy/dx=-x/y$. Moreover, $x=0$, $y=0$ is the only equilibrium point of (5). Consequently, every solution $x=x(t), y=y(t)$ of (5) is a periodic function of time.

EXERCISES

1. Show that all solutions $x(t), y(t)$ of

$$\frac{dx}{dt}=x^2+y\sin x, \qquad \frac{dy}{dt}=-1+xy+\cos y$$

which start in the first quadrant ($x>0$, $y>0$) must remain there for all time (both backwards and forwards).

2. Show that all solutions $x(t), y(t)$ of

$$\frac{dx}{dt}=y(e^x-1), \qquad \frac{dy}{dt}=x+e^y$$

which start in the right half plane ($x>0$) must remain there for all time.

3. Show that all solutions $x(t), y(t)$ of

$$\frac{dx}{dt}=1+x^2+y^2, \qquad \frac{dy}{dt}=xy+\tan y$$

which start in the upper half plane ($y>0$) must remain there for all time.

4. Show that all solutions $x(t), y(t)$ of

$$\frac{dx}{dt}=-1-y+x^2, \qquad \frac{dy}{dt}=x+xy$$

which start inside the unit circle $x^2+y^2=1$ must remain there for all time. *Hint*: Compute $d(x^2+y^2)/dt$.

5. Let $x(t), y(t)$ be a solution of

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

with $x(t_0) \neq y(t_0)$. Show that $x(t)$ can never equal $y(t)$.

6. Can a figure 8 ever be an orbit of

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

where f and g have continuous partial derivatives with respect to x and y?

7. Show that the curve $y^2 + x^2 + x^4/2 = 2c^2$ is closed. *Hint*: Show that there exist two points $y = 0$, $x = \pm\alpha$ which lie on this curve.

Prove that all solutions of the following second-order equations are periodic.

8. $\dfrac{d^2z}{dt^2} + z^3 = 0$

9. $\dfrac{d^2z}{dt^2} + z + z^5 = 0$

10. $\dfrac{d^2z}{dt^2} + e^{z^2} = 1$

11. $\dfrac{d^2z}{dt^2} + \dfrac{z}{1+z^2} = 0$

12. Show that all solutions $z(t)$ of

$$\frac{d^2z}{dt^2} + z - 2z^3 = 0$$

are periodic if $\dot{z}^2(0) + z^2(0) - z^4(0) < \frac{1}{4}$, and unbounded if

$$\dot{z}^2(0) + z^2(0) - z^4(0) > \tfrac{1}{4}.$$

13. (a) Let

$$L = n \times \max_{i,j=1,\ldots,n} |\partial f_i / \partial x_j|, \quad \text{for} \quad |\mathbf{x} - \mathbf{x}^0| \leq b.$$

Show that $|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| \leq L|\mathbf{x} - \mathbf{y}|$ if $|\mathbf{x} - \mathbf{x}^0| \leq b$ and $|\mathbf{y} - \mathbf{x}^0| \leq b$.

(b) Let $M = \max|\mathbf{f}(\mathbf{x})|$ for $|\mathbf{x} - \mathbf{x}^0| \leq b$. Show that the Picard iterates

$$\mathbf{x}_{j+1}(t) = \mathbf{x}^0 + \int_{t_0}^{t} \mathbf{f}(\mathbf{x}_j(s))\,ds, \qquad \mathbf{x}_0(t) = \mathbf{x}^0$$

converge to a solution $\mathbf{x}(t)$ of the initial-value problem $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, $\mathbf{x}(t_0) = \mathbf{x}^0$ on the interval $|t - t_0| \leq b/M$. *Hint*: The proof of Theorem 2, Section 1.10 carries over here word for word.

14. Compute the Picard iterates $\mathbf{x}_j(t)$ of the initial-value problem $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, $\mathbf{x}(0) = \mathbf{x}^0$, and verify that they approach $e^{\mathbf{A}t}\mathbf{x}^0$ as j approaches infinity.

4.7 Phase portraits of linear systems

In this section we present a complete picture of all orbits of the linear differential equation

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad \mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{1}$$

This picture is called a phase portrait, and it depends almost completely on the eigenvalues of the matrix **A**. It also changes drastically as the eigenvalues of **A** change sign or become imaginary.

When analyzing Equation (1), it is often helpful to visualize a vector

$$\mathbf{x}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

in R^2 as a direction, or directed line segment, in the plane. Let

$$\mathbf{x}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

be a vector in R^2 and draw the directed line segment $\vec{x}$ from the point (0,0) to the point (x_1,x_2), as in Figure 1a. This directed line segment is parallel to the line through (0,0) with direction numbers x_1,x_2 respectively. If we visualize the vector **x** as being this directed line segment $\vec{x}$, then we see that the vectors **x** and c**x** are parallel if c is positive, and antiparallel if c is negative. We can also give a nice geometric interpretation of vector addition. Let **x** and **y** be two vectors in R^2. Draw the directed line segment $\vec{x}$, and place the vector $\vec{y}$ at the tip of $\vec{x}$. The vector $\vec{x}+\vec{y}$ is then the composi-

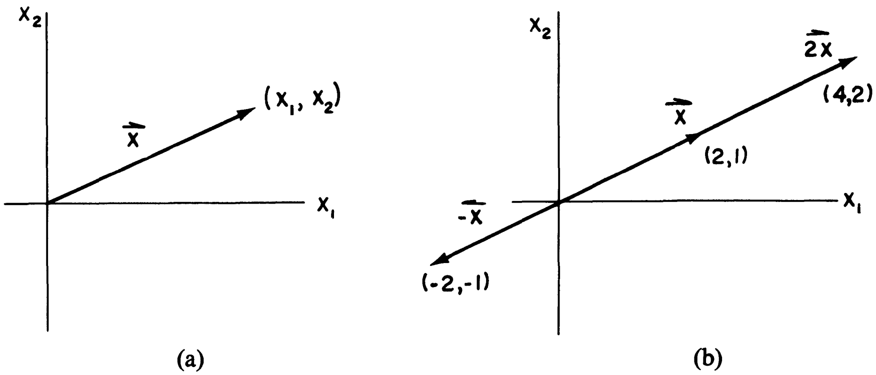

Figure 1

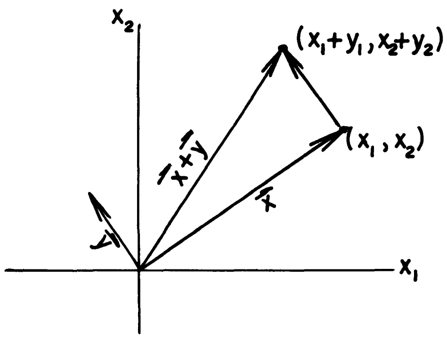

Figure 2

tion of these two directed line segments (see Figure 2). This construction is known as the parallelogram law of vector addition.

We are now in a position to derive the phase portraits of (1). Let λ_1 and λ_2 denote the two eigenvalues of **A**. We distinguish the following cases.

1. $\lambda_2 < \lambda_1 < 0$. Let $\mathbf{v}^1$ and $\mathbf{v}^2$ be eigenvectors of **A** with eigenvalues λ_1 and λ_2 respectively. In the $x_1 - x_2$ plane we draw the four half-lines l_1, l_1', l_2, and l_2', as shown in Figure 3. The rays l_1 and l_2 are parallel to $\mathbf{v}^1$ and $\mathbf{v}^2$, while the rays l_1' and l_2' are parallel to $-\mathbf{v}^1$ and $-\mathbf{v}^2$. Observe first that $\mathbf{x}(t) = ce^{\lambda_1 t}\mathbf{v}^1$ is a solution of (1) for any constant c. This solution is always proportional to $\mathbf{v}^1$, and the constant of proportionality, $ce^{\lambda_1 t}$, runs from $\pm\infty$ to 0, depending as to whether c is positive or negative. Hence, the orbit of this solution is the half-line l_1 for $c > 0$, and the half-line l_1' for $c < 0$. Similarly, the orbit of the solution $\mathbf{x}(t) = ce^{\lambda_2 t}\mathbf{v}^2$ is the half-line l_2 for $c > 0$, and the half-line l_2' for $c < 0$. The arrows on these four lines in Figure 3 indicate in what direction $\mathbf{x}(t)$ moves along its orbit.

Next, recall that every solution $\mathbf{x}(t)$ of (1) can be written in the form

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t}\mathbf{v}^1 + c_2 e^{\lambda_2 t}\mathbf{v}^2 \tag{2}$$

for some choice of constants c_1 and c_2. Obviously, every solution $\mathbf{x}(t)$ of (1) approaches $\binom{0}{0}$ as t approaches infinity. Hence, every orbit of (1) approaches the origin $x_1 = x_2 = 0$ as t approaches infinity. We can make an even stronger statement by observing that $e^{\lambda_2 t}\mathbf{v}^2$ is very small compared to $e^{\lambda_1 t}\mathbf{v}^1$ when t is very large. Therefore, $\mathbf{x}(t)$, for $c_1 \neq 0$, comes closer and closer to $c_1 e^{\lambda_1 t}\mathbf{v}^1$ as t approaches infinity. This implies that the tangent to the orbit of $\mathbf{x}(t)$ approaches l_1 if c_1 is positive, and l_1' if c_1 is negative. Thus,

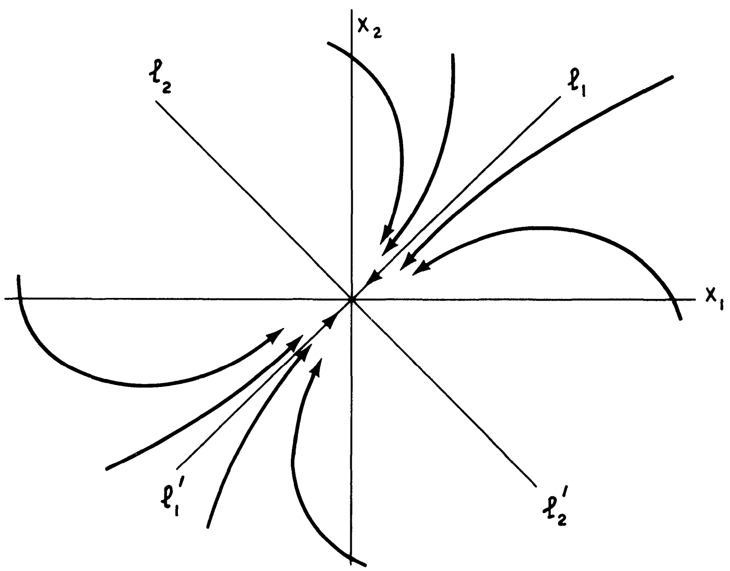

Figure 3. Phase portrait of a stable node

the phase portrait of (1) has the form described in Figure 3. The distinguishing feature of this phase portrait is that every orbit, with the exception of a single line, approaches the origin in a fixed direction (if we consider the directions $\mathbf{v}^1$ and $-\mathbf{v}^1$ equivalent). In this case we say that the equilibrium solution $\mathbf{x}=\mathbf{0}$ of (1) is a stable node.

Remark. The orbit of every solution $\mathbf{x}(t)$ of (1) approaches the origin $x_1=x_2=0$ as t approaches infinity. However, this point does not belong to the orbit of any nontrivial solution $\mathbf{x}(t)$.

1′. $0<\lambda_1<\lambda_2$. The phase portrait of (1) in this case is exactly the same as Figure 3, except that the direction of the arrows is reversed. Hence, the equilibrium solution $\mathbf{x}(t)=\mathbf{0}$ of (1) is an unstable node if both eigenvalues of $\mathbf{A}$ are positive.

2. $\lambda_1=\lambda_2<0$. In this case, the phase portrait of (1) depends on whether $\mathbf{A}$ has one or two linearly independent eigenvectors. (a) Suppose that $\mathbf{A}$ has two linearly independent eigenvectors $\mathbf{v}^1$ and $\mathbf{v}^2$ with eigenvalue $\lambda<0$. In this case, every solution $\mathbf{x}(t)$ of (1) can be written in the form

$$\mathbf{x}(t)=c_1e^{\lambda t}\mathbf{v}^1+c_2e^{\lambda t}\mathbf{v}^2=e^{\lambda t}\left(c_1\mathbf{v}^1+c_2\mathbf{v}^2\right) \tag{2}$$

for some choice of constants c_1 and c_2. Now, the vector $e^{\lambda t}(c_1\mathbf{v}^1+c_2\mathbf{v}^2)$ is parallel to $c_1\mathbf{v}^1+c_2\mathbf{v}^2$ for all t. Hence, the orbit of every solution $\mathbf{x}(t)$ of (1) is a half-line. Moreover, the set of vectors $\{c_1\mathbf{v}^1+c_2\mathbf{v}^2\}$, for all choices of c_1 and c_2, cover every direction in the x_1-x_2 plane, since $\mathbf{v}^1$ and $\mathbf{v}^2$ are linearly independent. Hence, the phase portrait of (1) has the form described in Figure 4a. (b) Suppose that $\mathbf{A}$ has only one linearly independent eigenvector $\mathbf{v}$, with eigenvalue λ. In this case, $\mathbf{x}^1(t)=e^{\lambda t}\mathbf{v}$ is one solution of (1). To find a second solution of (1) which is independent of $\mathbf{x}^1$, we observe that $(\mathbf{A}-\lambda\mathbf{I})^2\mathbf{u}=\mathbf{0}$ for every vector $\mathbf{u}$. Hence,

$$\mathbf{x}(t)=e^{\mathbf{A}t}\mathbf{u}=e^{\lambda t}e^{(\mathbf{A}-\lambda\mathbf{I})t}\mathbf{u}=e^{\lambda t}\left[\mathbf{u}+t(\mathbf{A}-\lambda\mathbf{I})\mathbf{u}\right] \tag{3}$$

is a solution of (1) for any choice of $\mathbf{u}$. Equation (3) can be simplified by observing that $(\mathbf{A}-\lambda\mathbf{I})\mathbf{u}$ must be a multiple k of $\mathbf{v}$. This follows immediately from the equation $(\mathbf{A}-\lambda\mathbf{I})[(\mathbf{A}-\lambda\mathbf{I})\mathbf{u}]=\mathbf{0}$, and the fact that $\mathbf{A}$ has only one linearly independent eigenvector $\mathbf{v}$. Choosing $\mathbf{u}$ independent of $\mathbf{v}$, we see that every solution $\mathbf{x}(t)$ of (1) can be written in the form

$$\mathbf{x}(t)=c_1e^{\lambda t}\mathbf{v}+c_2e^{\lambda t}(\mathbf{u}+kt\mathbf{v})=e^{\lambda t}(c_1\mathbf{v}+c_2\mathbf{u}+c_2kt\mathbf{v}), \tag{4}$$

for some choice of constants c_1 and c_2. Obviously, every solution $\mathbf{x}(t)$ of (1) approaches $\binom{0}{0}$ as t approaches infinity. In addition, observe that $c_1\mathbf{v}+c_2\mathbf{u}$ is very small compared to $c_2kt\mathbf{v}$ if c_2 is unequal to zero and t is very large. Hence, the tangent to the orbit of $\mathbf{x}(t)$ approaches $\pm\mathbf{v}$ (depending on the sign of c_2) as t approaches infinity, and the phase portrait of (1) has the form described in Figure 4b.

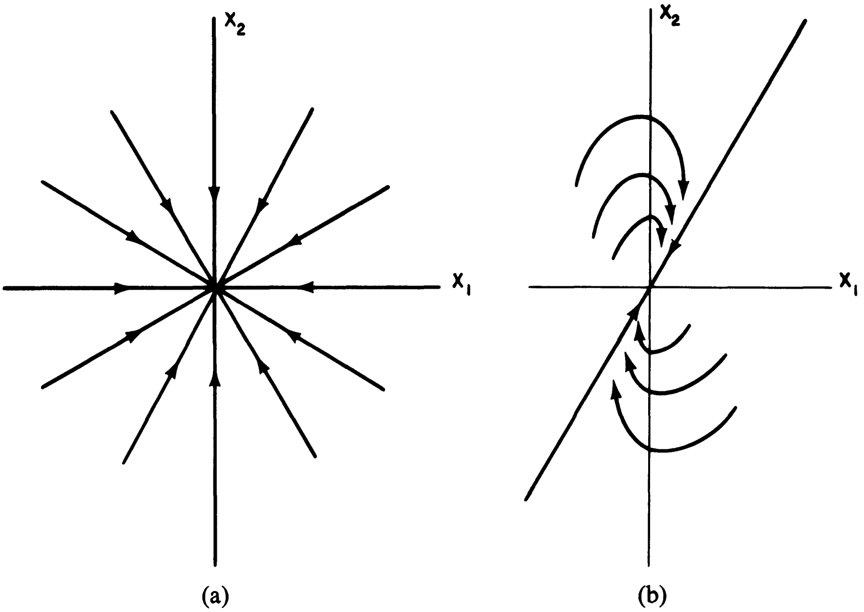

Figure 4

2′. $\lambda_1=\lambda_2>0$. The phase portraits of (1) in the cases (2a)′ and (2b)′ are exactly the same as Figures 4a and 4b, except that the direction of the arrows is reversed.

3. $\lambda_1<0<\lambda_2$. Let $\mathbf{v}^1$ and $\mathbf{v}^2$ be eigenvectors of $\mathbf{A}$ with eigenvalues λ_1 and λ_2 respectively. In the x_1-x_2 plane we draw the four half-lines l_1, l_1', l_2, and l_2'; the half-lines l_1 and l_2 are parallel to $\mathbf{v}^1$ and $\mathbf{v}^2$, while the half-lines l_1' and l_2' are parallel to $-\mathbf{v}^1$ and $-\mathbf{v}^2$. Observe first that every solution $\mathbf{x}(t)$ of (1) is of the form

$$\mathbf{x}(t)=c_1e^{\lambda_1 t}\mathbf{v}^1+c_2e^{\lambda_2 t}\mathbf{v}^2 \tag{5}$$

for some choice of constants c_1 and c_2. The orbit of the solution $\mathbf{x}(t)=c_1e^{\lambda_1 t}\mathbf{v}^1$ is l_1 for $c_1>0$ and l_1' for $c_1<0$, while the orbit of the solution $\mathbf{x}(t)=c_2e^{\lambda_2 t}\mathbf{v}^2$ is l_2 for $c_2>0$ and l_2' for $c_2<0$. Note, too, the direction of the arrows on l_1, l_1', l_2, and l_2'; the solution $\mathbf{x}(t)=c_1e^{\lambda_1 t}\mathbf{v}^1$ approaches $\binom{0}{0}$ as t approaches infinity, whereas the solution $\mathbf{x}(t)=c_2e^{\lambda_2 t}\mathbf{v}^2$ becomes unbounded (for $c_2\neq 0$) as t approaches infinity. Next, observe that $e^{\lambda_1 t}\mathbf{v}^1$ is very small compared to $e^{\lambda_2 t}\mathbf{v}^2$ when t is very large. Hence, every solution $\mathbf{x}(t)$ of (1) with $c_2\neq 0$ becomes unbounded as t approaches infinity, and its orbit approaches either l_2 or l_2'. Finally, observe that $e^{\lambda_2 t}\mathbf{v}^2$ is very small compared to $e^{\lambda_1 t}\mathbf{v}^1$ when t is very large negative. Hence, the orbit of any solution $\mathbf{x}(t)$ of (1), with $c_1\neq 0$, approaches either l_1 or l_1' as t approaches minus infinity. Consequently, the phase portrait of (1) has the form described in Figure 5. This phase portrait resembles a "saddle" near $x_1=x_2=$

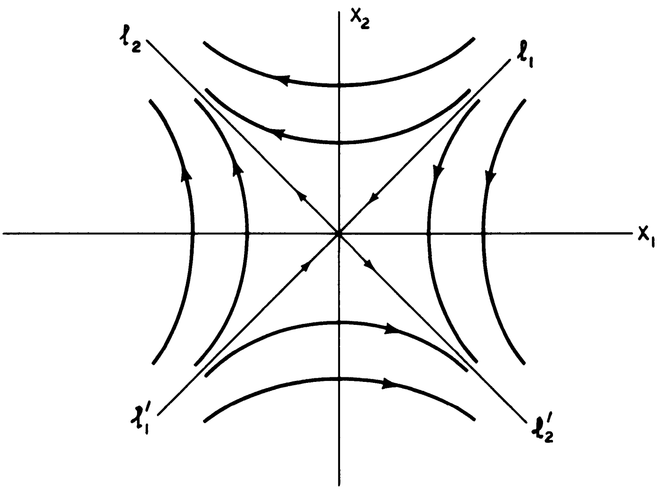

Figure 5. Phase portrait of a saddle point

0. For this reason, we say that the equilibrium solution $\mathbf{x}(t)=\mathbf{0}$ of (1) is a saddle point if the eigenvalues of $\mathbf{A}$ have opposite sign.

4. $\lambda_1=\alpha+i\beta, \lambda_2=\alpha-i\beta, \beta\neq 0$. Our first step in deriving the phase portrait of (1) is to find the general solution of (1). Let $\mathbf{z}=\mathbf{u}+i\mathbf{v}$ be an eigenvector of $\mathbf{A}$ with eigenvalue $\alpha+i\beta$. Then,

$$\begin{aligned}\mathbf{x}(t)&=e^{(\alpha+i\beta)t}(\mathbf{u}+i\mathbf{v})=e^{\alpha t}(\cos\beta t+i\sin\beta t)(\mathbf{u}+i\mathbf{v})\\&=e^{\alpha t}[\mathbf{u}\cos\beta t-\mathbf{v}\sin\beta t]+ie^{\alpha t}[\mathbf{u}\sin\beta t+\mathbf{v}\cos\beta t]\end{aligned}$$

is a complex-valued solution of (1). Therefore,

$$\mathbf{x}^1(t)=e^{\alpha t}[\mathbf{u}\cos\beta t-\mathbf{v}\sin\beta t]$$

and

$$\mathbf{x}^2(t)=e^{\alpha t}[\mathbf{u}\sin\beta t+\mathbf{v}\cos\beta t]$$

are two real-valued linearly independent solutions of (1), and every solution $\mathbf{x}(t)$ of (1) is of the form $\mathbf{x}(t)=c_1\mathbf{x}^1(t)+c_2\mathbf{x}^2(t)$. This expression can be written in the form (see Exercise 15)

$$\mathbf{x}(t)=e^{\alpha t}\begin{pmatrix}R_1\cos(\beta t-\delta_1)\\R_2\cos(\beta t-\delta_2)\end{pmatrix}\tag{6}$$

for some choice of constants $R_1\geqslant 0$, $R_2\geqslant 0$, δ_1, and δ_2. We distinguish the following cases.

(a) $\alpha=0$: Observe that both

$$x_1(t)=R_1\cos(\beta t-\delta_1)\quad\text{and}\quad x_2(t)=R_2\cos(\beta t-\delta_2)$$

are periodic functions of time with period $2\pi/\beta$. The function $x_1(t)$ varies between $-R_1$ and $+R_1$, while $x_2(t)$ varies between $-R_2$ and $+R_2$. Consequently, the orbit of any solution $\mathbf{x}(t)$ of (1) is a closed curve surrounding the origin $x_1 = x_2 = 0$, and the phase portrait of (1) has the form described in Figure 6a. For this reason, we say that the equilibrium solution $\mathbf{x}(t) = \mathbf{0}$ of (1) is a center when the eigenvalues of $\mathbf{A}$ are pure imaginary.

The direction of the arrows in Figure 6a must be determined from the differential equation (1). The simplest way of doing this is to check the sign of $\dot{x}_2$ when $x_2 = 0$. If $\dot{x}_2$ is greater than zero for $x_2 = 0$ and $x_1 > 0$, then all solutions $\mathbf{x}(t)$ of (1) move in the counterclockwise direction; if $\dot{x}_2$ is less than zero for $x_2 = 0$ and $x_1 > 0$, then all solutions $\mathbf{x}(t)$ of (1) move in the clockwise direction.

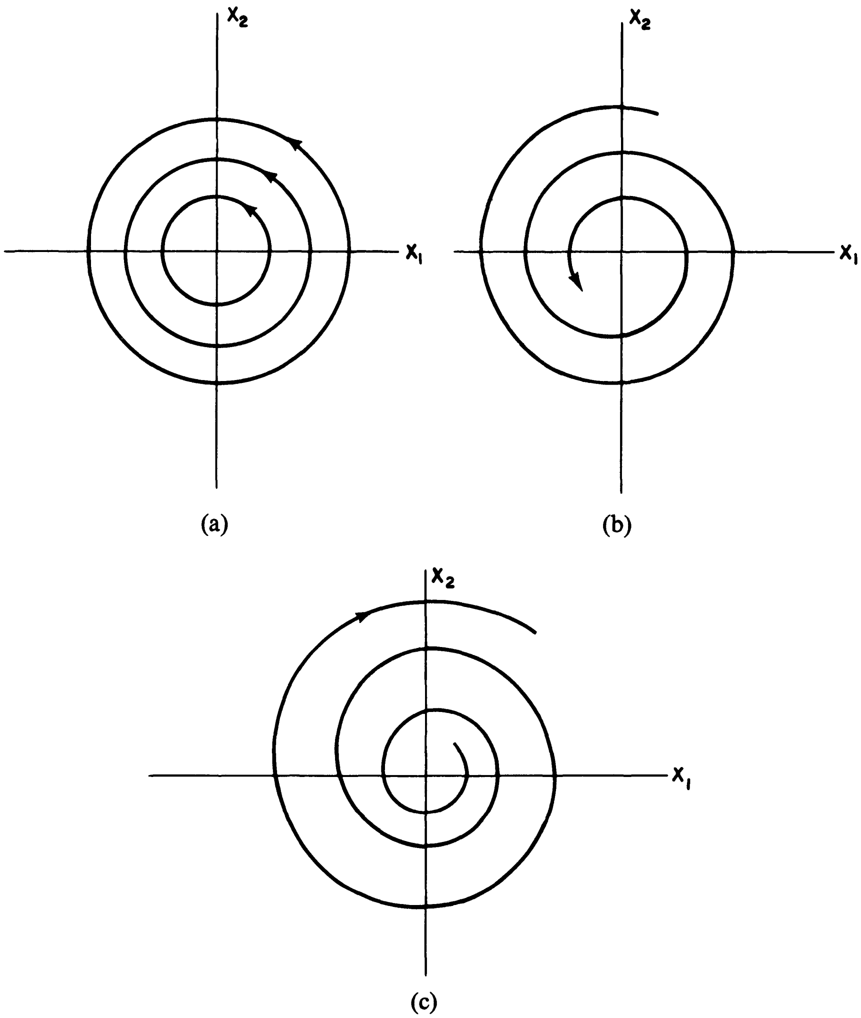

Figure 6. (a) $\alpha = 0$; (b) $\alpha < 0$; (c) $\alpha > 0$

(b) $\alpha<0$: In this case, the effect of the factor $e^{\alpha t}$ in Equation (6) is to change the simple closed curves of Figure 6a into the spirals of Figure 6b. This is because the point $\mathbf{x}(2\pi/\beta)=e^{2\pi\alpha/\beta}\mathbf{x}(0)$ is closer to the origin than $\mathbf{x}(0)$. Again, the direction of the arrows in Figure 6b must be determined directly from the differential equation (1). In this case, we say that the equilibrium solution $\mathbf{x}(t)=\mathbf{0}$ of (1) is a stable focus.

(c) $\alpha>0$: In this case, all orbits of (1) spiral away from the origin as t approaches infinity (see Figure 6c), and the equilibrium solution $\mathbf{x}(t)=\mathbf{0}$ of (1) is called an unstable focus.

Finally, we mention that the phase portraits of nonlinear systems, in the neighborhood of an equilibrium point, are often very similar to the phase portraits of linear systems. More precisely, let $\mathbf{x}=\mathbf{x}^0$ be an equilibrium solution of the nonlinear equation $\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x})$, and set $\mathbf{u}=\mathbf{x}-\mathbf{x}^0$. Then, (see Section 4.3) we can write the differential equation $\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x})$ in the form

$$\dot{\mathbf{u}}=\mathbf{A}\mathbf{u}+\mathbf{g}(\mathbf{u}) \tag{7}$$

where $\mathbf{A}$ is a constant matrix and $\mathbf{g}(\mathbf{u})$ is very small compared to $\mathbf{u}$. We state without proof the following theorem.

Theorem 4. *Suppose that* $\mathbf{u}=\mathbf{0}$ *is either a node, saddle, or focus point of the differential equation* $\dot{\mathbf{u}}=\mathbf{A}\mathbf{u}$. *Then, the phase portrait of the differential equation* $\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x})$, *in a neighborhood of* $\mathbf{x}=\mathbf{x}^0$, *has one of the forms described in Figures* 3, 5, *and* 6 (b *and* c), *depending as to whether* $\mathbf{u}=\mathbf{0}$ *is a node, saddle, or focus.*

Example 1. Draw the phase portrait of the linear equation

$$\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}=\begin{pmatrix} -2 & -1 \\ 4 & -7 \end{pmatrix}\mathbf{x}. \tag{8}$$

Solution. It is easily verified that

$$\mathbf{v}^1=\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{v}^2=\begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

are eigenvectors of $\mathbf{A}$ with eigenvalues -3 and -6, respectively. Therefore, $\mathbf{x}=\mathbf{0}$ is a stable node of (8), and the phase portrait of (8) has the form described in Figure 7. The half-line l_1 makes an angle of 45° with the x_1 axis, while the half-line l_2 makes an angle of θ degrees with the x_1-axis, where $\tan\theta=4$.

Example 2. Draw the phase portrait of the linear equation

$$\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}=\begin{pmatrix} 1 & -3 \\ -3 & 1 \end{pmatrix}\mathbf{x}. \tag{9}$$

Solution. It is easily verified that

$$\mathbf{v}^1=\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{v}^2=\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

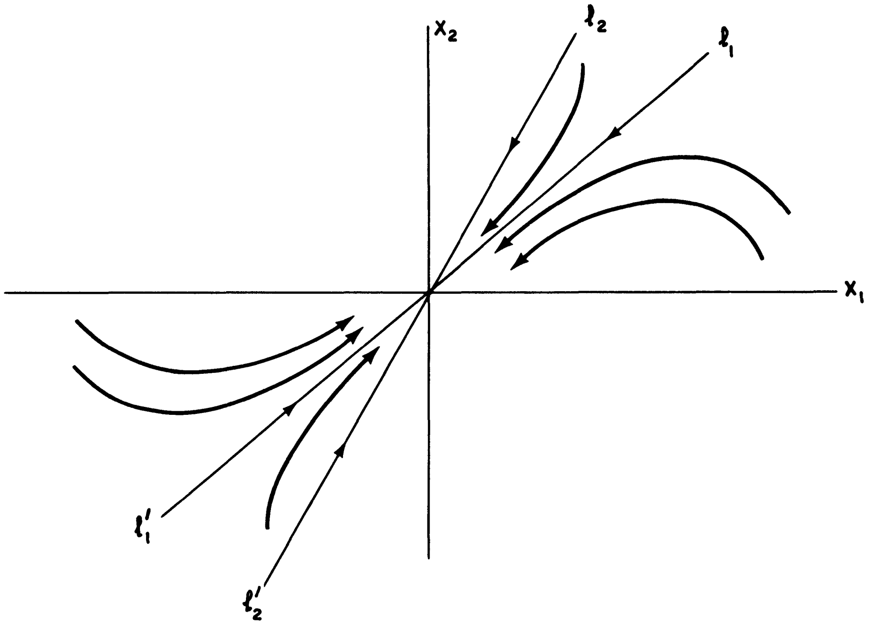

Figure 7. Phase portrait of (8)

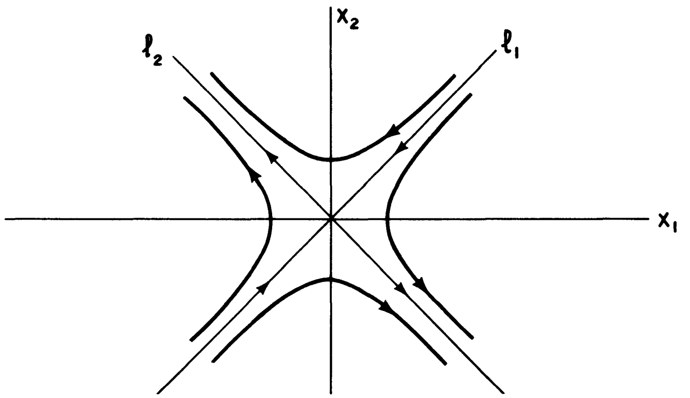

Figure 8. Phase portrait of (9)

are eigenvectors of **A** with eigenvalues -2 and 4, respectively. Therefore, $\mathbf{x}=\mathbf{0}$ is a saddle point of (9), and its phase portrait has the form described in Figure 8. The half-line l_1 makes an angle of 45° with the x_1-axis, and the half-line l_2 is at right angles to l_1.

Example 3. Draw the phase portrait of the linear equation

$$\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}=\begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix}\mathbf{x}. \tag{10}$$

Solution. The eigenvalues of $\mathbf{A}$ are $-1 \pm i$. Hence, $\mathbf{x}=\mathbf{0}$ is a stable focus of (10) and every nontrivial orbit of (10) spirals into the origin as t approaches infinity. To determine the direction of rotation of the spiral, we observe that $\dot{x}_2 = -x_1$ when $x_2 = 0$. Thus, $\dot{x}_2$ is negative for $x_1 > 0$ and $x_2 = 0$. Consequently, all nontrivial orbits of (10) spiral into the origin in the clockwise direction, as shown in Figure 9.

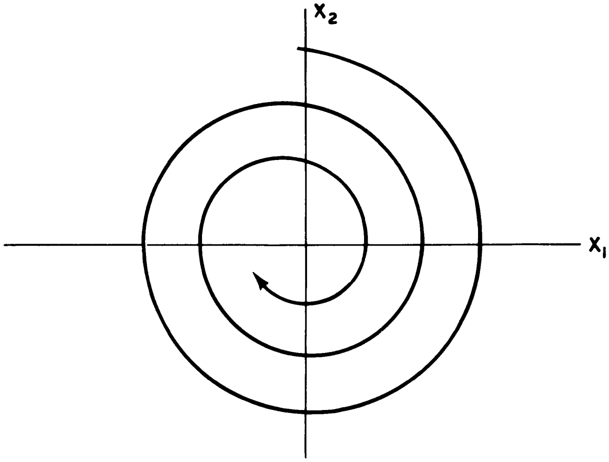

Figure 9. Phase portrait of (10)

Exercises

Draw the phase portraits of each of the following systems of differential equations.

1. $\dot{\mathbf{x}} = \begin{pmatrix} -5 & 1 \\ 1 & -5 \end{pmatrix}\mathbf{x}$ **2.** $\dot{\mathbf{x}} = \begin{pmatrix} 0 & -1 \\ 8 & -6 \end{pmatrix}\mathbf{x}$ **3.** $\dot{\mathbf{x}} = \begin{pmatrix} 4 & -1 \\ -2 & 5 \end{pmatrix}\mathbf{x}$

4. $\dot{\mathbf{x}} = \begin{pmatrix} -4 & -1 \\ 1 & -6 \end{pmatrix}\mathbf{x}$ **5.** $\dot{\mathbf{x}} = \begin{pmatrix} 1 & -4 \\ -8 & 4 \end{pmatrix}\mathbf{x}$ **6.** $\dot{\mathbf{x}} = \begin{pmatrix} 3 & -1 \\ 5 & -3 \end{pmatrix}\mathbf{x}$

7. $\dot{\mathbf{x}} = \begin{pmatrix} 0 & 2 \\ -2 & -1 \end{pmatrix}\mathbf{x}$ **8.** $\dot{\mathbf{x}} = \begin{pmatrix} 1 & -1 \\ 5 & -3 \end{pmatrix}\mathbf{x}$ **9.** $\dot{\mathbf{x}} = \begin{pmatrix} 2 & 1 \\ -5 & -2 \end{pmatrix}\mathbf{x}$

10. Show that every orbit of

$$\dot{\mathbf{x}} = \begin{pmatrix} 0 & 4 \\ -9 & 0 \end{pmatrix}\mathbf{x}$$

is an ellipse.

11. The equation of motion of a spring-mass system with damping (see Section 2.6) is $m\ddot{z} + c\dot{z} + kz = 0$, where m, c, and k are positive numbers. Convert this equation to a system of first-order equations for $x = z$, $y = \dot{z}$, and draw the phase portrait of this system. Distinguish the overdamped, critically damped, and underdamped cases.

12. Suppose that a 2×2 matrix $\mathbf{A}$ has 2 linearly independent eigenvectors with eigenvalue λ. Show that $\mathbf{A} = \lambda\mathbf{I}$.

13. This problem illustrates Theorem 4. Consider the system

$$\frac{dx}{dt}=y, \qquad \frac{dy}{dt}=x+2x^3. \tag{*}$$

(a) Show that the equilibrium solution $x=0, y=0$ of the linearized system $\dot{x}=y$, $\dot{y}=x$ is a saddle, and draw the phase portrait of the linearized system.
(b) Find the orbits of (*), and then draw its phase portrait.
(c) Show that there are exactly two orbits of (*) (one for $x>0$ and one for $x<0$) on which $x\to 0, y\to 0$ as $t\to\infty$. Similarly, there are exactly two orbits of (*) on which $x\to 0, y\to 0$ as $t\to-\infty$. Thus, observe that the phase portraits of (*) and the linearized system look the same near the origin.

14. Verify Equation (6). *Hint*: The expression $a\cos\omega t+b\sin\omega t$ can always be written in the form $R\cos(\omega t-\delta)$ for suitable choices of R and δ.

4.8 Long time behavior of solutions; the Poincaré–Bendixson Theorem

We consider now the problem of determining the long time behavior of all solutions of the differential equation

$$\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x}), \qquad \mathbf{x}=\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{f}(\mathbf{x})=\begin{bmatrix} f_1(x_1,\dots,x_n) \\ \vdots \\ f_n(x_1,\dots,x_n) \end{bmatrix}. \tag{1}$$

This problem has been solved completely in the special case that $\mathbf{f}(\mathbf{x})=\mathbf{A}\mathbf{x}$. As we have seen in Sections 4.2 and 4.7, all solutions $\mathbf{x}(t)$ of $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$ must exhibit one of the following four types of behavior: (i) $\mathbf{x}(t)$ is constant in time; (ii) $\mathbf{x}(t)$ is a periodic function of time; (iii) $\mathbf{x}(t)$ is unbounded as t approaches infinity; and (iv) $\mathbf{x}(t)$ approaches an equilibrium point as t approaches infinity.

A partial solution to this problem, in the case of nonlinear $\mathbf{f}(\mathbf{x})$, was given in Section 4.3. In that section we provided sufficient conditions that every solution $\mathbf{x}(t)$ of (1), whose initial value $\mathbf{x}(0)$ is sufficiently close to an equilibrium point $\boldsymbol{\xi}$, must ultimately approach $\boldsymbol{\xi}$ as t approaches infinity. In many applications it is often possible to go much further and prove that every physically (biologically) realistic solution approaches a single equilibrium point as time evolves. In this context, the following two lemmas play an extremely important role.

Lemma 1. *Let $g(t)$ be a monotonic increasing (decreasing) function of time for $t\geq t_0$, with $g(t)\leq c(\geq c)$ for some constant c. Then, $g(t)$ has a limit as t approaches infinity.*

PROOF. Suppose that $g(t)$ is monotonic increasing for $t\geq t_0$, and $g(t)$ is bounded from above. Let l be the least upper bound of g; that is, l is the smallest number which is not exceeded by the values of $g(t)$, for $t\geq t_0$. This number must be the limit of $g(t)$ as t approaches infinity. To prove this, let $\varepsilon>0$ be given, and observe that there exists a time $t_\varepsilon \geqslant t_0$ such that $l-g(t_\varepsilon)<\varepsilon$. (If no such time t_ε exists, then l is not the least upper bound of g.) Since $g(t)$ is monotonic, we see that $l-g(t)<\varepsilon$ for $t\geqslant t_\varepsilon$. This shows that $l=\lim_{t\to\infty} g(t)$. □

Lemma 2. *Suppose that a solution* $\mathbf{x}(t)$ *of* (1) *approaches a vector* $\boldsymbol{\xi}$ *as* t *approaches infinity. Then,* $\boldsymbol{\xi}$ *is an equilibrium point of* (1).

Proof. Suppose that $\mathbf{x}(t)$ approaches $\boldsymbol{\xi}$ as t approaches infinity. Then, $x_j(t)$ approaches ξ_j, where ξ_j is the jth component of $\boldsymbol{\xi}$. This implies that $|x_j(t_1)-x_j(t_2)|$ approaches zero as both t_1 and t_2 approach infinity, since

$$\begin{aligned}|x_j(t_1)-x_j(t_2)| &= |(x_j(t_1)-\xi_j)+(\xi_j-x_j(t_2))|\\ &\leqslant |x_j(t_1)-\xi_j|+|x_j(t_2)-\xi_j|.\end{aligned}$$

In particular, let $t_1=t$ and $t_2=t_1+h$, for some fixed positive number h. Then, $|x_j(t+h)-x_j(t)|$ approaches zero as t approaches infinity. But

$$x_j(t+h)-x_j(t)=h\frac{dx_j(\tau)}{dt}=hf_j(x_1(\tau),\dots,x_n(\tau)),$$

where τ is some number between t and $t+h$. Finally, observe that $f_j(x_1(\tau),\dots,x_n(\tau))$ must approach $f_j(\xi_1,\dots,\xi_n)$ as t approaches infinity. Hence, $f_j(\xi_1,\dots,\xi_n)=0, j=1,2,\dots,n$, and this proves Lemma 1. □

Example 1. Consider the system of differential equations

$$\frac{dx}{dt}=ax-bxy-ex^2,\qquad \frac{dy}{dt}=-cy+dxy-fy^2 \tag{2}$$

where a, b, c, d, e, and f are positive constants. This system (see Section 4.10) describes the population growth of two species x and y, where species y is dependent upon species x for its survival. Suppose that $c/d>a/e$. Prove that every solution $x(t)$, $y(t)$ of (2), with $x(0)$ and $y(0)>0$, approaches the equilibrium solution $x=a/e$, $y=0$, as t approaches infinity.
Solution. Our first step is to show that every solution $x(t), y(t)$ of (2) which starts in the first quadrant $(x>0, y>0)$ at $t=0$ must remain in the first quadrant for all future time. (If this were not so, then the model (2) could not correspond to reality.) To this end, recall from Section 1.5 that

$$x(t)=\frac{ax_0}{ex_0+(a-ex_0)e^{-at}},\qquad y(t)=0$$

is a solution of (2) for any choice of x_0. The orbit of this solution is the point $(0,0)$ for $x_0=0$; the line $0<x<a/e$ for $0<x_0<a/e$; the point $(a/e,0)$ for $x_0=a/e$; and the line $a/e<x<\infty$ for $x_0>a/e$. Thus, the x-axis, for $x\geqslant 0$, is the union of four disjoint orbits of (2). Similarly, (see Exercise 14), the positive y-axis is a single orbit of (2). Thus, if a solution

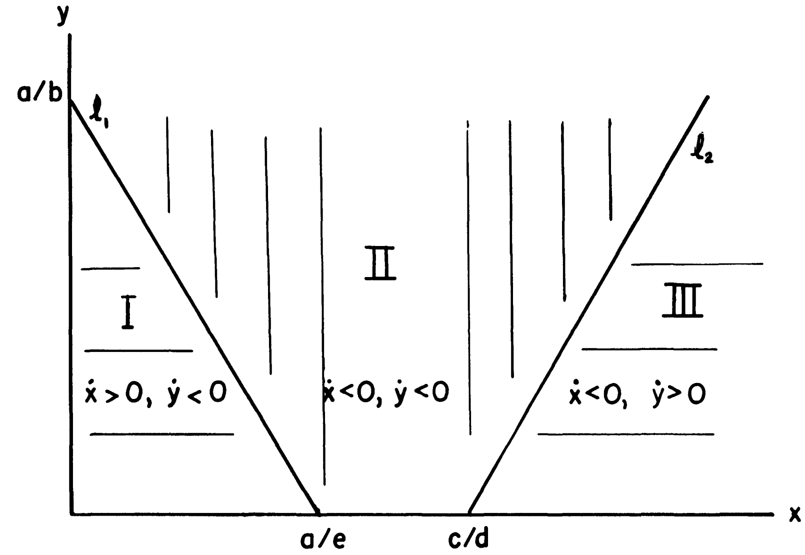

Figure 1

$x(t), y(t)$ of (2) leaves the first quadrant, its orbit must cross another orbit, and this is precluded by the uniqueness of orbits (Property 1, Section 4.6).

Our next step is to divide the first quadrant into regions where dx/dt and dy/dt have fixed signs. This is accomplished by drawing the lines $l_1: a - by - ex = 0$, and $l_2: -c + dx - fy = 0$, in the x–y plane. These lines divide the first quadrant into three regions I, II, and III as shown in Figure 1. (The lines l_1 and l_2 do not intersect in the first quadrant if $c/d > a/e$.) Now, observe that $ex + by$ is less than a in region I, while $ex + by$ is greater than a in regions II and III. Consequently, dx/dt is positive in region I and negative in regions II and III. Similarly, dy/dt is negative in regions I and II and positive in region III.

Next, we prove the following four simple lemmas.

Lemma 3. *Any solution $x(t), y(t)$ of (2) which starts in region I at time $t = t_0$ will remain in this region for all future time $t \geqslant t_0$ and ultimately approach the equilibrium solution $x = a/e, y = 0$.*

PROOF. Suppose that a solution $x(t), y(t)$ of (2) leaves region I at time $t = t^*$. Then, $\dot{x}(t^*) = 0$, since the only way a solution can leave region I is by crossing the line l_1. Differentiating both sides of the first equation of (2) with respect to t and setting $t = t^*$ gives

$$\frac{d^2x(t^*)}{dt^2} = -bx(t^*)\frac{dy(t^*)}{dt}.$$

This quantity is positive. Hence, $x(t)$ has a minimum at $t = t^*$. But this is impossible, since $x(t)$ is always increasing whenever $x(t), y(t)$ is in region

I. Thus, any solution $x(t), y(t)$ of (2) which starts in region I at time $t=t_0$ will remain in region I for all future time $t \geq t_0$. This implies that $x(t)$ is a monotonic increasing function of time, and $y(t)$ is a monotonic decreasing function of time for $t \geq t_0$, with $x(t) < a/e$ and $y(t) > 0$. Consequently, by Lemma 1, both $x(t)$ and $y(t)$ have limits ξ, η respectively, as t approaches infinity. Lemma 2 implies that (ξ, η) is an equilibrium point of (2). Now, it is easily verified that the only equilibrium points of (2) in the region $x \geq 0$, $y \geq 0$ are $x=0, y=0$, and $x=a/e, y=0$. Clearly, ξ cannot equal zero since $x(t)$ is increasing in region I. Therefore, $\xi = a/e$ and $\eta = 0$. □

Lemma 4. *Any solution $x(t), y(t)$ of* (2) *which starts in region* III *at time $t=t_0$ must leave this region at some later time.*

Proof. Suppose that a solution $x(t), y(t)$ of (2) remains in region III for all time $t \geq t_0$. Then, $x(t)$ is a monotonic decreasing function of time, and $y(t)$ is a monotonic increasing function of time, for $t \geq t_0$. Moreover, $x(t)$ is greater than c/d and $y(t)$ is less than $(dx(t_0)-c)/f$. Consequently, both $x(t)$ and $y(t)$ have limits ξ, η respectively, as t approaches infinity. Lemma 2 implies that (ξ, η) is an equilibrium point of (2). But (ξ, η) cannot equal $(0,0)$ or $(a/e, 0)$ if $x(t), y(t)$ is in region III for $t \geq t_0$. This contradiction establishes Lemma 4. □

Lemma 5. *Any solution $x(t), y(t)$ of* (2) *which starts in region* II *at time $t=t_0$ and remains in region II for all future time $t \geq t_0$ must approach the equilibrium solution $x=a/e, y=0$.*

Proof. Suppose that a solution $x(t), y(t)$ of (2) remains in region II for all time $t \geq t_0$. Then, both $x(t)$ and $y(t)$ are monotonic decreasing functions of time for $t \geq t_0$, with $x(t) > 0$ and $y(t) > 0$. Consequently, by Lemma 1, both $x(t)$ and $y(t)$ have limits ξ, η respectively, as t approaches infinity. Lemma 2 implies that (ξ, η) is an equilibrium point of (2). Now, (ξ, η) cannot equal $(0,0)$. Therefore, $\xi = a/e, \eta = 0$. □

Lemma 6. *A solution $x(t), y(t)$ of* (2) *cannot enter region* III *from region* II.

Proof. Suppose that a solution $x(t), y(t)$ of (2) leaves region II at time $t = t^*$ and enters region III. Then, $\dot{y}(t^*) = 0$. Differentiating both sides of the second equation of (2) with respect to t and setting $t = t^*$ gives

$$\frac{d^2 y(t^*)}{dt^2} = dy(t^*) \frac{dx(t^*)}{dt}.$$

This quantity is negative. Hence, $y(t)$ has a maximum at $t = t^*$. But this is impossible, since $y(t)$ is decreasing whenever $x(t), y(t)$ is in region II. □

Finally, observe that a solution $x(t), y(t)$ of (2) which starts on l_1 must immediately enter region I, and that a solution which starts on l_2 must immediately enter region II. It now follows immediately from Lemmas 3–6 that every solution $x(t), y(t)$ of (2), with $x(0)>0$ and $y(0)>0$, approaches the equilibrium solution $x=a/e, y=0$ as t approaches infinity.

Up to now, the solutions and orbits of the nonlinear equations that we have studied behaved very much like the solutions and orbits of linear equations. In actual fact, though, the situation is very different. The solutions and orbits of nonlinear equations, in general, exhibit a completely different behavior than the solutions and orbits of linear equations. A standard example is the system of equations

$$\frac{dx}{dt} = -y + x(1-x^2-y^2), \qquad \frac{dy}{dt} = x + y(1-x^2-y^2). \tag{3}$$

Since the term x^2+y^2 appears prominently in both equations, it suggests itself to introduce polar coordinates r, θ, where $x = r\cos\theta, y = r\sin\theta$, and to rewrite (3) in terms of r and θ. To this end, we compute

$$\begin{aligned}\frac{d}{dt}r^2 = 2r\frac{dr}{dt} &= 2x\frac{dx}{dt} + 2y\frac{dy}{dt}\\ &= 2(x^2+y^2) - 2(x^2+y^2)^2 = 2r^2(1-r^2).\end{aligned}$$

Similarly, we compute

$$\frac{d\theta}{dt} = \frac{d}{dt}\arctan\frac{y}{x} = \frac{1}{x^2}\frac{x\dfrac{dy}{dt} - y\dfrac{dx}{dt}}{1+(y/x)^2} = \frac{x^2+y^2}{x^2+y^2} = 1.$$

Consequently, the system of equations (3) is equivalent to the system of equations

$$\frac{dr}{dt} = r(1-r^2), \qquad \frac{d\theta}{dt} = 1. \tag{4}$$

The general solution of (4) is easily seen to be

$$r(t) = \frac{r_0}{\left[r_0^2 + (1-r_0^2)e^{-2t}\right]^{1/2}}, \qquad \theta = t + \theta_0 \tag{5}$$

where $r_0 = r(0)$ and $\theta_0 = \theta(0)$. Hence,

$$\begin{aligned} x(t) &= \frac{r_0}{\left[r_0^2 + (1-r_0^2)e^{-2t}\right]^{1/2}}\cos(t+\theta_0),\\ y(t) &= \frac{r_0}{\left[r_0^2 + (1-r_0^2)e^{-2t}\right]^{1/2}}\sin(t+\theta_0).\end{aligned}$$

Now, observe first that $x=0, y=0$ is the only equilibrium solution of (3). Second, observe that

$$x(t) = \cos(t+\theta_0), \qquad y(t) = \sin(t+\theta_0)$$

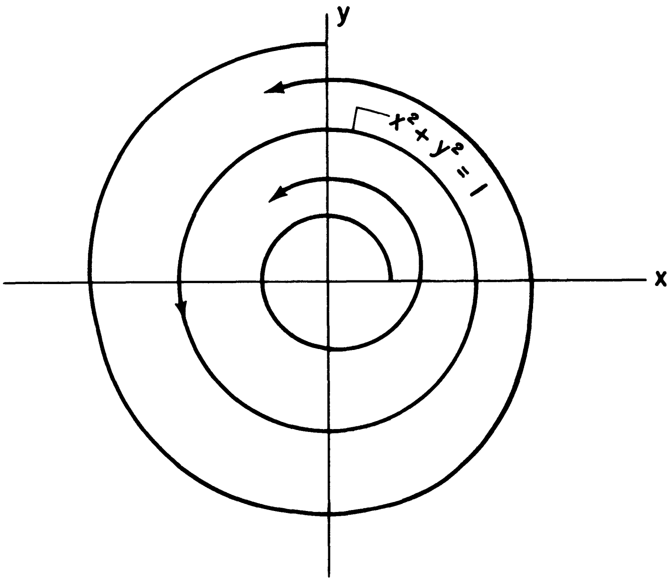

Figure 2. The phase portrait of (3)

when $r_0 = 1$. This solution is periodic with period 2π, and its orbit is the unit circle $x^2 + y^2 = 1$. Finally, observe from (5) that $r(t)$ approaches one as t approaches infinity, for $r_0 \neq 0$. Hence, all the orbits of (3), with the exception of the equilibrium point $x = 0$, $y = 0$, spiral into the unit circle. This situation is depicted in Figure 2.

The system of equations (3) shows that the orbits of a nonlinear system of equations may spiral into a simple closed curve. This, of course, is not possible for linear systems. Moreover, it is often possible to prove that orbits of a nonlinear system spiral into a closed curve even when we cannot explicitly solve the system of equations, or even find its orbits. This is the content of the following celebrated theorem.

Theorem 5. (Poincaré–Bendixson.) *Suppose that a solution $x = x(t)$, $y = y(t)$ of the system of differential equations*

$$\frac{dx}{dt} = f(x,y), \qquad \frac{dy}{dt} = g(x,y) \tag{6}$$

remains in a bounded region of the plane which contains no equilibrium points of (6). *Then, its orbit must spiral into a simple closed curve, which is itself the orbit of a periodic solution of* (6).

Example 2. Prove that the second-order differential equation

$$\ddot{z} + (z^2 + 2\dot{z}^2 - 1)\dot{z} + z = 0 \tag{7}$$

has a nontrivial periodic solution.

Solution. First, we convert Equation (7) to a system of two first-order equations by setting $x=z$ and $y=\dot{z}$. Then,

$$\frac{dx}{dt}=y, \qquad \frac{dy}{dt}=-x+(1-x^2-2y^2)y. \tag{8}$$

Next, we try and find a bounded region R in the $x-y$ plane, containing no equilibrium points of (8), and having the property that every solution $x(t)$, $y(t)$ of (8) which starts in R at time $t=t_0$, remains there for all future time $t \geqslant t_0$. It can be shown that a simply connected region such as a square or disc will never work. Therefore, we try and take R to be an annulus surrounding the origin. To this end, compute

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

and observe that $1-x^2-2y^2$ is positive for $x^2+y^2<\frac{1}{2}$ and negative for $x^2+y^2>1$. Hence, $x^2(t)+y^2(t)$ is increasing along any solution $x(t)$, $y(t)$ of (8) when $x^2+y^2<\frac{1}{2}$ and decreasing when $x^2+y^2>1$. This implies that any solution $x(t)$, $y(t)$ of (8) which starts in the annulus $\frac{1}{2}<x^2+y^2<1$ at time $t=t_0$ will remain in this annulus for all future time $t \geqslant t_0$. Now, this annulus contains no equilibrium points of (8). Consequently, by the Poincaré–Bendixson Theorem, there exists at least one periodic solution $x(t), y(t)$ of (8) lying entirely in this annulus, and then $z=x(t)$ is a nontrivial periodic solution of (7).

Exercises

1. *What Really Happened at the Paris Peace Talks*
The original plan developed by Henry Kissinger and Le Duc Tho to settle the Vietnamese war is described below. It was agreed that 1 million South Vietnamese ants and 1 million North Vietnamese ants would be placed in the backyard of the Presidential palace in Paris and be allowed to fight it out for a long period of time. If the South Vietnamese ants destroyed nearly all the North Vietnamese ants, then South Vietnam would retain control of all of its land. If the North Vietnamese ants were victorious, then North Vietnam would take over all of South Vietnam. If they appeared to be fighting to a standoff, then South Vietnam would be partitioned according to the proportion of ants remaining. Now, the South Vietnamese ants, denoted by S, and the North Vietnamese ants, denoted by N, compete against each other according to the following differential equations:

$$\begin{aligned}\frac{dS}{dt}&=\frac{1}{10}S-\frac{1}{20}S\times N\\ \frac{dN}{dt}&=\frac{1}{100}N-\frac{1}{100}N^2-\frac{1}{100}S\times N.\end{aligned} \tag{*}$$

Note that these equations correspond to reality since the South Vietnamese ants multiply much more rapidly than the North Vietnamese ants, but the North Vietnamese ants are much better fighters.

The battle began at 10:00 sharp on the morning of May 19, 1972, and was supervised by a representative of Poland and a representative of Canada. At 2:43 p.m. on the afternoon of May 21, the representative of Poland, being unhappy with the progress of the battle, slipped a bag of North Vietnamese ants into the backyard, but he was spotted by the eagle eyes of the representative of Canada. The South Vietnamese immediately claimed a foul and called off the agreement, thus setting the stage for the protracted talks that followed in Paris. The representative of Poland was hauled before a judge in Paris for sentencing. The judge, after making some remarks about the stupidity of the South Vietnamese, gave the Polish representative a very light sentence. Justify mathematically the judge's decision. *Hint*:

(a) Show that the lines $N=2$ and $N+S=1$ divide the first quadrant into three regions (see Figure 3) in which dS/dt and dN/dt have fixed signs.

(b) Show that every solution $S(t),N(t)$ of (*) which starts in either region I or region III must eventually enter region II.

(c) Show that every solution $S(t),N(t)$ of (*) which starts in region II must remain there for all future time.

(d) Conclude from (c) that $S(t)\to\infty$ for all solutions $S(t),N(t)$ of (*) with $S(t_0)$ and $N(t_0)$ positive. Conclude too that $N(t)$ has a finite limit ($\leqslant 2$) as $t\to\infty$.

(e) To prove that $N(t)\to 0$, observe that there exists t_0 such that $dN/dt \leqslant -N$ for $t \geqslant t_0$. Conclude from this inequality that $N(t)\to 0$ as $t\to\infty$.

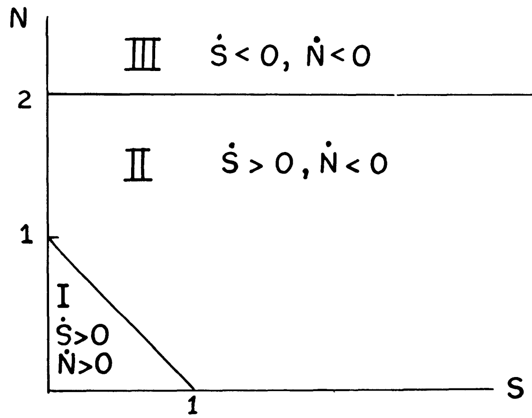

Figure 3

2. Consider the system of differential equations

$$\frac{dx}{dt} = ax - bxy, \qquad \frac{dy}{dt} = cy - dxy - ey^2 \qquad (*)$$

with $a/b > c/e$. Prove that $y(t)\to 0$ as $t\to\infty$, for every solution $x(t), y(t)$ of (*) with $x(t_0)$ and $y(t_0)$ positive. *Hint*: Follow the outline in Exercise 1.

3. (a) Without computing the eigenvalues of the matrix

$$\begin{pmatrix} -3 & 1 \\ 1 & -3 \end{pmatrix},$$

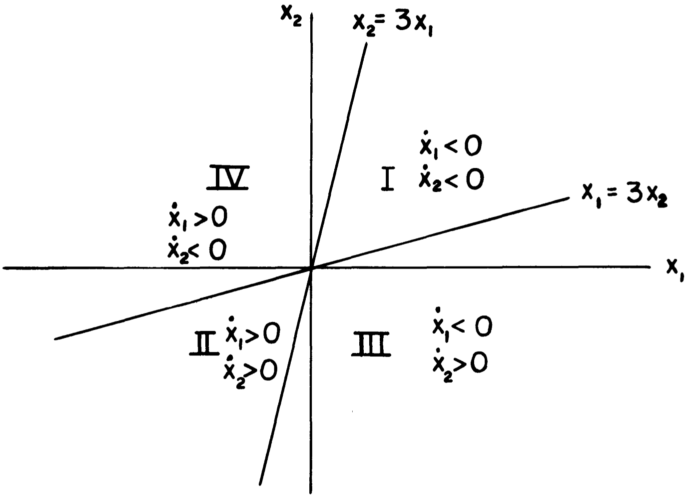

Figure 4

prove that every solution $\mathbf{x}(t)$ of

$$\dot{\mathbf{x}}=\begin{pmatrix} -3 & 1 \\ 1 & -3 \end{pmatrix}\mathbf{x}$$

approaches zero as t approaches infinity. *Hint*: (a) Show that the lines $x_2=3x_1$ and $x_1=3x_2$ divide the x_1-x_2 plane into four regions (see Figure 4) in which $\dot{x}_1$ and $\dot{x}_2$ have fixed signs.

(b) Show that every solution $\mathbf{x}(t)$ which starts in either region I or II must remain there for all future time and ultimately approach the equilibrium solution $\mathbf{x}=\mathbf{0}$.

(c) Show that every solution $\mathbf{x}(t)$ which remains exclusively in region III or IV must ultimately approach the equilibrium solution $\mathbf{x}=\mathbf{0}$.

A closed curve C is said to be a limit cycle of

$$\dot{x}=f(x,y), \qquad \dot{y}=g(x,y) \tag{*}$$

if orbits of (*) spiral into it, or away from it. It is stable if all orbits of (*) passing sufficiently close to it must ultimately spiral into it, and unstable otherwise. Find all limit cycles of each of the following systems of differential equations. (*Hint*: Compute $d(x^2+y^2)/dt$. Observe too, that C must be the orbit of a periodic solution of (*) if it contains no equilibrium points of (*).)

4. $\dot{x}=-y-\dfrac{x(x^2+y^2-2)}{\sqrt{x^2+y^2}}$

$\dot{y}=x-\dfrac{y(x^2+y^2-2)}{\sqrt{x^2+y^2}}$

5. $\dot{x}=x-x^3-xy^2$

$\dot{y}=y-y^3-yx^2$

6. $\dot{x}=y+x(x^2+y^2-1)(x^2+y^2-2)$
$\dot{y}=-x+y(x^2+y^2-1)(x^2+y^2-2)$

7. $\dot{x}=xy+x\cos(x^2+y^2)$
$\dot{y}=-x^2+y\cos(x^2+y^2)$

8. (a) Show that the system

$$\dot{x}=y+xf(r)/r,\qquad \dot{y}=-x+yf(r)/r\qquad (r^2=x^2+y^2)\tag{*}$$

has limit cycles corresponding to the zeros of $f(r)$. What is the direction of motion on these curves?

(b) Determine all limit cycles of (*) and discuss their stability if $f(r)=(r-3)^2(r^2-5r+4)$.

Use the Poincaré–Bendixson Theorem to prove the existence of a nontrivial periodic solution of each of the following differential equations.

9. $\ddot{z}+(z^2+\dot{z}^4-2)z=0$

10. $\ddot{z}+[\ln(z^2+4\dot{z}^2)]\dot{z}+z=0$

11. (a) According to Green's theorem in the plane, if C is a closed curve which is sufficiently "smooth," and if f and g are continuous and have continuous first partial derivatives, then

$$\oint_C [f(x,y)\,dy-g(x,y)\,dx]=\iint_R [f_x(x,y)+g_y(x,y)]\,dx\,dy$$

where R is the region enclosed by C. Assume that $x(t)$, $y(t)$ is a periodic solution of $\dot{x}=f(x,y)$, $\dot{y}=g(x,y)$, and let C be the orbit of this solution. Show that for this curve, the line integral above is zero.

(b) Suppose that f_x+g_y has the same sign throughout a simply connected region D in the $x-y$ plane. Show that the system of equations $\dot{x}=f(x,y)$, $\dot{y}=g(x,y)$ can have no periodic solution which is entirely in D.

12. Show that the system of differential equations

$$\dot{x}=x+y^2+x^3,\qquad \dot{y}=-x+y+yx^2$$

has no nontrivial periodic solution.

13. Show that the system of differential equations

$$\dot{x}=x-xy^2+y^3,\qquad \dot{y}=3y-yx^2+x^3$$

has no nontrivial periodic solution which lies inside the circle $x^2+y^2=4$.

14. (a) Show that $x=0, y=\psi(t)$ is a solution of (2) for any function $\psi(t)$ satisfying $\dot{\psi}=-c\psi-f\psi^2$.

(b) Choose $\psi(t_0)>0$. Show that the orbit of $x=0, y=\psi(t)$ (for all t for which ψ exists) is the positive y axis.

4.9 Introduction to bifurcation theory

Consider the system of equations

$$\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x},\varepsilon)\tag{1}$$

where

$$\mathbf{x}=\begin{pmatrix}x_1\\x_2\end{pmatrix},$$

and ε is a scalar. Intuitively speaking a bifurcation point of (1) is a value of ε at which the solutions of (1) change their behavior. More precisely, we say that $\varepsilon=\varepsilon_0$ is a bifurcation point of (1) if the phase portraits of (1) for $\varepsilon<\varepsilon_0$ and $\varepsilon>\varepsilon_0$ are different.

Remark. In the examples that follow we will appeal to our intuition in deciding whether two phase portraits are the same or are different. In more advanced courses we define two phase portraits to be the same, or topologically equivalent, if there exists a continuous transformation of the plane onto itself which maps one phase portrait onto the other.

Example 1. Find the bifurcation points of the system

$$\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}=\begin{pmatrix}1 & \varepsilon\\ 1 & -1\end{pmatrix}\mathbf{x} \tag{2}$$

Solution. The characteristic polynomial of the matrix $\mathbf{A}$ is

$$\begin{aligned} p(\lambda)&=\det(\mathbf{A}-\lambda\mathbf{I})\\ &=\det\begin{pmatrix}1-\lambda & \varepsilon\\ 1 & -1-\lambda\end{pmatrix}\\ &=(\lambda-1)(\lambda+1)-\varepsilon\\ &=\lambda^2-(1+\varepsilon). \end{aligned}$$

The roots of $p(\lambda)$ are $\pm\sqrt{1+\varepsilon}$ for $\varepsilon>-1$, and $\pm\sqrt{-\varepsilon-1}\,i$ for $\varepsilon<-1$. This

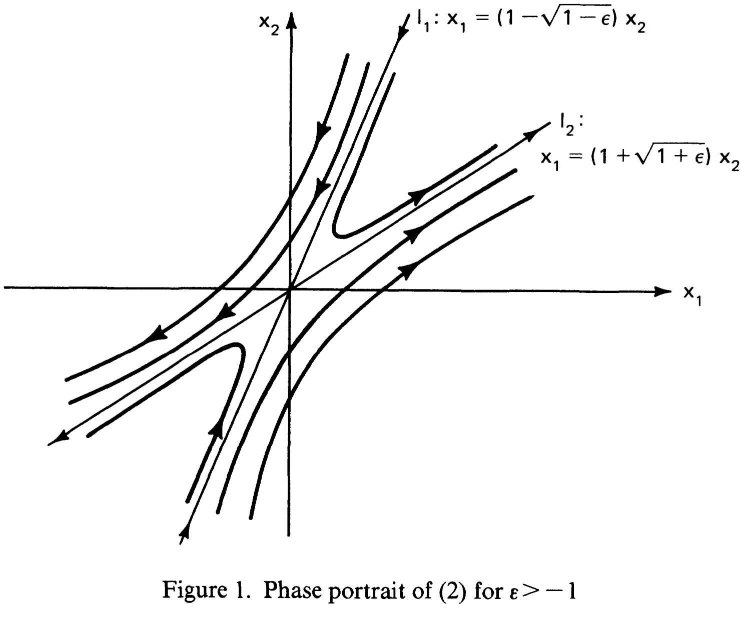

Figure 1. Phase portrait of (2) for $\varepsilon>-1$

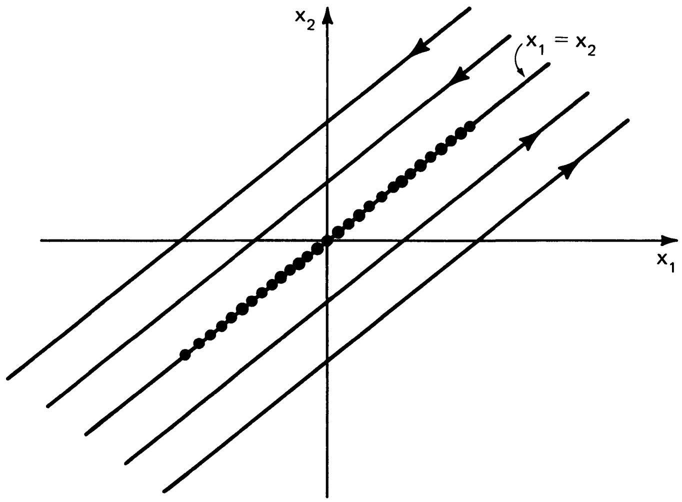

Figure 2. Phase portrait of (2) for $\varepsilon = -1$

implies that $\mathbf{x} = \mathbf{0}$ is a saddle for $\varepsilon > -1$, and a center for $\varepsilon < -1$. We conclude, therefore, that $\varepsilon = -1$ is a bifurcation point of (2). It is also clear that Eq. (2) has no other bifurcation points.

It is instructive to see how the solutions of (2) change as ε passes through the bifurcation value -1. For $\varepsilon > -1$, the eigenvalues of $\mathbf{A}$ are

$$\lambda_1 = \sqrt{1+\varepsilon}\,, \quad \lambda_2 = -\sqrt{1+\varepsilon}\,.$$

It is easily verified (see Exercise 10) that

$$\mathbf{x}^1 = \begin{pmatrix} 1+\sqrt{1+\varepsilon} \\ 1 \end{pmatrix}$$

is an eigenvector of $\mathbf{A}$ with eigenvalue $\sqrt{1+\varepsilon}$, while

$$\mathbf{x}^2 = \begin{pmatrix} 1-\sqrt{1+\varepsilon} \\ 1 \end{pmatrix}$$

is an eigenvector with eigenvalue $-\sqrt{1+\varepsilon}$. Hence, the phase portrait of (2) has the form shown in Figure 1. As $\varepsilon \to -1$ from the left, the lines l_1 and l_2 both approach the line $x_1 = x_2$. This line is a line of equilibrium points of (2) when $\varepsilon = -1$, while each line $x_1 - x_2 = c$ $(c \neq 0)$ is an orbit of (2) for $\varepsilon = -1$. The phase portrait of (2) for $\varepsilon = -1$ is given in Figure 2.

Example 2. Find the bifurcation points of the system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} = \begin{pmatrix} 0 & -1 \\ \varepsilon & -1 \end{pmatrix}\mathbf{x}. \tag{3}$$

Solution. The characteristic polynomial of the matrix $\mathbf{A}$ is

$$\begin{aligned} p(\lambda) = \det(\mathbf{A}-\lambda\mathbf{I}) &= \det\begin{pmatrix} -\lambda & -1 \\ \varepsilon & -1-\lambda \end{pmatrix} \\ &= \lambda(1+\lambda)+\varepsilon \\ &= \lambda^2+\lambda+\varepsilon \end{aligned}$$

and the roots of $p(\lambda)$ are

$$\lambda_1 = \frac{-1+\sqrt{1-4\varepsilon}}{2}, \lambda_2 = \frac{-1-\sqrt{1-4\varepsilon}}{2}.$$

Observe that λ_1 is positive and λ_2 is negative for $\varepsilon<0$. Hence, $\mathbf{x}=\mathbf{0}$ is a saddle for $\varepsilon<0$. For $0<\varepsilon<1/4$, both λ_1 and λ_2 are negative. Hence $\mathbf{x}=\mathbf{0}$ is a stable node for $0<\varepsilon<1/4$. Both λ_1 and λ_2 are complex, with negative real part, for $\varepsilon>1/4$. Hence $\mathbf{x}=\mathbf{0}$ is a stable focus for $\varepsilon>1/4$. Note that the phase portrait of (3) changes as ε passes through 0 and $1/4$. We conclude, therefore, that $\varepsilon=0$ and $\varepsilon=1/4$ are bifurcation points of (3).

Example 3. Find the bifurcation points of the system of equations

$$\begin{aligned} \frac{dx_1}{dt} &= x_2 \\ \frac{dx_2}{dt} &= x_1^2 - x_2 - \varepsilon. \end{aligned} \tag{4}$$

Solution. (i) We first find the equilibrium points of (4). Setting $dx_1/dt=0$ gives $x_2=0$, and then setting $dx_2/dt=0$ gives $x_1^2-\varepsilon=0$, so that $x_1=\pm\sqrt{\varepsilon}$, $\varepsilon>0$. Hence, $(\sqrt{\varepsilon},0)$ and $(-\sqrt{\varepsilon},0)$ are two equilibrium points of (4) for $\varepsilon>0$. The system (4) has no equilibrium points when $\varepsilon<0$. We conclude, therefore, that $\varepsilon=0$ is a bifurcation point of (4). (ii) We now analyze the behavior of the solutions of (4) near the equilibrium points $(\pm\sqrt{\varepsilon},0)$ to determine whether this system has any additional bifurcation points. Setting

$$u = x_1 \mp \sqrt{\varepsilon}, \quad v = x_2$$

gives

$$\begin{aligned} \frac{du}{dt} &= v \\ \frac{dv}{dt} &= \left(u \pm \sqrt{\varepsilon}\right)^2 - v - \varepsilon = \pm 2\sqrt{\varepsilon}\,u - v + u^2. \end{aligned} \tag{5}$$

The system (5) can be written in the form

$$\frac{d}{dt}\begin{pmatrix}u\\v\end{pmatrix}=\begin{pmatrix}0 & 1\\ \pm 2\sqrt{\varepsilon} & -1\end{pmatrix}\begin{pmatrix}u\\v\end{pmatrix}+\begin{pmatrix}0\\u^2\end{pmatrix}.$$

By Theorem 4, the phase portrait of (4) near the equilibrium solution

$$\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}\pm\sqrt{\varepsilon}\\0\end{pmatrix}$$

is determined by the phase portrait of the linearized system

$$\frac{d}{dt}\begin{pmatrix}u\\v\end{pmatrix}=\mathbf{A}\begin{pmatrix}u\\v\end{pmatrix}=\begin{pmatrix}0 & 1\\ \pm 2\sqrt{\varepsilon} & -1\end{pmatrix}\begin{pmatrix}u\\v\end{pmatrix}.$$

To find the eigenvalues of $\mathbf{A}$ we compute

$$\begin{aligned}p(\lambda)&=\det(\mathbf{A}-\lambda\mathbf{I})\\&=\det\begin{pmatrix}-\lambda & 1\\ \pm 2\sqrt{\varepsilon} & -1-\lambda\end{pmatrix}\\&=\lambda^2+\lambda\mp 2\sqrt{\varepsilon}.\end{aligned}$$

Hence, the eigenvalues of $\mathbf{A}$ when $u=x_1-\sqrt{\varepsilon}$ are

$$\lambda_1=\frac{-1+\sqrt{1+8\sqrt{\varepsilon}}}{2},\quad \lambda_2=\frac{-1-\sqrt{1+8\sqrt{\varepsilon}}}{2} \tag{6}$$

while the eigenvalues of $\mathbf{A}$ when $u=x_1+\sqrt{\varepsilon}$ are

$$\lambda_1=\frac{-1+\sqrt{1-8\sqrt{\varepsilon}}}{2},\quad \lambda_2=\frac{-1-\sqrt{1-8\sqrt{\varepsilon}}}{2}. \tag{7}$$

Observe from (6) that $\lambda_1>0$, while $\lambda_2<0$. Thus, the system (4) behaves like a saddle near the equilibrium points $\begin{pmatrix}\sqrt{\varepsilon}\\0\end{pmatrix}$. On the other hand, we see from (7) that both λ_1 and λ_2 are negative for $0<\varepsilon<1/64$, and complex for $\varepsilon>1/64$. Consequently, the system (4) near the equilibrium solution $\begin{pmatrix}-\sqrt{\varepsilon}\\0\end{pmatrix}$ behaves like a stable node for $0<\varepsilon<1/64$, and a stable focus for $\varepsilon>1/64$. It can be shown that the phase portraits of a stable node and a stable focus are equivalent. Consequently, $\varepsilon=1/64$ is *not* a bifurcation point of (4).

Another situation which is included in the context of bifurcation theory, and which is of much current research interest now, is when the system (1) has a certain number of equilibrium, or periodic, solutions for $\varepsilon=\varepsilon_0$, and a different number for $\varepsilon\neq\varepsilon_0$. Suppose, for example, that $\mathbf{x}(t)$ is an equilibrium, or periodic, solution of (1) for $\varepsilon=0$, and $\mathbf{x}^1(t),\mathbf{x}^2(t),\ldots,\mathbf{x}^k(t)$ are equilibria, or periodic solutions of (1) for $\varepsilon\neq 0$ which approach $\mathbf{x}(t)$ as $\varepsilon\to 0$. In this case we say that the solutions $\mathbf{x}^1(t),\mathbf{x}^2(t),\ldots,\mathbf{x}^k(t)$ *bifurcate* from $\mathbf{x}(t)$. We illustrate this situation with the following example.

Example 4. Find all equilibrium solutions of the system of equations

$$\begin{aligned}\frac{dx_1}{dt}&=3\varepsilon x_1-3\varepsilon x_2-x_1^2-x_2^2\\ \frac{dx}{dt}&=\varepsilon x_1-x_1x_2=x_1(\varepsilon-x_2).\end{aligned}\tag{8}$$

Solution. Let $\begin{pmatrix}x_1\\x_2\end{pmatrix}$ be an equilibrium solution of the system (8). The second equation of (8) implies that $x_1=0$ or $x_2=\varepsilon$.

$x_1=0$. In this case, the first equation of (8) implies that

$$0=3\varepsilon x_2+x_2^2=x_2(3\varepsilon+x_2)$$

so that $x_2=0$ or $x_2=-3\varepsilon$. Thus $\begin{pmatrix}0\\0\end{pmatrix}$ and $\begin{pmatrix}0\\-3\varepsilon\end{pmatrix}$ are two equilibrium points of (8).

$x_2=\varepsilon$. In this case, the first equation of (8) implies that

$$3\varepsilon x_1-3\varepsilon^2-x_1^2-\varepsilon^2=0$$

or

$$x_1^2-3\varepsilon x_1+4\varepsilon^2=0.\tag{9}$$

The solutions

$$x_1=\frac{3\varepsilon\pm\sqrt{9\varepsilon^2-16\varepsilon^2}}{2}$$

of (9) are complex. Thus,

$$\mathbf{x}^1=\begin{pmatrix}0\\0\end{pmatrix}\quad\text{and}\quad\mathbf{x}^2=\begin{pmatrix}0\\-3\varepsilon\end{pmatrix}$$

are two equilibrium points of (8), for $\varepsilon\neq 0$, which bifurcate from the single equilibrium point $\mathbf{x}=\begin{pmatrix}0\\0\end{pmatrix}$ when $\varepsilon=0$.

Exercises

Find the bifurcation points of each of the following systems of equations.

1. $\dot{\mathbf{x}}=\begin{pmatrix}1 & \varepsilon\\ \varepsilon & 1\end{pmatrix}\mathbf{x}$

2. $\dot{\mathbf{x}}=\begin{pmatrix}1 & 1\\ \varepsilon & 1\end{pmatrix}\mathbf{x}$

3. $\dot{\mathbf{x}}=\begin{pmatrix}0 & 2\\ -2 & \varepsilon\end{pmatrix}\mathbf{x}$

4. $\dot{\mathbf{x}}=\begin{pmatrix}0 & 2\\ 2 & \varepsilon\end{pmatrix}\mathbf{x}$

5. $\dot{\mathbf{x}}=\begin{pmatrix}0 & -\varepsilon\\ \varepsilon & 0\end{pmatrix}\mathbf{x}$

In each of Problems 6–8, show that more than one equilibrium solutions bifurcate from the equilibrium solution $\mathbf{x}=\mathbf{0}$ when $\varepsilon=0$.

6. $\dot{x}_1=\varepsilon x_1-\varepsilon x_2-x_1^2+x_2^2$
$\dot{x}_2=\varepsilon x_2+x_1x_2$

7. $\dot{x}_1=\varepsilon x_1-x_1^2-x_1x_2$
$\dot{x}_2=-2\varepsilon x_1+2\varepsilon x_2+x_1x_2-x_2^2$

8. $\dot{x}_1=\varepsilon x_2+x_1x_2$
$\dot{x}_2=-\varepsilon x_1+\varepsilon x_2+x_1^2+x_2^2$

9. Consider the system of equations

$$\begin{aligned}\dot{x}_1&=3\varepsilon x_1-5\varepsilon x_2-x_1^2+x_2^2\\ \dot{x}_2&=2\varepsilon x_1-\varepsilon x_2.\end{aligned}\qquad (*)$$

(a) Show that each point on the lines $x_2=x_1$ and $x_2=-x_1$ are equilibrium points of $(*)$ for $\varepsilon=0$.

(b) Show that

$$\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}\quad\text{and}\quad\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\tfrac{7}{3}\varepsilon\begin{pmatrix}1\\ 2\end{pmatrix}.$$

are the only equilibrium points of $(*)$ for $\varepsilon\neq 0$.

10. Show that

$$\mathbf{x}^1=\begin{pmatrix}1+\sqrt{1+\varepsilon}\\ 1\end{pmatrix}\quad\text{and}\quad\mathbf{x}^2=\begin{pmatrix}1-\sqrt{1+\varepsilon}\\ 1\end{pmatrix}$$

are eigenvectors of the matrix $\begin{pmatrix}1 & \varepsilon\\ 1 & -1\end{pmatrix}$ with eigenvalues $\sqrt{1+\varepsilon}$ and $-\sqrt{1+\varepsilon}$ respectively.

4.10 Predator–prey problems; or why the percentage of sharks caught in the Mediterranean Sea rose dramatically during World War I

In the mid 1920's the Italian biologist Umberto D'Ancona was studying the population variations of various species of fish that interact with each other. In the course of his research, he came across some data on percentages-of-total-catch of several species of fish that were brought into different Mediterranean ports in the years that spanned World War I. In particular, the data gave the percentage-of-total-catch of selachians, (sharks, skates, rays, etc.) which are not very desirable as food fish. The data for the port of Fiume, Italy, during the years 1914–1923 is given below.

1914	1915	1916	1917	1918
11.9%	21.4%	22.1%	21.2%	36.4%
1919	1920	1921	1922	1923
27.3%	16.0%	15.9%	14.8%	10.7%

D'Ancona was puzzled by the very large increase in the percentage of selachians during the period of the war. Obviously, he reasoned, the increase in the percentage of selachians was due to the greatly reduced level of fishing during this period. But how does the intensity of fishing affect the fish populations? The answer to this question was of great concern to D'Ancona in his research on the struggle for existence between competing species. It was also of concern to the fishing industry, since it would have obvious implications for the way fishing should be done.

Now, what distinguishes the selachians from the food fish is that the selachians are predators, while the food fish are their prey; the selachians depend on the food fish for their survival. At first, D'Ancona thought that this accounted for the large increase of selachians during the war. Since the level of fishing was greatly reduced during this period, there were more prey available to the selachians, who therefore thrived and multiplied rapidly. However, this explanation does not hold any water since there were also more food fish during this period. D'Ancona's theory only shows that there are more selachians when the level of fishing is reduced; it does not explain why a reduced level of fishing is *more* beneficial to the predators than to their prey.

After exhausting all possible biological explanations of this phenomenon, D'Ancona turned to his colleague, the famous Italian mathematician Vito Volterra. Hopefully, Volterra would formulate a mathematical model of the growth of the selachians and their prey, the food fish, and this model would provide the answer to D'Ancona's question. Volterra began his analysis of this problem by separating all the fish into the prey population $x(t)$ and the predator population $y(t)$. Then, he reasoned that the food fish do not compete very intensively among themselves for their food supply since this is very abundant, and the fish population is not very dense. Hence, in the absence of the selachians, the food fish would grow according to the Malthusian law of population growth $\dot{x}=ax$, for some positive constant a. Next, reasoned Volterra, the number of contacts per unit time between predators and prey is bxy, for some positive constant b. Hence, $\dot{x}=ax-bxy$. Similarly, Volterra concluded that the predators have a natural rate of decrease $-cy$ proportional to their present number, and that they also increase at a rate dxy proportional to their present number y and their food supply x. Thus,

$$\frac{dx}{dt} = ax - bxy, \qquad \frac{dy}{dt} = -cy + dxy. \tag{1}$$

The system of equations (1) governs the interaction of the selachians and food fish in the absence of fishing. We will carefully analyze this system and derive several interesting properties of its solutions. Then, we will include the effect of fishing in our model, and show why a reduced level of fishing is more beneficial to the selachians than to the food fish. In fact, we will derive the surprising result that a reduced level of fishing is actually harmful to the food fish.

Observe first that (1) has two equilibrium solutions $x(t)=0, y(t)=0$ and $x(t)=c/d, y(t)=a/b$. The first equilibrium solution, of course, is of no interest to us. This system also has the family of solutions $x(t)=x_0e^{at}, y(t)=0$ and $x(t)=0, y(t)=y_0e^{-ct}$. Thus, both the x and y axes are orbits of (1). This implies that every solution $x(t)$, $y(t)$ of (1) which starts in the first quadrant $x>0, y>0$ at time $t=t_0$ will remain there for all future time $t \geqslant t_0$.

The orbits of (1), for $x,y\neq 0$ are the solution curves of the first-order equation

$$\frac{dy}{dx} = \frac{-cy+dxy}{ax-bxy} = \frac{y(-c+dx)}{x(a-by)}. \tag{2}$$

This equation is separable, since we can write it in the form

$$\frac{a-by}{y}\frac{dy}{dx} = \frac{-c+dx}{x}.$$

Consequently, $a\ln y - by + c\ln x - dx = k_1$, for some constant k_1. Taking exponentials of both sides of this equation gives

$$\frac{y^a}{e^{by}}\frac{x^c}{e^{dx}} = K \tag{3}$$

for some constant K. Thus, the orbits of (1) are the family of curves defined by (3), and these curves are *closed* as we now show.

Lemma 1. *Equation* (3) *defines a family of closed curves for* $x,y>0$.

PROOF. Our first step is to determine the behavior of the functions $f(y)=y^a/e^{by}$ and $g(x)=x^c/e^{dx}$ for x and y positive. To this end, observe that $f(0)=0, f(\infty)=0$, and $f(y)$ is positive for $y>0$. Computing

$$f'(y) = \frac{ay^{a-1}-by^a}{e^{by}} = \frac{y^{a-1}(a-by)}{e^{by}},$$

we see that $f(y)$ has a single critical point at $y=a/b$. Consequently, $f(y)$ achieves its maximum value $M_y=(a/b)^a/e^a$ at $y=a/b$, and the graph of

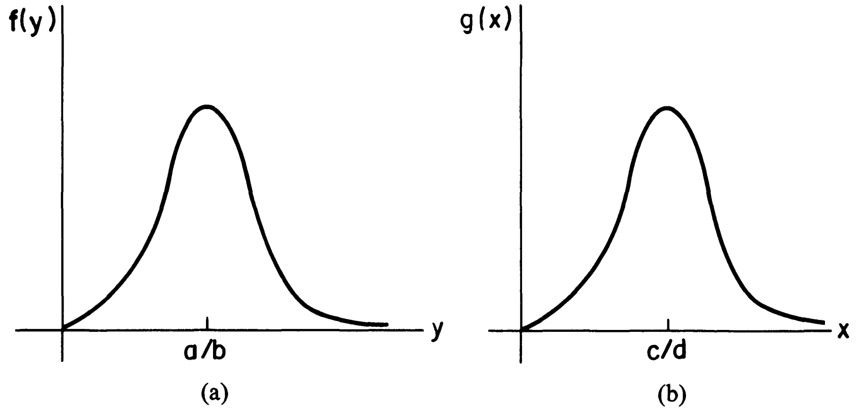

Figure 1. (a) Graph of $f(y)=y^ae^{-by}$; (b) Graph of $g(x)=x^ce^{-dx}$

$f(y)$ has the form described in Figure 1a. Similarly, $g(x)$ achieves its maximum value $M_x=(c/d)^c/e^c$ at $x=c/d$, and the graph of $g(x)$ has the form described in Figure 1b.

From the preceding analysis, we conclude that Equation (3) has no solution $x,y>0$ for $K>M_xM_y$, and the single solution $x=c/d$, $y=a/b$ for $K=M_xM_y$. Thus, we need only consider the case $K=\lambda M_y$, where λ is a positive number less than M_x. Observe first that the equation $x^c/e^{dx}=\lambda$ has one solution $x=x_m<c/d$, and one solution $x=x_M>c/d$. Hence, the equation

$$f(y)=y^ae^{-by}=\left(\frac{\lambda}{x^ce^{-dx}}\right)M_y$$

has no solution y when x is less than x_m or greater than x_M. It has the single solution $y=a/b$ when $x=x_m$ or x_M, and it has two solutions $y_1(x)$ and $y_2(x)$ for each x between x_m and x_M. The smaller solution $y_1(x)$ is always less than a/b, while the larger solution $y_2(x)$ is always greater than a/b. As x approaches either x_m or x_M, both $y_1(x)$ and $y_2(x)$ approach a/b. Consequently, the curves defined by (3) are closed for x and y positive, and have the form described in Figure 2. Moreover, none of these closed curves (with the exception of $x=c/d$, $y=a/b$) contain any equilibrium points of (1). Therefore, all solutions $x(t)$, $y(t)$ of (1), with $x(0)$ and $y(0)$ positive, are *periodic* functions of time. That is to say, each solution $x(t)$, $y(t)$ of (1), with $x(0)$ and $y(0)$ positive, has the property that $x(t+T)=x(t)$ and $y(t+T)=y(t)$ for some positive T. □

Now, the data of D'Ancona is really an *average* over each one year period of the proportion of predators. Thus, in order to compare this data with the predictions of (1), we must compute the "average values" of $x(t)$ and $y(t)$, for any solution $x(t)$, $y(t)$ of (1). Remarkably, we can find these average values even though we cannot compute $x(t)$ and $y(t)$ exactly. This is the content of Lemma 2.

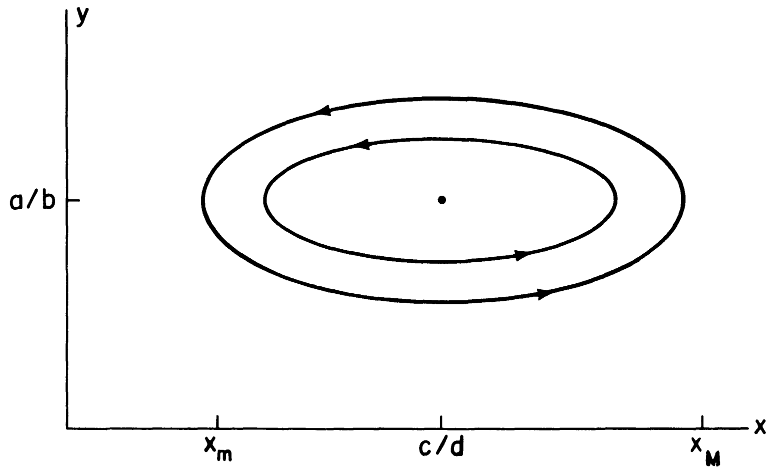

Figure 2. Orbits of (1) for x,y positive

Lemma 2. *Let $x(t), y(t)$ be a periodic solution of* (1), *with period $T>0$. Define the average values of x and y as*

$$\bar{x}=\frac{1}{T}\int_0^T x(t)\,dt, \qquad \bar{y}=\frac{1}{T}\int_0^T y(t)\,dt.$$

Then, $\bar{x}=c/d$ and $\bar{y}=a/b$. In other words, the average values of $x(t)$ and $y(t)$ are the equilibrium values.

PROOF. Dividing both sides of the first equation of (1) by x gives $\dot{x}/x=a-by$, so that

$$\frac{1}{T}\int_0^T \frac{\dot{x}(t)}{x(t)}\,dt=\frac{1}{T}\int_0^T\left[a-by(t)\right]dt.$$

Now, $\int_0^T \dot{x}(t)/x(t)\,dt=\ln x(T)-\ln x(0)$, and this equals zero since $x(T)=x(0)$. Consequently,

$$\frac{1}{T}\int_0^T by(t)\,dt=\frac{1}{T}\int_0^T a\,dt=a,$$

so that $\bar{y}=a/b$. Similarly, by dividing both sides of the second equation of (1) by $Ty(t)$ and integrating from 0 to T, we obtain that $\bar{x}=c/d$. □

We are now ready to include the effects of fishing in our model. Observe that fishing decreases the population of food fish at a rate $\varepsilon x(t)$, and decreases the population of selachians at a rate $\varepsilon y(t)$. The constant ε reflects the intensity of fishing; i.e., the number of boats at sea and the number of nets in the water. Thus, the true state of affairs is described by the modified system of differential equations

$$\begin{aligned}\frac{dx}{dt}&=ax-bxy-\varepsilon x=(a-\varepsilon)x-bxy\\ \frac{dy}{dt}&=-cy+dxy-\varepsilon y=-(c+\varepsilon)y+dxy.\end{aligned}\tag{4}$$

This system is exactly the same as (1) (for $a-\varepsilon>0$), with a replaced by $a-\varepsilon$, and c replaced by $c+\varepsilon$. Hence, the average values of $x(t)$ and $y(t)$ are now

$$\bar{x}=\frac{c+\varepsilon}{d},\qquad \bar{y}=\frac{a-\varepsilon}{b}.\tag{5}$$

Consequently, a moderate amount of fishing ($\varepsilon<a$) actually increases the number of food fish, on the average, and decreases the number of selachians. Conversely, a reduced level of fishing increases the number of selachians, on the average, and *decreases* the number of food fish. This remarkable result, which is known as Volterra's principle, explains the data of D'Ancona, and completely solves our problem.

Volterra's principle has spectacular applications to insecticide treatments, which destroy both insect predators and their insect prey. It implies that the application of insecticides will actually increase the population of those insects which are kept in control by other predatory insects. A remarkable confirmation comes from the cottony cushion scale insect (Icerya purchasi), which, when accidentally introduced from Australia in 1868, threatened to destroy the American citrus industry. Thereupon, its natural Australian predator, a ladybird beetle (Novius Cardinalis) was introduced, and the beetles reduced the scale insects to a low level. When DDT was discovered to kill scale insects, it was applied by the orchardists in the hope of further reducing the scale insects. However, in agreement with Volterra's principle, the effect was an increase of the scale insect!

Oddly enough, many ecologists and biologists refused to accept Volterra's model as accurate. They pointed to the fact that the oscillatory behavior predicted by Volterra's model is not observed in most predator–prey systems. Rather, most predator–prey systems tend to equilibrium states as time evolves. Our answer to these critics is that the system of differential equations (1) is not intended as a model of the general predator–prey interaction. This is because the food fish and selachians do not compete intensively among themselves for their available resources. A more general model of predator–prey interactions is the system of differential equations

$$\dot{x}=ax-bxy-ex^2,\qquad \dot{y}=-cy+dxy-fy^2.\tag{6}$$

Here, the term ex^2 reflects the internal competition of the prey x for their limited external resources, and the term fy^2 reflects the competition among the predators for the limited number of prey. The solutions of (6) are not, in general, periodic. Indeed, we have already shown in Example 1 of Section 4.8 that all solutions $x(t)$, $y(t)$ of (6), with $x(0)$ and $y(0)$ positive, ultimately approach the equilibrium solution $x=a/e$, $y=0$ if c/d is greater than a/e. In this situation, the predators die out, since their available food supply is inadequate for their needs.

Surprisingly, some ecologists and biologists even refuse to accept the more general model (6) as accurate. As a counterexample, they cite the experiments of the mathematical biologist G. F. Gause. In these experiments, the population was composed of two species of protozoa, one of which, Didinium nasatum, feeds on the other, Paramecium caudatum. In all of Gause's experiments, the Didinium quickly destroyed the Paramecium and then died of starvation. This situation cannot be modeled by the system of equations (6), since no solution of (6) with $x(0)y(0)\neq 0$ can reach $x=0$ or $y=0$ in finite time.

Our answer to these critics is that the Didinium are a special, and atypical type of predator. On the one hand, they are ferocious attackers and require a tremendous amount of food; a Didinium demands a fresh Paramecium every three hours. On the other hand, the Didinium don't perish from an insufficient supply of Paramecium. They continue to multiply, but give birth to smaller offspring. Thus, the system of equations (6) does not accurately model the interaction of Paramecium and Didinium. A better model, in this case, is the system of differential equations

$$\frac{dx}{dt}=ax-b\sqrt{x}\,y,\qquad \frac{dy}{dt}=\begin{cases} d\sqrt{x}\,y, & x\neq 0\\ -cy, & x=0\end{cases}. \tag{7}$$

It can be shown (see Exercise 6) that every solution $x(t)$, $y(t)$ of (7) with $x(0)$ and $y(0)$ positive reaches $x=0$ in finite time. This does not contradict the existence–uniqueness theorem, since the function

$$g(x,y)=\begin{cases} d\sqrt{x}\,y, & x\neq 0\\ -cy, & x=0\end{cases}$$

does not have a partial derivative with respect to x or y, at $x=0$.

Finally, we mention that there are several predator–prey interactions in nature which cannot be modeled by any system of ordinary differential equations. These situations occur when the prey are provided with a refuge that is inaccessible to the predators. In these situations, it is impossible to make any definitive statements about the future number of predators and prey, since we cannot predict how many prey will be stupid enough to leave their refuge. In other words, this process is now *random*, rather than *deterministic*, and therefore cannot be modeled by a system of ordinary differential equations. This was verified directly in a famous experiment of Gause. He placed five Paramecium and three Didinium in each of thirty identical test tubes, and provided the Paramecium with a refuge from the Didinium. Two days later, he found the predators dead in four tubes, and a mixed population containing from two to thirty-eight Paramecium in the remaining twenty-six tubes.

Reference

Volterra, V: "Leçons sur la théorie mathématique de la lutte pour la vie." Paris, 1931.

Exercises

1. Find all biologically realistic equilibrium points of (6) and determine their stability.

2. We showed in Section 4.8 that $y(t)$ ultimately approaches zero for all solutions $x(t), y(t)$ of (6), if $c/d > a/e$. Show that there exist solutions $x(t), y(t)$ of (6) for which $y(t)$ increases at first to a maximum value, and then decreases to zero. (To an observer who sees only the predators without noticing the prey, such a case of a population passing through a maximum to total extinction would be very difficult to explain.)

3. In many instances, it is the adult members of the prey who are chiefly attacked by the predators, while the young members are better protected, either by their smaller size, or by their living in a different station. Let x_1 be the number of adult prey, x_2 the number of young prey, and y the number of predators. Then,

$$\dot{x}_1 = -a_1x_1 + a_2x_2 - bx_1y$$
$$\dot{x}_2 = nx_1 - (a_1 + a_2)x_2$$
$$\dot{y} = -cy + dx_1y$$

   where a_2x_2 represents the number of young (per unit time) growing into adults, and n represents the birth rate proportional to the number of adults. Find all equilibrium solutions of this system.

4. There are several situations in nature where species 1 preys on species 2 which in turn preys on species 3. One case of this kind of population is the Island of Komodo in Malaya which is inhabited by giant carnivorous reptiles, and by mammals—their food—which feed on the rich vegetation of the island. We assume that the reptiles have no direct influence on the vegetation, and that only the plants compete among themselves for their available resources. A system of differential equations governing this interaction is

$$\dot{x}_1 = -a_1x_1 - b_{12}x_1x_2 + c_{13}x_1x_3$$
$$\dot{x}_2 = -a_2x_2 + b_{21}x_1x_2$$
$$\dot{x}_3 = a_3x_3 - a_4x_3^2 - c_{31}x_1x_3$$

   Find all equilibrium solutions of this system.

5. Consider a predator–prey system where the predator has alternate means of support. This system can be modelled by the differential equations

$$\dot{x}_1 = \alpha_1x_1(\beta_1 - x_1) + \gamma_1x_1x_2$$
$$\dot{x}_2 = \alpha_2x_2(\beta_2 - x_2) - \gamma_2x_1x_2$$

   where $x_1(t)$ and $x_2(t)$ are the predators and prey populations, respectively, at time t.

(a) Show that the change of coordinates $\beta_i y_i(t) = x_i(t/\alpha_i\beta_i)$ reduces this system of equations to

$$\dot{y}_1 = y_1(1-y_1) + a_1 y_1 y_2, \qquad \dot{y}_2 = y_2(1-y_2) - a_2 y_1 y_2$$

where $a_1 = \gamma_1\beta_2/\alpha_1\beta_1$ and $a_2 = \gamma_2\beta_1/\alpha_2\beta_2$.

(b) What are the stable equilibrium populations when (i) $0 < a_2 < 1$, (ii) $a_2 > 1$?

(c) It is observed that $a_1 = 3a_2$ (a_2 is a measure of the aggressiveness of the predator). What is the value of a_2 if the predator's instinct is to maximize its stable equilibrium population?

6. (a) Let $x(t)$ be a solution of $\dot{x} = ax - M\sqrt{x}$, with $M > a\sqrt{x(t_0)}$. Show that

$$a\sqrt{x} = M - \left(M - a\sqrt{x(t_0)}\right)e^{a(t-t_0)/2}.$$

(b) Conclude from (a) that $x(t)$ approaches zero in finite time.

(c) Let $x(t), y(t)$ be a solution of (7), with $by(t_0) > a\sqrt{x(t_0)}$. Show that $x(t)$ reaches zero in finite time. *Hint*: Observe that $y(t)$ is increasing for $t \geqslant t_0$.

(d) It can be shown that $by(t)$ will eventually exceed $a\sqrt{x(t)}$ for every solution $x(t), y(t)$ of (7) with $x(t_0)$ and $y(t_0)$ positive. Conclude, therefore, that all solutions $x(t)$, $y(t)$ of (7) achieve $x = 0$ in finite time.

4.11 The principle of competitive exclusion in population biology

It is often observed, in nature, that the struggle for existence between two similar species competing for the same limited food supply and living space nearly always ends in the complete extinction of one of the species. This phenomenon is known as the "principle of competitive exclusion." It was first enunciated, in a slightly different form, by Darwin in 1859. In his paper 'The origin of species by natural selection' he writes: "As the species of the same genus usually have, though by no means invariably, much similarity in habits and constitutions and always in structure, the struggle will generally be more severe between them, if they come into competition with each other, than between the species of distinct genera."

There is a very interesting biological explanation of the principle of competitive exclusion. The cornerstone of this theory is the idea of a "niche." A niche indicates what place a given species occupies in a community; i.e., what are its habits, food and mode of life. It has been observed that as a result of competition two similar species rarely occupy the same niche. Rather, each species takes possession of those kinds of food and modes of life in which it has an advantage over its competitor. If the two species tend to occupy the same niche then the struggle for existence between them will be very intense and result in the extinction of the weaker species.

An excellent illustration of this theory is the colony of terns inhabiting the island of Jorilgatch in the Black Sea. This colony consists of four different species of terns: sandwich-tern, common-tern, blackbeak-tern, and little-tern. These four species band together to chase away predators from the colony. However, there is a sharp difference between them as regards the procuring of food. The sandwich-tern flies far out into the open sea to hunt certain species, while the blackbeak-tern feeds exclusively on land. On the other hand, common-tern and little-tern catch fish close to the shore. They sight the fish while flying and dive into the water after them. The little-tern seizes his fish in shallow swampy places, whereas the common-tern hunts somewhat further from shore. In this manner, these four similar species of tern living side by side upon a single small island differ sharply in all their modes of feeding and procuring food. Each has a niche in which it has a distinct advantage over its competitors.

In this section we present a rigorous mathematical proof of the law of competitive exclusion. This will be accomplished by deriving a system of differential equations which govern the interaction between two similar species, and then showing that every solution of the system approaches an equilibrium state in which one of the species is extinct.

In constructing a mathematical model of the struggle for existence between two competing species, it is instructive to look again at the logistic law of population growth

$$\frac{dN}{dt}=aN-bN^2. \tag{1}$$

This equation governs the growth of the population $N(t)$ of a single species whose members compete among themselves for a limited amount of food and living space. Recall (see Section 1.5) that $N(t)$ approaches the limiting population $K=a/b$, as t approaches infinity. This limiting population can be thought of as the maximum population of the species which the microcosm can support. In terms of K, the logistic law (1) can be rewritten in the form

$$\frac{dN}{dt}=aN\left(1-\frac{b}{a}N\right)=aN\left(1-\frac{N}{K}\right)=aN\left(\frac{K-N}{K}\right). \tag{2}$$

Equation (2) has the following interesting interpretation. When the population N is very low, it grows according to the Malthusian law $dN/dt=aN$. The term aN is called the "biotic potential" of the species. It is the potential rate of increase of the species under ideal conditions, and it is realized if there are no restrictions on food and living space, and if the individual members of the species do not excrete any toxic waste products. As the population increases though, the biotic potential is reduced by the factor $(K-N)/K$, which is the relative number of still vacant places in the microcosm. Ecologists call this factor the environmental resistance to growth.

Now, let $N_1(t)$ and $N_2(t)$ be the population at time t of species 1 and 2 respectively. Further, let K_1 and K_2 be the maximum population of species 1 and 2 which the microcosm can support, and let a_1N_1 and a_2N_2 be the biotic potentials of species 1 and 2. Then, $N_1(t)$ and $N_2(t)$ satisfy the system of differential equations

$$\frac{dN_1}{dt}=a_1N_1\left(\frac{K_1-N_1-m_2}{K_1}\right),\qquad \frac{dN_2}{dt}=a_2N_2\left(\frac{K_2-N_2-m_1}{K_2}\right),\tag{3}$$

where m_2 is the total number of places of the first species which are taken up by members of the second species, and m_1 is the total number of places of the second species which are taken up by members of the first species. At first glance it would appear that $m_2=N_2$ and $m_1=N_1$. However, this is not generally the case, for it is highly unlikely that two species utilize the environment in identical ways. Equal numbers of individuals of species 1 and 2 do not, on the average, consume equal quantities of food, take up equal amounts of living space and excrete equal amounts of waste products of the same chemical composition. In general, we must set $m_2=\alpha N_2$ and $m_1=\beta N_1$, for some constants α and β. The constants α and β indicate the degree of influence of one species upon the other. If the interests of the two species do not clash, and they occupy separate niches, then both α and β are zero. If the two species lay claim to the same niche and are very similar, then α and β are very close to one. On the other hand, if one of the species, say species 2, utilizes the environment very unproductively; i.e., it consumes a great deal of food or excretes very poisonous waste products, then one individual of species 2 takes up the place of many individuals of species 1. In this case, then, the coefficient α is very large.

We restrict ourselves now to the case where the two species are nearly identical, and lay claim to the same niche. Then, $\alpha=\beta=1$, and $N_1(t)$ and $N_2(t)$ satisfy the system of differential equations

$$\frac{dN_1}{dt}=a_1N_1\left(\frac{K_1-N_1-N_2}{K_1}\right),\qquad \frac{dN_2}{dt}=a_2N_2\left(\frac{K_2-N_1-N_2}{K_2}\right).\tag{4}$$

In this instance, we expect the struggle for existence between species 1 and 2 to be very intense, and to result in the extinction of one of the species. This is indeed the case as we now show.

Theorem 6 (Principle of competitive exclusion). *Suppose that K_1 is greater than K_2. Then, every solution $N_1(t)$, $N_2(t)$ of (4) approaches the equilibrium solution $N_1=K_1$, $N_2=0$ as t approaches infinity. In other words, if species 1 and 2 are very nearly identical, and the microcosm can support more members of species 1 than of species 2, then species 2 will ultimately become extinct.*

Our first step in proving Theorem 6 is to show that $N_1(t)$ and $N_2(t)$ can never become negative. To this end, recall from Section 1.5 that

$$N_1(t)=\frac{K_1N_1(0)}{N_1(0)+(K_1-N_1(0))e^{-a_1t}},\qquad N_2(t)=0$$

is a solution of (4) for any choice of $N_1(0)$. The orbit of this solution in the N_1–N_2 plane is the point $(0,0)$ for $N_1(0)=0$; the line $0<N_1<K_1$, $N_2=0$ for $0<N_1(0)<K_1$; the point $(K_1,0)$ for $N_1(0)=K_1$; and the line $K_1<N_1<\infty$, $N_2=0$ for $N_1(0)>K_1$. Thus, the N_1 axis, for $N_1\geqslant 0$, is the union of four distinct orbits. Similarly, the N_2 axis, for $N_2\geqslant 0$, is the union of four distinct orbits of (4). This implies that all solutions $N_1(t)$, $N_2(t)$ of (4) which start in the first quadrant $(N_1>0, N_2>0)$ of the N_1–N_2 plane must remain there for all future time.

Our second step in proving Theorem 6 is to split the first quadrant into regions in which both dN_1/dt and dN_2/dt have fixed signs. This is accomplished in the following manner. Let l_1 and l_2 be the lines $K_1-N_1-N_2=0$ and $K_2-N_1-N_2=0$, respectively. Observe that dN_1/dt is negative if (N_1,N_2) lies above l_1, and positive if (N_1,N_2) lies below l_1. Similarly, dN_2/dt is negative if (N_1,N_2) lies above l_2, and positive if (N_1,N_2) lies below l_2. Thus, the two parallel lines l_1 and l_2 split the first quadrant of the N_1–N_2 plane into three regions (see Figure 1) in which both dN_1/dt and dN_2/dt have fixed signs. Both $N_1(t)$ and $N_2(t)$ increase with time (along any solution of (4)) in region I; $N_1(t)$ increases, and $N_2(t)$ decreases, with time in region II; and both $N_1(t)$ and $N_2(t)$ decrease with time in region III.

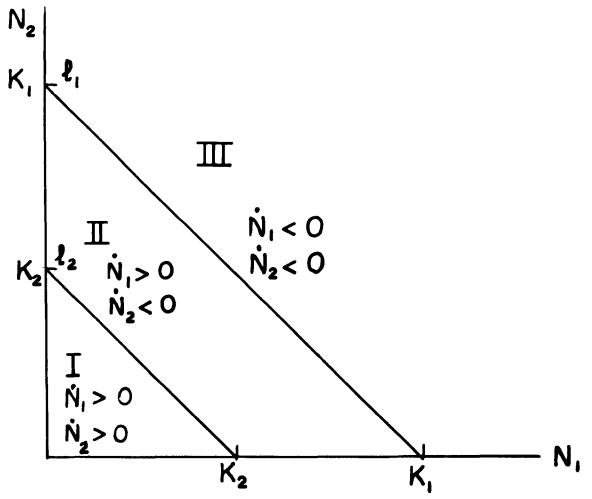

Figure 1

Lemma 1. *Any solution $N_1(t)$, $N_2(t)$ of* (4) *which starts in region* I *at $t=t_0$ must leave this region at some later time.*

PROOF. Suppose that a solution $N_1(t)$, $N_2(t)$ of (4) remains in region I for all time $t\geqslant t_0$. This implies that both $N_1(t)$ and $N_2(t)$ are monotonic increasing functions of time for $t\geqslant t_0$, with $N_1(t)$ and $N_2(t)$ less than K_2. Consequently, by Lemma 1 of Section 4.8, both $N_1(t)$ and $N_2(t)$ have limits

ξ,η respectively, as t approaches infinity. Lemma 2 of Section 4.8 implies that (ξ,η) is an equilibrium point of (4). Now, the only equilibrium points of (4) are $(0,0)$, $(K_1,0)$, and $(0,K_2)$, and (ξ,η) obviously cannot equal any of these three points. We conclude therefore, that any solution $N_1(t)$, $N_2(t)$ of (4) which starts in region I must leave this region at a later time. □

Lemma 2. *Any solution* $N_1(t)$, $N_2(t)$ *of* (4) *which starts in region* II *at time* $t=t_0$ *will remain in this region for all future time* $t\geqslant t_0$, *and ultimately approach the equilibrium solution* $N_1=K_1$, $N_2=0$.

PROOF. Suppose that a solution $N_1(t)$, $N_2(t)$ of (4) leaves region II at time $t=t^*$. Then, either $\dot{N}_1(t^*)$ or $\dot{N}_2(t^*)$ is zero, since the only way a solution of (4) can leave region II is by crossing l_1 or l_2. Assume that $\dot{N}_1(t^*)=0$. Differentiating both sides of the first equation of (4) with respect to t and setting $t=t^*$ gives

$$\frac{d^2N_1(t^*)}{dt^2}=\frac{-a_1N_1(t^*)}{K_1}\frac{dN_2(t^*)}{dt}.$$

This quantity is positive. Hence, $N_1(t)$ has a minimum at $t=t^*$. But this is impossible, since $N_1(t)$ is increasing whenever a solution $N_1(t)$, $N_2(t)$ of (4) is in region II. Similarly, if $\dot{N}_2(t^*)=0$, then

$$\frac{d^2N_2(t^*)}{dt^2}=\frac{-a_2N_2(t^*)}{K_2}\frac{dN_1(t^*)}{dt}.$$

This quantity is negative, implying that $N_2(t)$ has a maximum at $t=t^*$. But this is impossible, since $N_2(t)$ is decreasing whenever a solution $N_1(t)$, $N_2(t)$ of (4) is in region II.

The previous argument shows that any solution $N_1(t)$, $N_2(t)$ of (4) which starts in region II at time $t=t_0$ will remain in region II for all future time $t\geqslant t_0$. This implies that $N_1(t)$ is monotonic increasing and $N_2(t)$ is monotonic decreasing for $t\geqslant t_0$, with $N_1(t)<K_1$ and $N_2(t)>K_2$. Consequently, by Lemma 1 of Section 4.8, both $N_1(t)$ and $N_2(t)$ have limits ξ,η respectively, as t approaches infinity. Lemma 2 of Section 4.8 implies that (ξ,η) is an equilibrium point of (4). Now, (ξ,η) obviously cannot equal $(0,0)$ or $(0,K_2)$. Consequently, $(\xi,\eta)=(K_1,0)$, and this proves Lemma 2. □

Lemma 3. *Any solution* $N_1(t)$, $N_2(t)$ *of* (4) *which starts in region* III *at time* $t=t_0$ *and remains there for all future time must approach the equilibrium solution* $N_1(t)=K_1$, $N_2(t)=0$ *as* t *approaches infinity.*

PROOF. If a solution $N_1(t)$, $N_2(t)$ of (4) remains in region III for $t\geqslant t_0$, then both $N_1(t)$ and $N_2(t)$ are monotonic decreasing functions of time for $t\geqslant t_0$, with $N_1(t)>0$ and $N_2(t)>0$. Consequently, by Lemma 1 of Section 4.8, both $N_1(t)$ and $N_2(t)$ have limits ξ,η respectively, as t approaches infinity. Lemma 2 of Section 4.8 implies that (ξ,η) is an equilibrium point of (4). Now, (ξ,η) obviously cannot equal $(0,0)$ or $(0,K_2)$. Consequently, $(\xi,\eta)=(K_1,0)$. □

Proof of theorem 6. Lemmas 1 and 2 above state that every solution $N_1(t)$, $N_2(t)$ of (4) which starts in regions I or II at time $t=t_0$ must approach the equilibrium solution $N_1=K_1$, $N_2=0$ as t approaches infinity. Similarly, Lemma 3 shows that every solution $N_1(t)$, $N_2(t)$ of (4) which starts in region III at time $t=t_0$ and remains there for all future time must also approach the equilibrium solution $N_1=K_1$, $N_2=0$. Next, observe that any solution $N_1(t)$, $N_2(t)$ of (4) which starts on l_1 or l_2 must immediately afterwards enter region II. Finally, if a solution $N_1(t), N_2(t)$ of (4) leaves region III, then it must cross the line l_1 and immediately afterwards enter region II. Lemma 2 then forces this solution to approach the equilibrium solution $N_1=K_1$, $N_2=0$. □

Theorem 6 deals with the case of identical species; i.e., $\alpha=\beta=1$. By a similar analysis (see Exercises 4–6) we can predict the outcome of the struggle for existence for all values of α and β.

Reference

Gause, G. F., 'The Struggle for Existence,' Dover Publications, New York, 1964.

Exercises

1. Rewrite the system of equations (4) in the form

$$\frac{K_1}{a_1N_1}\frac{dN_1}{dt}=K_1-N_1-N_2, \qquad \frac{K_2}{a_2N_2}\frac{dN_2}{dt}=K_2-N_1-N_2.$$

Then, subtract these two equations and integrate to obtain directly that $N_2(t)$ approaches zero for all solutions $N_1(t)$, $N_2(t)$ of (4) with $N_1(t_0)>0$.

2. The system of differential equations

$$\begin{aligned}\frac{dN_1}{dt}&=N_1[-a_1+c_1(1-b_1N_1-b_2N_2)]\\ \frac{dN_2}{dt}&=N_2[-a_2+c_2(1-b_1N_1-b_2N_2)]\end{aligned} \qquad (*)$$

is a model of two species competing for the same limited resource. Suppose that $c_1>a_1$ and $c_2>a_2$. Deduce from Theorem 6 that $N_1(t)$ ultimately approaches zero if $a_1c_2>a_2c_1$, and $N_2(t)$ ultimately approaches zero if $a_1c_2<a_2c_1$.

3. In 1926, Volterra presented the following model of two species competing for the same limited food supply:

$$\begin{aligned}\frac{dN_1}{dt}&=[b_1-\lambda_1(h_1N_1+h_2N_2)]N_1\\ \frac{dN_2}{dt}&=[b_2-\lambda_2(h_1N_1+h_2N_2)]N_2.\end{aligned}$$

Suppose that $b_1/\lambda_1>b_2/\lambda_2$. (The coefficient b_i/λ_i is called the susceptibility of species i to food shortages.) Prove that species 2 will ultimately become extinct if $N_1(t_0)>0$.

Problems 4–6 are concerned with the system of equations

$$\frac{dN_1}{dt}=\frac{a_1N_1}{K_1}(K_1-N_1-\alpha N_2),\qquad \frac{dN_2}{dt}=\frac{a_2N_2}{K_2}(K_2-N_2-\beta N_1).\quad (*)$$

4. (a) Assume that $K_1/\alpha > K_2$ and $K_2/\beta < K_1$. Show that $N_2(t)$ approaches zero as t approaches infinity for every solution $N_1(t)$, $N_2(t)$ of (*) with $N_1(t_0)>0$.

(b) Assume that $K_1/\alpha < K_2$ and $K_2/\beta > K_1$. Show that $N_1(t)$ approaches zero as t approaches infinity for every solution $N_1(t)$, $N_2(t)$ of (*) with $N_1N_2(t_0)>0$. *Hint*: Draw the lines $l_1: N_1+\alpha N_2=K_1$ and $l_2: N_2+\beta N_1=K_2$, and follow the proof of Theorem 6.

5. Assume that $K_1/\alpha > K_2$ and $K_2/\beta > K_1$. Prove that all solutions $N_1(t)$, $N_2(t)$ of (*), with both $N_1(t_0)$ and $N_2(t_0)$ positive, ultimately approach the equilibrium solution

$$N_1=N_1^0=\frac{K_1-\alpha K_2}{1-\alpha\beta},\qquad N_2=N_2^0=\frac{K_2-\beta K_1}{1-\alpha\beta}.$$

Hint:

(a) Draw the lines $l_1: N_1+\alpha N_2=K_1$ and $l_2: N_2+\beta N_1=K_2$. The two lines divide the first quadrant into four regions (see Figure 2) in which both $\dot N_1$ and $\dot N_2$ have fixed signs.

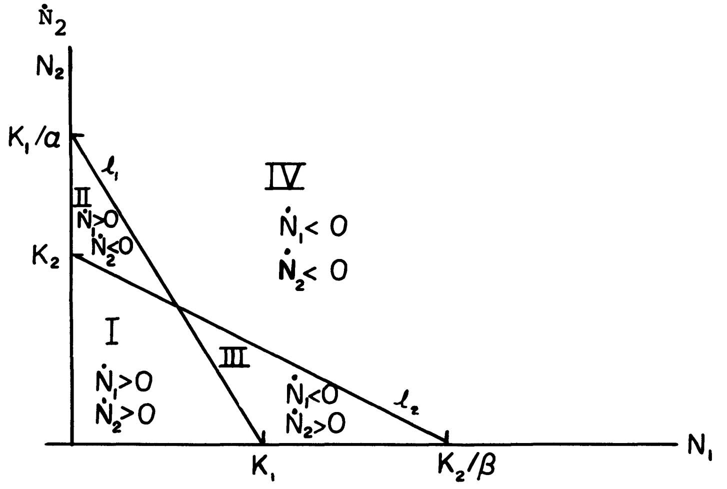

Figure 2

(b) Show that all solutions $N_1(t)$, $N_2(t)$ of (*) which start in either region II or III must remain in these regions and ultimately approach the equilibrium solution $N_1=N_1^0$, $N_2=N_2^0$.

(c) Show that all solutions $N_1(t)$, $N_2(t)$ of (*) which remain exclusively in region I or region IV for all time $t\geqslant t_0$ must ultimately approach the equilibrium solution $N_1=N_1^0$, $N_2=N_2^0$.

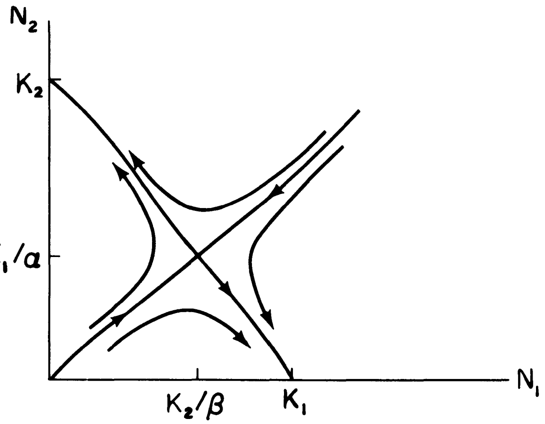

Figure 3

6. Assume that $K_1/\alpha < K_2$ and $K_2/\beta < K_1$.
(a) Show that the equilibrium solution $N_1 = 0$, $N_2 = 0$ of (*) is unstable.
(b) Show that the equilibrium solutions $N_1 = K_1$, $N_2 = 0$ and $N_1 = 0$, $N_2 = K_2$ of (*) are asymptotically stable.
(c) Show that the equilibrium solution $N_1 = N_1^0$, $N_2 = N_2^0$ (see Exercise 5) of (*) is a saddle point. (This calculation is very cumbersome.)
(d) It is not too difficult to see that the phase portrait of (*) must have the form described in Figure 3.

4.12 The Threshold Theorem of epidemiology

Consider the situation where a small group of people having an infectious disease is inserted into a large population which is capable of catching the disease. What happens as time evolves? Will the disease die out rapidly, or will an epidemic occur? How many people will ultimately catch the disease? To answer these questions we will derive a system of differential equations which govern the spread of an infectious disease within a population, and analyze the behavior of its solutions. This approach will also lead us to the famous Threshold Theorem of epidemiology which states that an epidemic will occur only if the number of people who are susceptible to the disease exceeds a certain threshold value.

We begin with the assumptions that the disease under consideration confers permanent immunity upon any individual who has completely recovered from it, and that it has a negligibly short incubation period. This latter assumption implies that an individual who contracts the disease becomes infective immediately afterwards. In this case we can divide the population into three classes of individuals: the infective class (I), the susceptible class (S) and the removed class (R). The infective class consists of those individuals who are capable of transmitting the disease to others.

The susceptible class consists of those individuals who are not infective, but who are capable of catching the disease and becoming infective. The removed class consists of those individuals who have had the disease and are dead, or have recovered and are permanently immune, or are isolated until recovery and permanent immunity occur.

The spread of the disease is presumed to be governed by the following rules.

Rule 1: The population remains at a fixed level N in the time interval under consideration. This means, of course, that we neglect births, deaths from causes unrelated to the disease under consideration, immigration and emigration.

Rule 2: The rate of change of the susceptible population is proportional to the product of the number of members of (S) and the number of members of (I).

Rule 3: Individuals are removed from the infectious class (I) at a rate proportional to the size of (I).

Let $S(t), I(t)$, and $R(t)$ denote the number of individuals in classes (S), (I), and (R), respectively, at time t. It follows immediately from Rules 1–3 that $S(t), I(t), R(t)$ satisfies the system of differential equations

$$\begin{aligned} \frac{dS}{dt} &= -rSI \\ \frac{dI}{dt} &= rSI - \gamma I \\ \frac{dR}{dt} &= \gamma I \end{aligned} \tag{1}$$

for some positive constants r and γ. The proportionality constant r is called the infection rate, and the proportionality constant γ is called the removal rate.

The first two equations of (1) do not depend on R. Thus, we need only consider the system of equations

$$\frac{dS}{dt} = -rSI, \qquad \frac{dI}{dt} = rSI - \gamma I \tag{2}$$

for the two unknown functions $S(t)$ and $I(t)$. Once $S(t)$ and $I(t)$ are known, we can solve for $R(t)$ from the third equation of (1). Alternately, observe that $d(S+I+R)/dt = 0$. Thus,

$$S(t) + I(t) + R(t) = \text{constant} = N$$

so that $R(t) = N - S(t) - I(t)$.

The orbits of (2) are the solution curves of the first-order equation

$$\frac{dI}{dS} = \frac{rSI - \gamma I}{-rSI} = -1 + \frac{\gamma}{rS}. \tag{3}$$

Integrating this differential equation gives

$$I(S) = I_0 + S_0 - S + \rho \ln \frac{S}{S_0}, \tag{4}$$

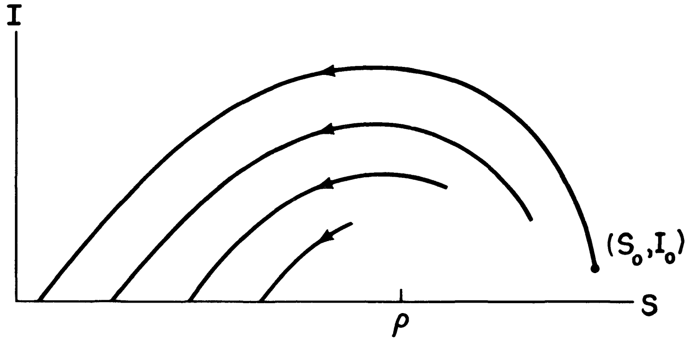

Figure 1. The orbits of (2)

where S_0 and I_0 are the number of susceptibles and infectives at the initial time $t=t_0$, and $\rho=\gamma/r$. To analyze the behavior of the curves (4), we compute $I'(S)=-1+\rho/S$. The quantity $-1+\rho/S$ is negative for $S>\rho$, and positive for $S<\rho$. Hence, $I(S)$ is an increasing function of S for $S<\rho$, and a decreasing function of S for $S>\rho$.

Next, observe that $I(0)=-\infty$ and $I(S_0)=I_0>0$. Consequently, there exists a unique point S_∞, with $0<S_\infty<S_0$, such that $I(S_\infty)=0$, and $I(S)>0$ for $S_\infty<S\leqslant S_0$. The point $(S_\infty,0)$ is an equilibrium point of (2) since both dS/dt and dI/dt vanish when $I=0$. Thus, the orbits of (2), for $t_0\leqslant t<\infty$, have the form described in Figure 1.

Let us see what all this implies about the spread of the disease within the population. As t runs from t_0 to ∞, the point $(S(t),I(t))$ travels along the curve (4), and it moves along the curve in the direction of decreasing S, since $S(t)$ decreases monotonically with time. Consequently, if S_0 is less than ρ, then $I(t)$ decreases monotonically to zero, and $S(t)$ decreases monotonically to S_∞. Thus, if a small group of infectives I_0 is inserted into a group of susceptibles S_0, with $S_0<\rho$, then the disease will die out rapidly. On the other hand, if S_0 is greater than ρ, then $I(t)$ increases as $S(t)$ decreases to ρ, and it achieves a maximum value when $S=\rho$. It only starts decreasing when the number of susceptibles falls below the threshold value ρ. From these results we may draw the following conclusions.

Conclusion 1: An epidemic will occur only if the number of susceptibles in a population exceeds the threshold value $\rho=\gamma/r$.

Conclusion 2: The spread of the disease does not stop for lack of a susceptible population; it stops only for lack of infectives. In particular, some individuals will escape the disease altogether.

Conclusion 1 corresponds to the general observation that epidemics tend to build up more rapidly when the density of susceptibles is high due to overcrowding, and the removal rate is low because of ignorance, inadequate isolation and inadequate medical care. On the other hand, outbreaks tend to be of only limited extent when good social conditions entail lower densities of susceptibles, and when removal rates are high because of good public health vigilance and control.

If the number of susceptibles S_0 is initially greater than, but close to, the threshold value ρ, then we can estimate the number of individuals who ultimately contract the disease. Specifically, if $S_0-\rho$ is small compared to ρ, then the number of individuals who ultimately contract the disease is approximately $2(S_0-\rho)$. This is the famous Threshold Theorem of epidemiology, which was first proven in 1927 by the mathematical biologists Kermack and McKendrick.

Theorem 7 (Threshold Theorem of epidemiology). *Let $S_0=\rho+\nu$ and assume that ν/ρ is very small compared to one. Assume moreover, that the number of initial infectives I_0 is very small. Then, the number of individuals who ultimately contract the disease is 2ν. In other words, the level of susceptibles is reduced to a point as far below the threshold as it originally was above it.*

PROOF. Letting t approach infinity in (4) gives

$$0=I_0+S_0-S_\infty+\rho\ln\frac{S_\infty}{S_0}.$$

If I_0 is very small compared to S_0, then we can neglect it, and write

$$\begin{aligned}0&=S_0-S_\infty+\rho\ln\frac{S_\infty}{S_0}\\&=S_0-S_\infty+\rho\ln\left[\frac{S_0-(S_0-S_\infty)}{S_0}\right]\\&=S_0-S_\infty+\rho\ln\left[1-\left(\frac{S_0-S_\infty}{S_0}\right)\right].\end{aligned}$$

Now, if $S_0-\rho$ is small compared to ρ, then S_0-S_∞ will be small compared to S_0. Consequently, we can truncate the Taylor series

$$\ln\left[1-\left(\frac{S_0-S_\infty}{S_0}\right)\right]=-\left(\frac{S_0-S_\infty}{S_0}\right)-\frac{1}{2}\left(\frac{S_0-S_\infty}{S_0}\right)^2+\dots$$

after two terms. Then,

$$\begin{aligned}0&=S_0-S_\infty-\rho\left(\frac{S_0-S_\infty}{S_0}\right)-\frac{\rho}{2}\left(\frac{S_0-S_\infty}{S_0}\right)^2\\&=(S_0-S_\infty)\left[1-\frac{\rho}{S_0}-\frac{\rho}{2S_0^2}(S_0-S_\infty)\right].\end{aligned}$$

Solving for $S_0 - S_\infty$, we see that

$$S_0 - S_\infty = 2S_0\left(\frac{S_0}{\rho} - 1\right) = 2(\rho+\nu)\left[\frac{\rho+\nu}{\rho} - 1\right]$$

$$= 2(\rho+\nu)\frac{\nu}{\rho} = 2\rho\left(1+\frac{\nu}{\rho}\right)\frac{\nu}{\rho} \cong 2\nu. \qquad \square$$

During the course of an epidemic it is impossible to accurately ascertain the number of new infectives each day or week, since the only infectives who can be recognized and removed from circulation are those who seek medical aid. Public health statistics thus record only the number of new removals each day or week, not the number of new infectives. Therefore, in order to compare the results predicted by our model with data from actual epidemics, we must find the quantity dR/dt as a function of time. This is accomplished in the following manner. Observe first that

$$\frac{dR}{dt} = \gamma I = \gamma(N - R - S).$$

Second, observe that

$$\frac{dS}{dR} = \frac{dS/dt}{dR/dt} = \frac{-rSI}{\gamma I} = \frac{-S}{\rho}.$$

Hence, $S(R) = S_0 e^{-R/\rho}$ and

$$\frac{dR}{dt} = \gamma\left(N - R - S_0 e^{-R/\rho}\right). \tag{5}$$

Equation (5) is separable, but cannot be solved explicitly. However, if the epidemic is not very large, then R/ρ is small and we can truncate the Taylor series

$$e^{-R/\rho} = 1 - \frac{R}{\rho} + \frac{1}{2}\left(\frac{R}{\rho}\right)^2 + \dots$$

after three terms. With this approximation,

$$\frac{dR}{dt} = \gamma\left[N - R - S_0\left[1 - R/\rho + \frac{1}{2}(R/\rho)^2\right]\right]$$

$$= \gamma\left[N - S_0 + \left(\frac{S_0}{\rho} - 1\right)R - \frac{S_0}{2}\left(\frac{R}{\rho}\right)^2\right].$$

The solution of this equation is

$$R(t) = \frac{\rho^2}{S_0}\left[\frac{S_0}{\rho} - 1 + \alpha\tanh\left(\frac{1}{2}\alpha\gamma t - \phi\right)\right] \tag{6}$$

where

$$\alpha=\left[\left(\frac{S_0}{\rho}-1\right)^2+\frac{2S_0(N-S_0)}{\rho^2}\right]^{1/2},\qquad \phi=\tanh^{-1}\frac{1}{\alpha}\left(\frac{S_0}{\rho}-1\right)$$

and the hyperbolic tangent function $\tanh z$ is defined by

$$\tanh z=\frac{e^z-e^{-z}}{e^z+e^{-z}}.$$

It is easily verified that

$$\frac{d}{dz}\tanh z=\operatorname{sech}^2 z=\frac{4}{(e^z+e^{-z})^2}.$$

Hence,

$$\frac{dR}{dt}=\frac{\gamma\alpha^2\rho^2}{2S_0}\operatorname{sech}^2\left(\frac{1}{2}\alpha\gamma t-\phi\right). \tag{7}$$

Equation (7) defines a symmetric bell shaped curve in the t–dR/dt plane (see Figure 2). This curve is called the epidemic curve of the disease. It illustrates very well the common observation that in many actual epidemics, the number of new cases reported each day climbs to a peak value and then dies away again.

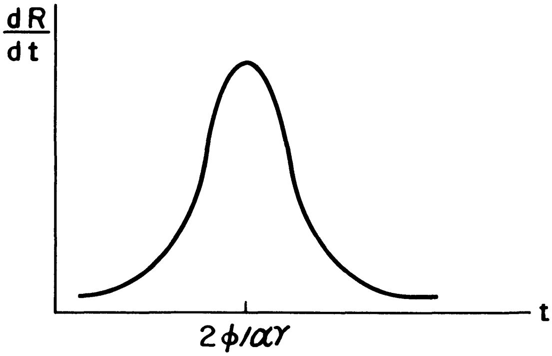

Figure 2

Kermack and McKendrick compared the values predicted for dR/dt from (7) with data from an actual plague in Bombay which spanned the last half of 1905 and the first half of 1906. They set

$$\frac{dR}{dt}=890\operatorname{sech}^2(0.2t-3.4)$$

with t measured in weeks, and compared these values with the number of deaths per week from the plague. This quantity is a very good approximation of dR/dt, since almost all cases terminated fatally. As can be seen from Figure 3, there is excellent agreement between the actual values of

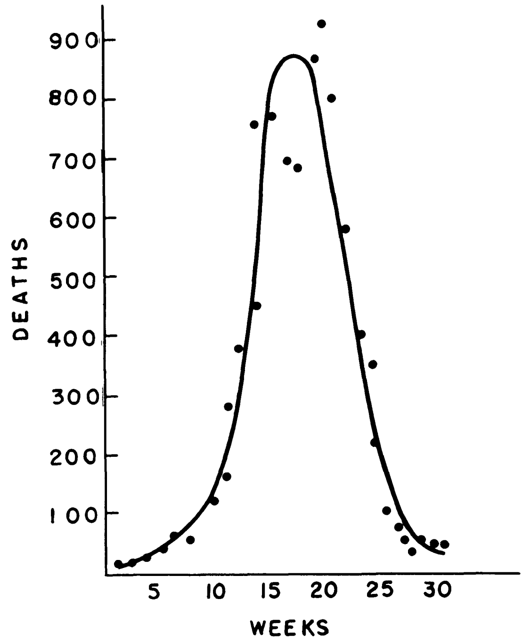

Figure 3

dR/dt, denoted by •, and the values predicted by (7). This indicates, of course, that the system of differential equations (1) is an accurate and reliable model of the spread of an infectious disease within a population of fixed size.

References

Bailey, N. T. J., 'The mathematical theory of epidemics,' 1957, New York.

Kermack, W. O. and McKendrick, A. G., Contributions to the mathematical theory of epidemics, Proceedings Roy. Stat. Soc., A, 115, 700–721, 1927.

Waltman, P., 'Deterministic threshold models in the theory of epidemics,' Springer-Verlag, New York, 1974.

Exercises

1. Derive Equation (6).

2. Suppose that the members of (S) are vaccinated against the disease at a rate λ proportional to their number. Then,

$$\frac{dS}{dt} = -rSI - \lambda S, \qquad \frac{dI}{dt} = rSI - \gamma I. \tag{*}$$

   (a) Find the orbits of (*).
   (b) Conclude from (a) that $S(t)$ approaches zero as t approaches infinity, for every solution $S(t)$, $I(t)$ of (*).

3. Suppose that the members of (S) are vaccinated against the disease at a rate λ proportional to the product of their numbers and the square of the members of (I). Then,

$$\frac{dS}{dt} = -rSI - \lambda SI^2, \qquad \frac{dI}{dt} = I(rS - \gamma). \tag{*}$$

(a) Find the orbits of (*).
(b) Will any susceptibles remain after the disease dies out?

4. The *intensity* i of an epidemic is the proportion of the total number of susceptibles that finally contracts the disease. Show that

$$i = \frac{I_0 + S_0 - S_\infty}{S_0}$$

where S_∞ is a root of the equation

$$S = S_0 e^{(S - S_0 - I_0)/\rho}.$$

5. Compute the intensity of the epidemic if $\rho = 1000$, $I_0 = 10$, and (a) $S_0 = 1100$, (b) $S_0 = 1200$, (c) $S_0 = 1300$, (d) $S_0 = 1500$, (e) $S_0 = 1800$, (f) $S_0 = 1900$. (This cannot be done analytically.)

6. Let R_∞ denote the total number of individuals who contract the disease.
(a) Show that $R_\infty = I_0 + S_0 - S_\infty$.
(b) Let R_1 denote the members of (R) who are removed from the population prior to the peak of the epidemic. Compute R_1/R_∞ for each of the values of S_0 in 5a–5f. Notice that most of the removals occur after the peak. This type of asymmetry is often found in actual notifications of infectious diseases.

7. It was observed in London during the early 1900's, that large outbreaks of measles epidemics recurred about once every two years. The mathematical biologist H. E. Soper tried to explain this phenomenon by assuming that the stock of susceptibles is constantly replenished by new recruits to the population. Thus, he assumed that

$$\frac{dS}{dt} = -rSI + \mu, \qquad \frac{dI}{dt} = rSI - \gamma I \tag{*}$$

for some positive constants r, γ, and μ.
(a) Show that $S = \gamma/r$, $I = \mu/\gamma$ is the only equilibrium solution of (*).
(b) Show that every solution $S(t)$, $I(t)$ of (*) which starts sufficiently close to this equilibrium point must ultimately approach it as t approaches infinity.
(c) It can be shown that every solution $S(t)$, $I(t)$ of (*) approaches the equilibrium solution $S = \gamma/r$, $I = \mu/\gamma$ as t approaches infinity. Conclude, therefore, that the system (*) does not predict recurrent outbreaks of measles epidemics. Rather, it predicts that the disease will ultimately approach a steady state.

4.13 A model for the spread of gonorrhea

Gonorrhea ranks first today among reportable communicable diseases in the United States. There are more reported cases of gonorrhea every year than the combined totals for syphilis, measles, mumps, and infectious hepatitis. Public health officials estimate that more than 2,500,000 Americans contract gonorrhea every year. This painful and dangerous disease, which is caused by the gonococcus germ, is spread from person to person by sexual contact. A few days after the infection there is usually itching and burning of the genital area, particularly while urinating. About the same time a discharge develops which males will notice, but which females may not notice. Infected women may have no easily recognizable symptoms, even while the disease does substantial internal damage. Gonorrhea can only be cured by antibiotics (usually penicillin). However, treatment must be given early if the disease is to be stopped from doing serious damage to the body. If untreated, gonorrhea can result in blindness, sterility, arthritis, heart failure, and ultimately, death.

In this section we construct a mathematical model of the spread of gonorrhea. Our work is greatly simplified by the fact that the incubation period of gonorrhea is very short (3–7 days) compared to the often quite long period of active infectiousness. Thus, we will assume in our model that an individual becomes infective immediately after contracting gonorrhea. In addition, gonorrhea does not confer even partial immunity to those individuals who have recovered from it. Immediately after recovery, an individual is again susceptible. Thus, we can split the sexually active and promiscuous portion of the population into two groups, susceptibles and infectives. Let $c_1(t)$ be the total number of promiscuous males, $c_2(t)$ the total number of promiscuous females, $x(t)$ the total number of infective males, and $y(t)$ the total number of infective females, at time t. Then, the total numbers of susceptible males and susceptible females are $c_1(t)-x(t)$ and $c_2(t)-y(t)$ respectively. The spread of gonorrhea is presumed to be governed by the following rules:

1. Male infectives are cured at a rate a_1 proportional to their total number, and female infectives are cured at a rate a_2 proportional to their total number. The constant a_1 is larger than a_2 since infective males quickly develop painful symptoms and therefore seek prompt medical attention. Female infectives, on the other hand, are usually asymptomatic, and therefore are infectious for much longer periods.

2. New infectives are added to the male population at a rate b_1 proportional to the total number of male susceptibles and female infectives. Similarly, new infectives are added to the female population at a rate b_2 proportional to the total number of female susceptibles and male infectives.

3. The total numbers of promiscuous males and promiscuous females remain at constant levels c_1 and c_2, respectively.

It follows immediately from rules 1–3 that

$$\begin{aligned}\frac{dx}{dt} &= -a_1x+b_1(c_1-x)y\\ \frac{dy}{dt} &= -a_2y+b_2(c_2-y)x.\end{aligned}\tag{1}$$

Remark. The system of equations (1) treats only those cases of gonorrhea which arise from heterosexual contacts; the case of homosexual contacts (assuming no interaction between heterosexuals and homosexuals) is treated in Exercises 5 and 6. The number of cases of gonorrhea which arise from homosexual encounters is a small percentage of the total number of incidents of gonorrhea. Interestingly enough, this situation is completely reversed in the case of syphilis. Indeed, more than 90% of all cases of syphilis reported in the state of Rhode Island during 1973 resulted from homosexual encounters. (This statistic is not as startling as it first appears. Within ten to ninety days after being infected with syphilis, an individual usually develops a chancre sore at the spot where the germs entered the body. A homosexual who contracts syphilis as a result of anal intercourse with an infective will develop a chancre sore on his rectum. This individual, naturally, will be reluctant to seek medical attention, since he will then have to reveal his identity as a homosexual. Moreover, he feels no sense of urgency, since the chancre sore is usually painless and disappears after several days. With gonorrhea, on the other hand, the symptoms are so painful and unmistakable that a homosexual will seek prompt medical attention. Moreover, he need not reveal his identity as a homosexual since the symptoms of gonorrhea appear in the genital area.)

Our first step in analyzing the system of differential equations (1) is to show that they are realistic. Specifically, we must show that $x(t)$ and $y(t)$ can never become negative, and can never exceed c_1 and c_2, respectively. This is the content of Lemmas 1 and 2.

Lemma 1. *If $x(t_0)$ and $y(t_0)$ are positive, then $x(t)$ and $y(t)$ are positive for all $t \geqslant t_0$.*

Lemma 2. *If $x(t_0)$ is less than c_1 and $y(t_0)$ is less than c_2, then $x(t)$ is less than c_1 and $y(t)$ is less than c_2 for all $t \geqslant t_0$.*

PROOF OF LEMMA 1. Suppose that Lemma 1 is false. Let $t^* > t_0$ be the first time at which either x or y is zero. Assume that x is zero first. Then, evaluating the first equation of (1) at $t = t^*$ gives $\dot{x}(t^*) = b_1 c_1 y(t^*)$. This quantity is positive. (Note that $y(t^*)$ cannot equal zero since $x = 0, y = 0$ is an equilibrium solution of (1).) Hence, $x(t)$ is less than zero for t close to, and less than t^*. But this contradicts our assumption that t^* is the first time at which $x(t)$ equals zero. We run into the same contradiction if $y(t^*) = 0$. Thus, both $x(t)$ and $y(t)$ are positive for $t \geqslant t_0$. □

PROOF OF LEMMA 2. Suppose that Lemma 2 is false. Let $t^* > t_0$ be the first time at which either $x = c_1$, or $y = c_2$. Suppose that $x(t^*) = c_1$. Evaluating the first equation of (1) at $t = t^*$ gives $\dot{x}(t^*) = -a_1 c_1$. This quantity is negative. Hence, $x(t)$ is greater than c_1 for t close to, and less than t^*. But this contradicts our assumption that t^* is the first time at which $x(t)$ equals c_1. We run into the same contradiction if $y(t^*)=c_2$. Thus, $x(t)$ is less than c_1 and $y(t)$ is less than c_2 for $t \geq t_0$. □

Having shown that the system of equations (1) is a realistic model of gonorrhea, we now see what predictions it makes concerning the future course of this disease. Will gonorrhea continue to spread rapidly and uncontrollably as the data in Figure 1 seems to suggest, or will it level off eventually? The following extremely important theorem of epidemiology provides the answer to this question.

Theorem 8.

(a) *Suppose that a_1a_2 is less than $b_1b_2c_1c_2$. Then, every solution $x(t)$, $y(t)$ of (1) with $0<x(t_0)<c_1$ and $0<y(t_0)<c_2$, approaches the equilibrium solution*

$$x=\frac{b_1b_2c_1c_2-a_1a_2}{a_1b_2+b_1b_2c_2}, \qquad y=\frac{b_1b_2c_1c_2-a_1a_2}{a_2b_1+b_1b_2c_1}$$

as t approaches infinity. In other words, the total numbers of infective males and infective females will ultimately level off.

(b) *Suppose that a_1a_2 is greater than $b_1b_2c_1c_2$. Then every solution $x(t)$, $y(t)$ of (1) with $0<x(t_0)<c_1$ and $0<y(t_0)<c_2$, approaches zero as t approaches infinity. In other words, gonorrhea will ultimately die out.*

Our first step in proving part (a) of Theorem 8 is to split the rectangle $0<x<c_1$, $0<y<c_2$ into regions in which both dx/dt and dy/dt have fixed signs. This is accomplished in the following manner. Setting $dx/dt=0$ in (1), and solving for y as a function of x gives

$$y=\frac{a_1x}{b_1(c_1-x)}\equiv\phi_1(x).$$

Similarly, setting $dy/dt=0$ in (1) gives

$$x=\frac{a_2y}{b_2(c_2-y)}, \quad \text{or} \quad y=\frac{b_2c_2x}{a_2+b_2x}\equiv\phi_2(x).$$

Observe first that $\phi_1(x)$ and $\phi_2(x)$ are monotonic increasing functions of x; $\phi_1(x)$ approaches infinity as x approaches c_1, and $\phi_2(x)$ approaches c_2 as x approaches infinity. Second, observe that the curves $y=\phi_1(x)$ and $y=\phi_2(x)$ intersect at (0,0) and at (x_0, y_0) where

$$x_0=\frac{b_1b_2c_1c_2-a_1a_2}{a_1b_2+b_1b_2c_2}, \qquad y_0=\frac{b_1b_2c_1c_2-a_1a_2}{a_2b_1+b_1b_2c_1}.$$

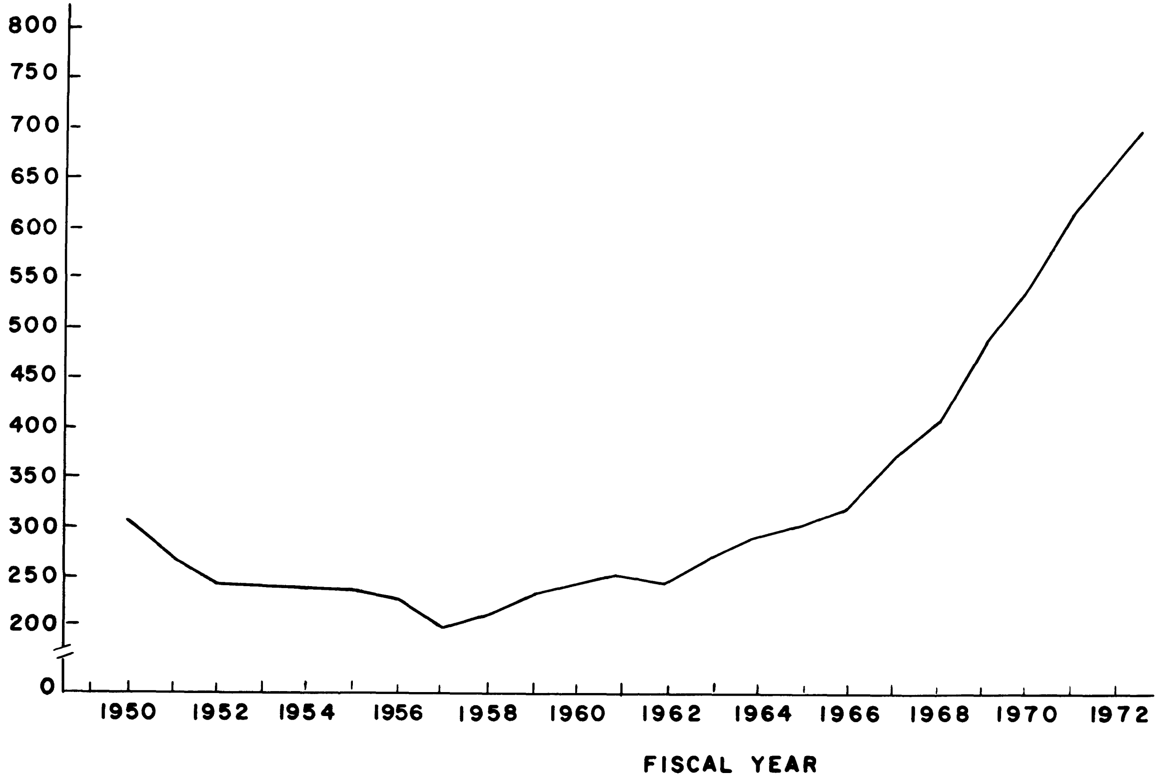

Figure 1. Reported cases of gonorrhea, in thousands, for 1950–1973.

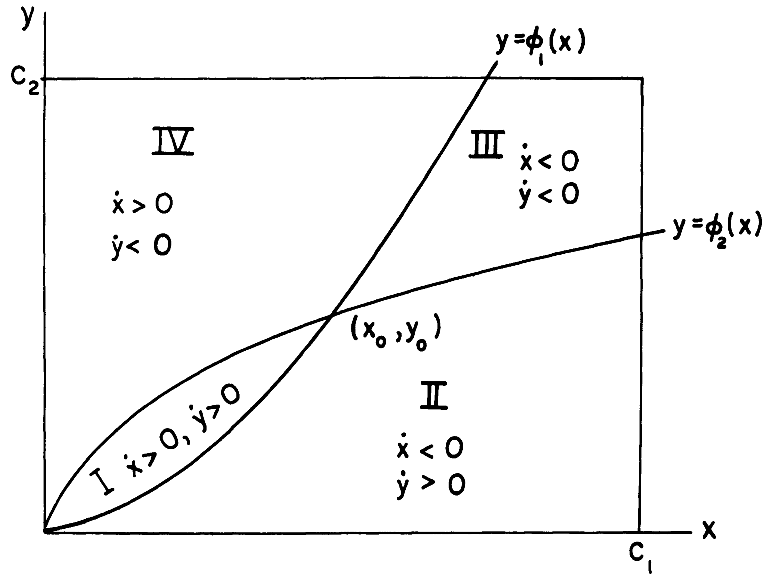

Figure 2

Third, observe that $\phi_2(x)$ is increasing faster than $\phi_1(x)$ at $x=0$, since

$$\phi_2'(0)=\frac{b_2c_2}{a_2}>\frac{a_1}{b_1c_1}=\phi_1'(0).$$

Hence, $\phi_2(x)$ lies above $\phi_1(x)$ for $0<x<x_0$, and $\phi_2(x)$ lies below $\phi_1(x)$ for $x_0<x<c_1$, as shown in Figure 2. The point (x_0, y_0) is an equilibrium point of (1) since both dx/dt and dy/dt are zero when $x=x_0$ and $y=y_0$.

Finally, observe that dx/dt is positive at any point (x,y) above the curve $y=\phi_1(x)$, and negative at any point (x,y) below this curve. Similarly, dy/dt is positive at any point (x,y) below the curve $y=\phi_2(x)$, and negative at any point (x,y) above this curve. Thus, the curves $y=\phi_1(x)$ and $y=\phi_2(x)$ split the rectangle $0<x<c_1$, $0<y<c_2$ into four regions in which dx/dt and dy/dt have fixed signs (see Figure 2).

Next, we require the following four simple lemmas.

Lemma 3. *Any solution $x(t), y(t)$ of (1) which starts in region I at time $t=t_0$ will remain in this region for all future time $t \geqslant t_0$ and approach the equilibrium solution $x=x_0, y=y_0$ as t approaches infinity.*

PROOF. Suppose that a solution $x(t), y(t)$ of (1) leaves region I at time $t=t^*$. Then, either $\dot{x}(t^*)$ or $\dot{y}(t^*)$ is zero, since the only way a solution of (1) can leave region I is by crossing the curve $y=\phi_1(x)$ or $y=\phi_2(x)$. Assume that $\dot{x}(t^*)=0$. Differentiating both sides of the first equation of (1) with respect to t and setting $t=t^*$ gives

$$\frac{d^2x(t^*)}{dt^2}=b_1(c_1-x(t^*))\frac{dy(t^*)}{dt}.$$

This quantity is positive, since $x(t^*)$ is less than c_1, and dy/dt is positive on the curve $y=\phi_1(x)$, $0<x<x_0$. Hence, $x(t)$ has a minimum at $t=t^*$. But this is impossible, since $x(t)$ is increasing whenever the solution $x(t),y(t)$ is in region I. Similarly, if $\dot{y}(t^*)=0$, then

$$\frac{d^2y(t^*)}{dt^2}=b_2(c_2-y(t^*))\frac{dx(t^*)}{dt}.$$

This quantity is positive, since $y(t^*)$ is less than c_2, and dx/dt is positive on the curve $y=\phi_2(x)$, $0<x<x_0$. Hence, $y(t)$ has a minimum at $t=t^*$. But this is impossible, since $y(t)$ is increasing whenever the solution $x(t),y(t)$ is in region I.

The previous argument shows that any solution $x(t),y(t)$ of (1) which starts in region I at time $t=t_0$ will remain in region I for all future time $t\geqslant t_0$. This implies that $x(t)$ and $y(t)$ are monotonic increasing functions of time for $t\geqslant t_0$, with $x(t)<x_0$ and $y(t)<y_0$. Consequently, by Lemma 1 of Section 4.8, both $x(t)$ and $y(t)$ have limits ξ,η, respectively, as t approaches infinity. Lemma 2 of Section 4.8 implies that (ξ,η) is an equilibrium point of (1). Now, it is easily seen from Figure 2 that the only equilibrium points of (1) are $(0,0)$ and (x_0,y_0). But (ξ,η) cannot equal $(0,0)$ since both $x(t)$ and $y(t)$ are increasing functions of time. Hence, $(\xi,\eta)=(x_0,y_0)$, and this proves Lemma 3. □

Lemma 4. *Any solution $x(t),y(t)$ of* (1) *which starts in region* III *at time $t=t_0$ will remain in this region for all future time and ultimately approach the equilibrium solution $x=x_0$, $y=y_0$.*

PROOF. Exactly the same as Lemma 3 (see Exercise 1). □

Lemma 5. *Any solution $x(t),y(t)$ of* (1) *which starts in region* II *at time $t=t_0$, and remains in region* II *for all future time, must approach the equilibrium solution $x=x_0$, $y=y_0$ as t approaches infinity.*

PROOF. If a solution $x(t)$, $y(t)$ of (1) remains in region II for $t\geqslant t_0$, then $x(t)$ is monotonic decreasing and $y(t)$ is monotonic increasing for $t\geqslant t_0$. Moreover, $x(t)$ is positive and $y(t)$ is less than c_2, for $t\geqslant t_0$. Consequently, by Lemma 1 of Section 4.8, both $x(t)$ and $y(t)$ have limits ξ,η respectively, as t approaches infinity. Lemma 2 of Section 4.8 implies that (ξ,η) is an equilibrium point of (1). Now, (ξ,η) cannot equal $(0,0)$ since $y(t)$ is increasing for $t\geqslant t_0$. Therefore, $(\xi,\eta)=(x_0,y_0)$, and this proves Lemma 5. □

Lemma 6. *Any solution $x(t), y(t)$ of* (1) *which starts in region* IV *at time $t = t_0$ and remains in region* IV *for all future time, must approach the equilibrium solution $x = x_0, y = y_0$ as t approaches infinity.*

PROOF. Exactly the same as Lemma 5 (see Exercise 2). □

We are now in a position to prove Theorem 8.

PROOF OF THEOREM 8. (a) Lemmas 3 and 4 state that every solution $x(t)$, $y(t)$ of (1) which starts in region I or III at time $t = t_0$ must approach the equilibrium solution $x = x_0$, $y = y_0$ as t approaches infinity. Similarly, Lemmas 5 and 6 state that every solution $x(t), y(t)$ of (1) which starts in region II or IV and which remains in these regions for all future time, must also approach the equilibrium solution $x = x_0, y = y_0$. Now, observe that if a solution $x(t)$, $y(t)$ of (1) leaves region II or IV, then it must cross the curve $y = \phi_1(x)$ or $y = \phi_2(x)$, and immediately afterwards enter region I or region III. Consequently, all solutions $x(t)$, $y(t)$ of (1) which start in regions II and IV or on the curves $y = \phi_1(x)$ and $y = \phi_2(x)$, must also approach the equilibrium solution $x(t) = x_0, y(t) = y_0$. □

(b) PROOF #1. If a_1a_2 is greater than $b_1b_2c_1c_2$, then the curves $y = \phi_1(x)$ and $y = \phi_2(x)$ have the form described in Figure 3 below. In region I, dx/dt is positive and dy/dt is negative; in region II, both dx/dt and dy/dt are negative; and in region III, dx/dt is negative and dy/dt is positive. It is a simple matter to show (see Exercise 3) that every solution $x(t), y(t)$ of (1) which starts in region II at time $t = t_0$ must remain in this region for all

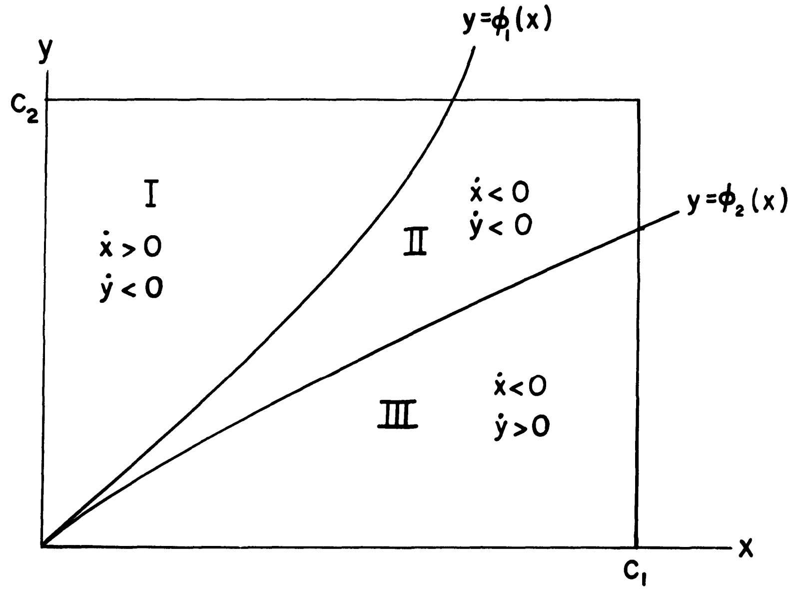

Figure 3

future time, and approach the equilibrium solution $x=0$, $y=0$ as t approaches infinity. It is also trivial to show that every solution $x(t), y(t)$ of (1) which starts in region I or region III at time $t=t_0$ must cross the curve $y=\phi_1(x)$ or $y=\phi_2(x)$, and immediately afterwards enter region II (see Exercise 4). Consequently, every solution $x(t), y(t)$ of (1), with $0<x(t_0)<c_1$ and $0<y(t_0)<c_2$, approaches the equilibrium solution $x=0, y=0$ as t approaches infinity. □

PROOF #2. We would now like to show how we can use the Poincaré–Bendixson theorem to give an elegant proof of part (b) of Theorem 8. Observe that the system of differential equations (1) can be written in the form

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -a_1 & b_1c_1 \\ b_2c_2 & -a_2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} b_1xy \\ b_2xy \end{pmatrix}. \tag{2}$$

Thus, by Theorem 2 of Section 4.3, the stability of the solution $x=0, y=0$ of (2) is determined by the stability of the equilibrium solution $x=0, y=0$ of the linearized system

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{A}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -a_1 & b_1c_1 \\ b_2c_2 & -a_2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

The characteristic polynomial of the matrix $\mathbf{A}$ is

$$\lambda^2+(a_1+a_2)\lambda+a_1a_2-b_1b_2c_1c_2$$

whose roots are

$$\lambda = \frac{-(a_1+a_2)\pm\left[(a_1+a_2)^2-4(a_1a_2-b_1b_2c_1c_2)\right]^{1/2}}{2}.$$

It is easily verified that both these roots are real and negative. Hence, the equilibrium solution $x=0, y=0$ of (2) is asymptotically stable. This implies that any solution $x(t), y(t)$ of (1) which starts sufficiently close to the origin $x=y=0$ will approach the origin as t approaches infinity. Now, suppose that a solution $x(t), y(t)$ of (1), with $0<x(t_0)<c_1$ and $0<y(t_0)<c_2$, does not approach the origin as t approaches infinity. By the previous remark, this solution must always remain a minimum distance from the origin. Consequently, its orbit for $t\geqslant t_0$ lies in a bounded region in the $x-y$ plane which contains no equilibrium points of (1). By the Poincaré–Bendixson Theorem, therefore, its orbit must spiral into the orbit of a periodic solution of (1). But the system of differential equations (1) has no periodic solution in the first quadrant $x\geqslant 0$, $y\geqslant 0$. This follows immediately from Exercise 11, Section 4.8, and the fact that

$$\frac{\partial}{\partial x}\left[-a_1x+b_1(c_1-x)y\right]+\frac{\partial}{\partial y}\left[-a_2y+b_2(c_2-y)x\right]$$
$$=-(a_1+a_2+b_1y+b_2x)$$

is strictly negative if both x and y are nonnegative. Consequently, every solution $x(t)$, $y(t)$ of (1), with $0 < x(t_0) < c_1$ and $0 < y(t_0) < c_2$ approaches the equilibrium solution $x = 0, y = 0$ as t approaches infinity. □

Now, it is quite difficult to evaluate the coefficients a_1, a_2, b_1, b_2, c_1, and c_2. Indeed, it is impossible to obtain even a crude estimate of a_2, which should be interpreted as the average amount of time that a female remains infective. (Similarly, a_1 should be interpreted as the average amount of time that a male remains infective.) This is because most females do not exhibit symptoms. Thus, a female can be infective for an amount of time varying from just one day to well over a year. Nevertheless, it is still possible to ascertain from public health data that $a_1 a_2$ is less than $b_1 b_2 c_1 c_2$, as we now show. Observe that the condition $a_1 a_2 < b_1 b_2 c_1 c_2$ is equivalent to

$$1 < \left(\frac{b_1 c_1}{a_2} \right) \left(\frac{b_2 c_2}{a_1} \right).$$

The quantity $b_1 c_1 / a_2$ can be interpreted as the average number of males that one female infective contacts during her infectious period, if every male is susceptible. Similarly, the quantity $b_2 c_2 / a_1$ can be interpreted as the average number of females that one male infective contacts during his infectious period, if every female is susceptible. The quantities $b_1 c_1 / a_2$ and $b_2 c_2 / a_1$ are called the maximal female and male contact rates, respectively. Theorem 8 can now be interpreted in the following manner.

(a) If the product of the maximal male and female contact rates is greater than one, then gonorrhea will approach a nonzero steady state.
(b) If the product of the maximal male and female contact rates is less than one, then gonorrhea will die out eventually.

In 1973, the average number of female contacts named by a male infective during his period of infectiousness was 0.98, while the average number of male contacts named by a female infective during her period of infectiousness was 1.15. These numbers are very good approximations of the maximal male and female contact rates, respectively, and their product does not exceed the product of the maximal male and female contact rates. (The number of contacts of a male or female infective during their period of infectiousness is slightly less than the maximal male or female contact rates. However, the *actual* number of contacts is often greater than the number of contacts named by an infective.) The product of 1.15 with 0.98 is 1.0682. Thus, gonorrhea will ultimately approach a nonzero steady state.

Remark. Our model of gonorrhea is rather crude since it lumps all promiscuous males and all promiscuous females together, regardless of age. A more accurate model can be obtained by separating the male and female populations into different age groups and then computing the rate of change of infectives in each age group. This has been done recently, but the analysis is too difficult to present here. We just mention that a result completely analogous to Theorem 8 is obtained: either gonorrhea dies out in each age group, or it approaches a constant, positive level in each age group.

Exercises

In Problems 1 and 2, we assume that $a_1a_2 < b_1b_2c_1c_2$.

1. (a) Suppose that a solution $x(t), y(t)$ of (1) leaves region III of Figure 2 at time $t = t^*$ by crossing the curve $y = \phi_1(x)$ or $y = \phi_2(x)$. Conclude that either $x(t)$ or $y(t)$ has a maximum at $t = t^*$. Then, show that this is impossible. Conclude, therefore, that any solution $x(t), y(t)$ of (1) which starts in region III at time $t = t_0$ must remain in region III for all future time $t > t_0$.
   (b) Conclude from (a) that any solution $x(t), y(t)$ of (1) which starts in region III has a limit ξ, η as t approaches infinity. Then, show that (ξ, η) must equal (x_0, y_0).

2. Suppose that a solution $x(t), y(t)$ of (1) remains in region IV of Figure 2 for all time $t \geqslant t_0$. Prove that $x(t)$ and $y(t)$ have limits ξ, η respectively, as t approaches infinity. Then conclude that (ξ, η) must equal (x_0, y_0).

In Problems 3 and 4, we assume that $a_1a_2 > b_1b_2c_1c_2$.

3. Suppose that a solution $x(t), y(t)$ of (1) leaves region II of Figure 3 at time $t = t^*$ by crossing the curve $y = \phi_1(x)$ or $y = \phi_2(x)$. Show that either $x(t)$ or $y(t)$ has a maximum at $t = t^*$. Then, show that this is impossible. Conclude, therefore, that every solution $x(t), y(t)$ of (1) which starts in region II at time $t = t_0$ must remain in region II for all future time $t \geqslant t_0$.

4. (a) Suppose that a solution $x(t), y(t)$ of (1) remains in either region I or III of Figure 3 for all time $t \geqslant t_0$. Show that $x(t)$ and $y(t)$ have limits ξ, η respectively, as t approaches infinity.
   (b) Conclude from Lemma 1 of Section 4.8 that $(\xi, \eta) = (0, 0)$.
   (c) Show that (ξ, η) cannot equal $(0, 0)$ if $x(t), y(t)$ remains in region I or region III for all time $t \geqslant t_0$.
   (d) Show that any solution $x(t), y(t)$ of (1) which starts on either $y = \phi_1(x)$ or $y = \phi_2(x)$ will immediately afterwards enter region II.

5. Assume that $a_1a_2 < b_1b_2c_1c_2$. Prove directly, using Theorem 2 of Section 4.3, that the equilibrium solution $x = x_0, y = y_0$ of (1) is asymptotically stable. *Warning*: The calculations are extremely tedious.

6. Assume that the number of homosexuals remains constant in time. Call this constant c. Let $x(t)$ denote the number of homosexuals who have gonorrhea at time t. Assume that homosexuals are cured of gonorrhea at a rate α_1, and that new infectives are added at a rate $\beta_1(c - x)x$.
   (a) Show that $\dot{x} = -\alpha_1 x + \beta_1 x(c - x)$.
   (b) What happens to $x(t)$ as t approaches infinity?

7. Suppose that the number of homosexuals $c(t)$ grows according to the logistic law $\dot{c} = c(a - bc)$, for some positive constants a and b. Let $x(t)$ denote the number of homosexuals who have gonorrhea at time t, and assume (see Problem 6) that $\dot{x} = -\alpha_1 x + \beta_1 x(c - x)$. What happens to $x(t)$ as t approaches infinity?