People's Democratic Republic of Algeria

Ministry of Higher Education and scientific Research

University Center Abdelhafid Boussouf-MILA

Institute of mathematics and computer sciences

Department of mathematics

Course Of Introduction To Dynamics Systems.

Modélisation mathématiques et technique de décision

By: Rami AMIRA

University year: 2024 – 2025

CONTENTS

INTRODUCTION

This course covers those topics necessary for a clear understanding of the qualitative theory of ordinary differential equations and the concept of a dynamical system. It is written for ferst yers master students.

It begins with a study of linear systems of ordinary differential equations, a topic already familiar to the student who has completed a first course in differential equations. An efficient method for solving any linear system of ordinary differential equations is presented in Chapter 1. The major part of this course is devoted to a study of nonlinear systems of ordinary differential equations and dynamical systems. Since most nonlinear differential equations cannot be solved, this course focuses on the qualitative or geometrical theory of nonlinear systems of differential equations originated by Henri Poincarc in his work on differential equations at the end of the nineteenth century as well as on the functional properties inherent in the solution set of a system of nonlinear differential equations embodied in the more recent concept of a dynamical system. Our primary goal is to describe the qualitative behavior of the solution set of a given system of differential equations including the invariant sets and limiting behavior of the dynamical system or flow defined by the system of differential equations. In order to achieve this goal, it is first necessary to develop the local theory for nonlinear systems. This is done in Chapter 2 which includes the fundamental local existence-uniqueness theorem, the Hartman-Grobman Theorem and the Stable Manifold Theorem. These latter two theorems establish that the qualitative behavior of the solution set of a nonlinear system of ordinary differential equations near an equilibrium point is typically the same as the qualitative behavior of the solution set of the corresponding linearized system near the equilibrium point.

CHAPTER 1

CONTINUOUS DYNAMICAL SYSTEMS

1.1 Dynamical Systems

Dynamics is primarily the study of the time-evolutionary process and the corresponding system of equations is known as dynamical system. Generally, a system of n first-order differential equations in the space $\mathbb{R}^*$ is called a dynamical system of dimension n which determines the time behavior of evolutionary process. Evolutionary processes may possess the properties of determinacy/non-determinacy, finite/infinite dimensionality, and differentiability. A process is called deterministic if its entire future course and its entire past are uniquely determined by its state at the present time. Otherwise, the process is called nondeterministic. However, the process may be semi-deterministic (determined, but not uniquely). In classical mechanics the motion of a system whose future and past are uniquely determined by the initial positions and the initial velocities is an example of a deterministic dynamical system. The evolutionary process may describe, viz. (i) a continuous-time process and (ii) a discrete-time process. The continuous-time process is represented by differential equations, whereas the discrete-time process is by difference equations (or maps). The continuous-time dynamical systems may be described mathematically as follows:

Let $x = x(t) \in \mathbb{R}^n, t \in I \subseteq \mathbb{R}$ be the vector representing the dynamics of a continuous system (continuous-time system). The mathematical representation of the system may be written as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \dot{x} = f(x, t) \tag{1.1}$$

where $f(x,t)$ is a sufficiently smooth function defined on some subset $U \subset \mathbb{R}^n \times \mathbb{R}$. Schematically, this can be shown as

$$\underset{\text{(state space)}}{\mathbb{R}^n} \times \underset{\text{(time)}}{\mathbb{R}} = \underset{\text{(space of motions)}}{\mathbb{R}^{n+1}}$$

The variable t is usually interpreted as time and the function $f(x,t)$ is generally nonlinear. The time interval may be finite, semi-finite or infinite. On the other hand, the discrete system is related to a discrete map (given only at equally spaced points of time) such that from a point x_0, one can obtain a point x_1 which in turn maps into x_2, and so on. In other words, $x_{n+1} = g(x_n) = g(g(x_{n-1}))$, etc. This is also written in the form $x_{n+1} = g(x_n) = g^2(x_{n-1}) = \cdots$. The discrete system will be discussed in the later course .

If the right-hand side of Eq.1.1 is explicitly time independent then the system is called **autonomous**. The trajectories of such a system do not change in time. On the other hand, if the right-hand side of Eq. 1.1 has explicit dependence on time then the system is called **nonautonomous**.

An n-dimensional nonautonomous system can be converted into autonomous form by introducing a new dependent variable x_{n+1} such that $x_{n+1} = t$. In general, the solution of Eq. 1.1 is difficult or sometimes impossible to obtain when the function $f(x,t)$ is nonlinear, except in some special cases. Examples of autonomous and nonautonomous systems are given below.

(i) Autonomous systems

(a) $\ddot{x} + \alpha\dot{x} + \beta x = 0, \alpha, \beta > 0$. This is a damped linear harmonic oscillator. The parameters α and β are, respectively, the strength of damping and the strength of linear restoring force.

(b) $\ddot{x} + \omega^2 \sin x = 0, \omega = \sqrt{g/L}$.g is the gravitational acceleration, L the string length. This is a simple undamped nonlinear oscillator (pendulum).

(c) $\left.\begin{aligned} \dot{x} &= \alpha x - \beta xy \\ \dot{y} &= -\gamma y + \delta xy \end{aligned}\right\}$. . This is the well-known Lotka-Volterra predator-prey model, where $\alpha, \beta, \gamma, \delta$ are all positive constants.

(d) $\ddot{x} - \mu\left(1 - x^2\right)\dot{x} + \beta x = 0, \mu > 0$. This is the well-known van der Pol oscillator.

(ii) Nonautonomous systems

(a) $\ddot{x} + \alpha\dot{x} + \beta x = f \cos \omega t, \alpha, \beta > 0$. This is an example of linear oscillator with external time-dependent force. f and ω are the amplitude and frequency of driving force, respectively.

(b) $\ddot{x} + \alpha\dot{x} + \omega_0^2 x + \beta x^3 = f \sin \omega t$. This is a Duffing nonlinear oscillator with cubic restoring force. α is the strength of damping, ω_0 is the natural frequency and β is the strength of the nonlinear restoring force.

(c) $\ddot{x} - \mu\left(1 - x^2\right)\dot{x} + \beta x = f \cos \omega t, \mu > 0$. This is a van der Pol nonlinear forced oscillator.

(d) $\ddot{x} - \mu\left(1 - x^2\right)\dot{x} + \omega_0^2 x + \beta x^3 = f \cos \omega x$. This is a Duffing-van der Pol nonlinear forced oscillator.

1.2 Flows

The time-evolutionary process may be described as a flow of a vector field.
Generally, flow is frequently used for discussing the dynamics as a whole rather than the evolution of a system at a particular point. The solution $x(t)$ of a system $\dot{x} = f(x)$ which satisfies $x(t_0) = x_0$ gives the past ($t < t_0$) and future ($t > t_0$) evolutions of the system. Mathematically, the flow is defined by $\phi_t(x) : U \to \mathbb{R}^n$ where $\phi_t(x) = \phi(t, x)$ is a smooth vector function of $x \in U \subseteq \mathbb{R}^n$ and $t \in I \subseteq \mathbb{R}$ satisfying the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_t(x) = f\left(\phi_t(x)\right)$$

for all t such that the solution through x exists and $\phi(0, x) = x$. The flow $\phi_t(x)$ satisfies the following properties:
(a) $\phi_o = I_d$,
(b) $\phi_{t+s} = \phi_t \circ \phi_s$.
Some flows may also satisfy the property (c)

$$\phi(t + s, x) = \phi(t, \phi(s, x)) = \phi(s, \phi(t, x)) = \phi(s + t, x).$$

Flows in $\mathbb{R}$: Consider a one-dimensional autonomous system represented by $\dot{x} = f(x), x \in \mathbb{R}$. We can imagine that a fluid is flowing along the real line with local velocity $f(x)$. This imaginary fluid is called **the phase fluid** and the real line is called the **phase line**.

For solution of the system $\dot{x} = f(x)$ starting from an arbitrary initial position x_0, we place an imaginary particle, called a **phase point**, at x_0 and watch how it moves along with the flow in phase line in varying time t. As time goes on, the phase point (x, t) in the one-dimensional system $\dot{x} = f(x)$ with $x(0) = x_0$ moves along the x-axis according to some function $\phi(t, x_0)$. The function $\phi(t, x_0)$ is called the **trajectory** for a given initial state x_0, and the set $\left\{\phi(t, x_0) \mid t \in I \subseteq \mathbb{R}\right\}$ is the orbit of $x_0 \in \mathbb{R}$. The set of all qualitative trajectories of the system is called phase portrait.
Flows in $\mathbb{R}^2$: Consider a two-dimensional system represented by the following equations $\dot{x} = f(x, y), \dot{y} = g(x, y), (x, y) \in \mathbb{R}^2$. An imaginary fluid particle flows in the plane $\mathbb{R}^2$, known as phase plane of the system. The succession of states given parametrically by $x = x(t), y = y(t)$ trace out a curve through some initial point $P(x(t_0), y(t_0))$ is called a **phase path**.
The set $\left\{\phi(t, x_0) \mid t \in I \subseteq \mathbb{R}\right\}$ is the orbit of x in $\mathbb{R}^2$. There are an infinite number of trajectories that would fill the phase plane when they are plotted. But the qualitative behavior can be determined by plotting a few trajectories with different initial conditions. The phase portrait displays how the qualitative behavior of a system is changing as x and y varies with time t. An orbit is called periodic if $x(t + p) = x(t)$ for some $p > 0$, for all t. The smallest integer p for which the relation is satisfied is called the prime period of the

orbit. Flows in $\mathbb{R}$ cannot have oscillatory or closed path.

Flows in $\mathbb{R}^n$: Let us now define an autonomous system representing n ordinary differential equations as

$$\left.\begin{aligned} \dot{x}_1 &= f_1\left(x_1, x_2, \ldots, x_n\right) \\ \dot{x}_2 &= f_2\left(x_1, x_2, \ldots, x_n\right) \\ &\vdots \\ \dot{x}_n &= f_n\left(x_1, x_2, \ldots, x_n\right) \end{aligned}\right\}$$

which can also be written in symbolic notation as $\dot{x} = f(x)$, where $x = (x_1, x_2, \ldots, x_n)$ and $f = (f_1, f_2, \ldots, f_n)$. The solution of this system with the initial condition $x(0) = x_0$ can be thought as a continuous curve in the phase space $\mathbb{R}^n$ parameterized by time $t \in I \subseteq \mathbb{R}$.

So the set of all states of the evolutionary process is represented by an n-valued vector field in $\mathbb{R}^n$. The solutions of the system with different initial conditions describe a family of phase curves in the phase space, called the phase portrait of the system. The vector field $f(x)$ is everywhere tangent to these curves and their orientation is directed by the direction of the tangent vector of $f(x)$.

1.3 Evolution

Consider a system $\dot{x} = f(x), x \in \mathbb{R}^n$ with initial conditions $x\left(t_0\right) = x_0$. Let $E \subset \mathbb{R}^n$ be an open set and $f \in C^1(E)$. For $x_0 \in E$, let $\phi\left(t, x_0\right)$ be a solution of the above system on the maximum interval of existence $I\left(x_0\right) \subset \mathbb{R}$. The mapping $\phi_t : \mathbb{R}^n \to \mathbb{R}^n$ defined by $\phi_t(x_0) = \phi(t, x_0)$ is known as **evolution operator** of the system.

The linear flow for the system $\dot{x} = Ax$ with $x\left(t_0\right) = x_0$, is defined by $\phi_t : \mathbb{R}^n \to \mathbb{R}^n$ and $\phi_t = e^{At}$, the exponential matrix. The mappings ϕ_t for both linear and nonlinear systems satisfy the following properties:

(i) $\phi_0(x) = x$

(ii) $\phi_s\left(\phi_t(x)\right) = \phi_{s+t}(x), \forall s, t \in \mathbb{R}$

(iii) $\phi_t\left(\phi_{-t}(x)\right) = \phi_{-t}\left(\phi_t(x)\right) = x, \forall t \in \mathbb{R}$

In general a dynamical system may be viewed as group of nonlinear / linear operators evolving as $\left\{\phi_t(x), t \in \mathbb{R}, x \in \mathbb{R}^n\right\}$. The following dynamical group properties hold good:

(i) $\phi_t\phi_s \in \left\{\phi_t(x), t \in \mathbb{R}, x \in \mathbb{R}^n\right\}$ (Closure property)

(ii) $\phi_t\left(\phi_s\phi_r\right) = (\phi_t\phi_s)\phi_r$ (Associative property)

(iii) $\phi_0(x) = x, \phi_0$ being the Identity operator.

(iv) $\phi_t\phi_{-t} = \phi_{-t}\phi_t = \phi_0$, where ϕ_{-t} is the Inverse of ϕ_t.

For some cases the flow satisfies the commutative property $\phi_t\phi_s = \phi_s\phi_t$.

1.4 Fixed Points of a System

The notion of fixed point is important in analyzing the local behavior of a system. The fixed point is nothing but a constant or equilibrium or invariant solution of a system. A point is a fixed point of the flow generated by an autonomous system $\dot{x} = f(x), x \in \mathbb{R}^n$ if and only if $\phi(t, x) = x$ for all $t \in \mathbb{R}$. Consequently in continuous system, this gives $\dot{x} = 0 \Rightarrow f(x) = 0$. For nonautonomous systems fixed point can be defined for a fixed time interval. $\tilde{A}$ fixed point is also known as a **critical point** or an **equilibrium point** or a **stationary point**. This point is also called **stagnation point** with respect to the flow ϕ_t in $\mathbb{R}^n$. Flows on line may have no fixed points, only one fixed point, finite number of fixed points, and infinite number of fixed points. For example, the flow $\dot{x} = 5$ (no fixed points), $\dot{x} = x$ (only one fixed point), $\dot{x} = x^2 - 1$ (two fixed points), and $\dot{x} = \sin x$ (infinite number of fixed points).

1.5 Linear Stability Analysis

A fixed point, say x_0 is said to be stable if for a given $\varepsilon > 0$, there exists a $\delta > 0$ depending upon ε such that for all $t \geq t_0$, $\|x(t) - x_0(t)\| < \varepsilon$, whenever $\|x(t_0) - x_0(t_0)\| < \delta$, where $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ denotes the norm of a vector in $\mathbb{R}^n$. Otherwise, the fixed point is called unstable. In linear stability analysis the quadratic and higher order terms in the Taylor series expansion about a fixed point x^* of a system $\dot{x} = f(x), x \in \mathbb{R}$ are neglected due to the smallness of the terms. Consider a small perturbation quantity $\xi(t)$, away from the fixed point x^*, such that $x(t) = x^* + \xi(t)$. We see whether the perturbation grows or decays as time goes on. So we get the perturbation equation as

$$\dot{\xi} = \dot{x} = f(x) = f(x^* + \xi).$$

Taylor series expansion of $f(x^* + \xi)$ gives

$$\dot{\xi} = f(x^*) + \xi f'(x^*) + \frac{\xi^2}{2} f''(x^*) + \cdots$$

According to linear stability analysis, we get

$$\dot{\xi} = \xi f'(x^*) \, [\because f(x^*) = 0]$$

Assuming $f'(x^*) \neq 0$, the perturbation $\xi(t)$ grows exponentially if $f'(x^*) > 0$ and decays exponentially if $f'(x^*) < 0$. Linear theory fails if $f'(x^*) = 0$ and then higher order derivatives must be considered in the neighborhood of fixed point for stability analysis of the system.

Example 1.1 Find the evolution operator ϕ_t for the one-dimensional flow $\dot{x} = -x^2$. Show that ϕ_t forms a dynamical group. Is it a commutative group?

Solution The solutions of the given system are obtained as below:

$$\dot{x} = \frac{\mathrm{d}x}{\mathrm{d}t} = -x^2 \Rightarrow \frac{1}{x} = t + A \Rightarrow x(t) = \frac{1}{t+A}$$

in any interval of $\mathbb{R}$ that does not contain the point $x = 0$, where A is a constant. If we take starting point $x(0) = x_0$, then $A = 1/x_0$ and so we get

$$x(t) = \frac{x_0}{1 + x_0 t}, \quad t \neq -1/x_0.$$

The point $x = 0$ is not included in this solution. But it is the fixed point of the given system, because $\dot{x} = 0 \Leftrightarrow x = 0$. Therefore, $\phi_t(0) = 0$ for all $t \in \mathbb{R}$. So the evolution operator of the system is given as $\phi_t(x) = \frac{x}{1+xt}$ for all $x \in \mathbb{R}$.

The evolution operator ϕ_t is not defined for all $t \in \mathbb{R}$. For example, if $t = -1/x, x \neq 0$, then ϕ_t is undefined. Thus we see that the interval in which ϕ_t is defined is completely dependent on x.

We shall now examine the group properties of the evolution operator ϕ_t below:

(i) $\phi_r \phi_s \in \left\{\phi_t(x), t \in \mathbb{R}, x \in \mathbb{R}\right\} \forall r, s \in \mathbb{R}$ (Closure property)

Now,

$$\begin{aligned}
\phi_r(y) &= \frac{y}{1+yr}. \text{ Take } y = \frac{x}{1+xs} \\
&= \frac{x/1+sx}{1+\frac{x}{1+x}} = \frac{x}{1+xs+xr} = \frac{x}{1+x(s+r)} \\
&= \phi_{s+r} \in \left\{\phi_t(x), t \in \mathbb{R}, x \in \mathbb{R}\right\}
\end{aligned}$$

(ii) $\phi_t\left(\phi_s \phi_r\right) = \left(\phi_t \phi_s\right)\phi_r$ (Associative property)

$$\begin{aligned}
\text{L.H.S. } &= \phi_t\left(\left(\phi_s\phi_r\right)(x)\right) = \phi_t(y) = \frac{y}{1+yt} = \frac{z}{1+zs} = \frac{x}{1+x(r+s)}, y = \phi_s\left(\phi_r(x)\right) \\
&\left(\text{where } y = \phi_s(z), z = \phi_r(x) = \frac{x}{1+rx}\right) \\
\therefore\ \text{L.H.S } &= \frac{x}{1+x(t+r+s)} = \phi_{r+r+s}(x) \\
\text{R.H.S. } &= \left(\left(\phi_t\phi_s\right)\phi_r(x)\right)
\end{aligned}$$

Now,

$$\begin{aligned}
\phi_t(y) &= \frac{y}{1+yt}, y = \phi_s(x) = \frac{x}{1+sx} \\
&= \frac{x}{1+x(t+s)} = \phi_{t+s}(x) \\
\phi_{t+s}\left(\phi_r\right)(x) &= \phi_{t+s}(z) = \frac{z}{1+z(t+s)}, z = \phi_r(x) = \frac{x}{1+rx} \\
\phi_{t+s}\left(\phi_r\right)(x) &= \frac{x}{1+x(t+s+r)} = \phi_{t+s+r}(x)
\end{aligned}$$

Hence, $\phi_t\left(\phi_s\phi_r\right)(x) = \left(\phi_s\phi_r\right)\phi_t(x), \forall x \in \mathbb{R}$.

(iii) $\phi_0(x) = \frac{x}{1+x.0} = x, \phi_0$ is the identity operator.

$$\begin{aligned}\phi_t\phi_{-t}(x) = \phi_t(y) &= \frac{y}{1+ty}, y = \phi_{-t}(x) = \frac{x}{1-tx} \\ &= \frac{x}{1-tx+tx} = x = \phi_0(x) \quad \left(\phi_{-t} \text{ is the inverse of } \phi_t\right)\end{aligned}$$

Hence the flow evolution operator forms a dynamical group.

(v) $\phi_t\phi_s = \phi_s\phi_t$

Now,

$$\begin{aligned}\left(\phi_t\phi_s\right)(x) = \phi_t(y) &= \frac{y}{1+ty}, y = \phi_s(x) = \frac{x}{1+xs} \\ &= \frac{x}{1+x(t+s)} = \phi_{t+s}(x) \\ \phi_s\phi_t(x) = \phi_s(z) &= \frac{z}{1+sz}, z = \phi_t(x) = \frac{x}{1+tx} \\ &= \frac{x}{1+tx+sx} = \frac{x}{1+(s+t)x} = \phi_{s+t}(x)\end{aligned}$$

So, $\phi_t\phi_s = \phi_s\phi_t$ (Commutative property).

Thus, the evolution operator ϕ_t forms a commutative group.

Example 1.2 Using linear stability analysis determine the stability of the critical points for the following systems

$$\text{(i) } \dot{x} = \sin x, \quad \text{(ii) } \dot{x} = x^2$$

Solution (i) The given system has infinite numbers of critical points. The critical points are $x_n^* = n\pi, n = 0, \pm 1, \pm 2, \ldots$. When n is even, $f'\left(x_n^*\right) = \cos\left(x_n^*\right) = \cos(n\pi) = (-1)^n = 1 > 0$. So, these critical points are unstable. When n is odd, $f'\left(x_n^*\right) = -1 < 0$, and so these critical points are stable.

(ii) The critical point of the system is at $x^* = 0$. Now, $f'\left(x^*\right) = 0$ and $f''\left(x^*\right) = 2 > 0$. Hence, x^* is attracting when $x < 0$ and repelling when $x > 0$. Actually, the critical point is semi-stable in nature.

1.6 Analysis of One-Dimensional Flows

As we know qualitative approach is the combination of analysis and geometry and is a powerful tool for analyzing solution behaviors of a system qualitatively. By drawing trajectories in phase line/plane/space, the behaviors of phase points may be found easily. In qualitative analysis we mainly look for the following solution behaviors:

(i) Local stabilities of fixed points for a system;

(ii) Analyzing the existence of periodic/quasi-periodic solutions, limit cycle, relaxation oscillation, hysteresis, etc.;

(iii) Local and asymptotic solution behaviors of a system;

(iv) Topological features of flows such as bifurcations, catastrophe, topological equivalence, transitiveness, etc.

We shall now analyze a simple one-dimensional system as follows.

Consider a one-dimensional system represented as $\dot{x}(t) = \sin x$ with the initial condition $x(t = 0) = x(0) = x_0$. The characteristic features of the system are (i) it is a one-dimensional system, (ii) nonlinear system (iii) autonomous system, and (iv) its closed-form solution (analytical solution) exists. This is a one-dimensional flow and we analyze the system on the basis of flow. The analytical solution of the system is obtained easily

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sin x \Rightarrow \mathrm{d}t = \mathrm{cosec}(x)\mathrm{d}x$$

Integrating, we get

$$\begin{aligned} t &= \int \mathrm{cosec}(x)\mathrm{d}x \\ &= -\log|\mathrm{cosec}(x) + \cot(x)| + c \end{aligned}$$

where c is an integrating constant. Using the initial condition $x(0) = x_0$, we get the integrating constant c as

$$c = \log\left|\mathrm{cosec}\,(x_0) + \cot\,(x_0)\right|.$$

Thus the solution of the system is given as

$$t = \log\left|\frac{\mathrm{cosec}\,(x_0) + \cot\,(x_0)}{\mathrm{cosec}(x) + \cot(x)}\right|$$

From this closed-form solution, the behaviors of solutions for any initial conditions are difficult to analyse. Moreover, the asymptotic values of the system are
also difficult to obtain. The qualitative approach can give better dynamical behavior about this simple system.

We consider t as time, x as the position of an imaginary particle moving along the flow in real line and $\dot{x}$ as the velocity of that particle. The differential equation $\dot{x} = \sin x$ represents a vector field on the line. It gives the velocity vector $\dot{x}$ at each position x. The arrows point to the right when $\dot{x} > 0$ and to the left when $\dot{x} < 0$. We shall draw the graph of $\sin x$ versus x in $x\dot{x}$ - plane which gives the flow in the x-axis (see Fig. 1.1).

We may imagine that fluid is flowing steadily along the x-axis with a velocity $\dot{x}$ which varies from place to place, according to equation $\dot{x} = \sin x$. At points $\dot{x} = 0$, there is no flow and such points are called equilibrium points (fixed points). According to the definition of fixed point, the equilibrium points of this system are obtained as $\sin x = 0 \Rightarrow x = n\pi (n = 0, \pm 1, \pm 2, \ldots)$. This simple looking autonomous system has infinite numbers of equilibrium points in $\mathbb{R}$. We can see that there are two kinds of equilibrium points. The equilibrium point where the flow is toward the point is called sink or attractor (neighboring trajectories approach asymptotically to the point as $t \to \infty$). On the other hand, when the flow is away from the point, the point is called source or repellor (neighboring trajectories move away from the point as $t \to \infty$). From the above figure the solid circles represent the sinks that are stable equilibrium points and the open circles are the sources, which are unstable equilibrium points. The names are given because the sinks and sources are common in fluid flow problems. From the geometric approach one can get local stability behavior of the equilibrium points of the system easily and is valid for all time. We shall now re-look the analytical solution of the system. The analytical solution can be expressed as

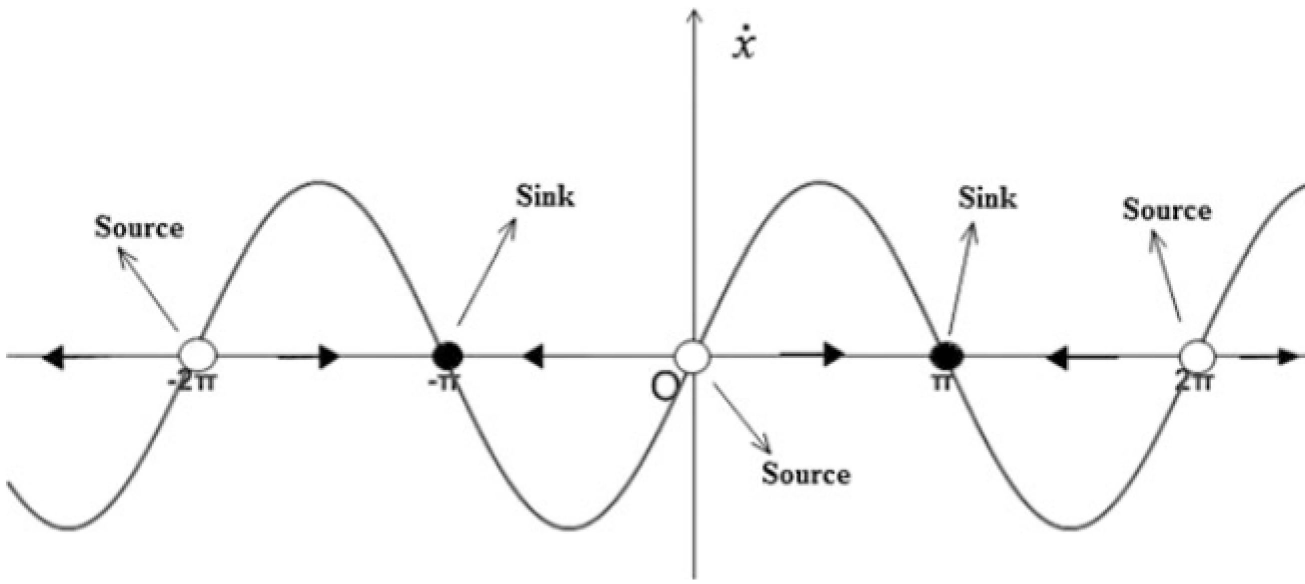

Figure 1.1: Graphical representation of flow generated by $\sin(x)$

Let the initial condition be $x_0 = x(0) = \pi/4$. Then from the above solution we obtain

$$A = \tan(\pi/8) = -1 + \sqrt{2} = 1/(1 + \sqrt{2})$$

So the solution is expressed as

$$x(t) = 2\tan^{-1}\left(\frac{e^t}{1 + \sqrt{2}}\right)$$

We see that the solution $x(t) \to \pi$ and $t \to \infty$.

Without using analytical solution for this particular initial condition the same result can be found by drawing the graph of x versus t. So the solution's behavior at any initial condition can be obtained easily by geometric approach. This simple one-dimensional system also has an interesting application. For a slow motion of a spring immersed in a highly viscous fluid such as grease or viscoelastic fluid (the combined effects of fluid viscosity and elasticity for example, synovial fluid in the joints of human bones), the viscous damping force is very strong compared to the inertia of motion. So one can neglect acceleration term (that is, inertia) and the spring-mass system may be governed by the equation $\alpha\dot{x} = \sin x$, where $\alpha > 0$ (string constant) is a real number and the dynamics can be obtained using this approach for different values of α (see the book Strogatz [5] for more physical examples and explanations).

We shall discuss a few worked out examples presented below.

Example 1.5 With the help of flow concept discuss the local stability of the fixed points of $\dot{x} = f(x) = \left(x^2 - 1\right)$.

Solution The fixed points of the given autonomous system are given by setting $f(x) = 0$. This gives $x = \pm 1$. So the fixed points of the system are 1 and -1 . For the local stability of the system about these fixed points we plot the graph of the function $f(x)$ and then sketch the vector field. The flow is to the right direction, indicated by the symbol $' \rightarrow '$, where the velocity $\dot{x} > 0$, that is, where $\left(x^2 - 1\right) > 0$ and to the left direction, indicated by the symbol $' \leftarrow '$, where $\dot{x} < 0$, that is, $\left(x^2 - 1\right) < 0$. We also use solid circles to represent stable fixed points and open circles for unstable fixed points.

In Fig.1.2 the arrows indicate the flow of the system. From the figure, we see that the fixed point $x = 1$ is unstable, since it acts as a source point and the fixed point $x = -1$ is stable, since it acts as a sink point.

Example 1.6 Discuss the stability character of the fixed points for the system $\dot{x} = x(1 - x)$ using the concept of flow.

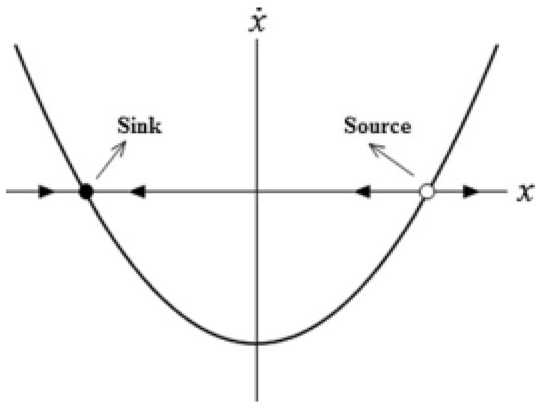

Figure 1.2: Graphical representation of $f(x) = \left(x^2 - 1\right)$

Solution Here $f(x) = x(1 - x)$. Then for the fixed points, we have

$$f(x) = 0 \Rightarrow x(1-x) = 0 \Rightarrow x = 0, 1$$

Thus the fixed points are 0 and 1 . To discuss the stability of these fixed points we plot the system (x versus $\dot{x}$) and then sketch the vector field. The flow is to the right direction, indicated by the symbol $' \rightarrow '$, when the velocity $\dot{x} > 0$, and to the left direction, indicated by the symbol $' \leftarrow '$, when $\dot{x} < 0$. We also use solid circle to represent stable fixed point and open circle to represent unstable fixed point.

From Fig. 1.3 we see that the fixed point $x = 1$ is stable whereas the fixed point $x = 0$ is unstable.

Example Find the fixed points and analyze the local stability of the following systems (i) $\dot{x} = x + x^3$ (ii) $\dot{x} = x - x^3$ (iii) $\dot{x} = -x - x^3$

Solution (i) Here $f(x) = x + x^3$. Then for fixed points $f(x) = 0 \Rightarrow x + x^3 = 0 \Rightarrow x = 0$, as $x \in \mathbb{R}$. So, 0 is the only fixed point of the system. We now see that when $x > 0, \dot{x} > 0$ and when $x < 0, \dot{x} < 0$. Hence the fixed point $x = 0$ is unstable. The graphical representation of the flow generated by the system is displayed in Fig.1.4.

(ii) Here $f(x) = x - x^3$. Then $f(x) = 0 \Rightarrow x - x^3 = 0 \Rightarrow x = 0, 1, -1$. Therefore, the fixed points of the system are $0, 1, -1$. We now see that

(a) when $x < -1$, then $\dot{x} > 0$

(b) when $-1 < x < 0, \dot{x} < 0$

(c) when $0 < x < 1, \dot{x} > 0$

(d) when $x > 1$, then $\dot{x} < 0$.

This shows that the fixed points 1 and -1 are stable whereas the fixed point 0 is unstable (Fig. 1.5).

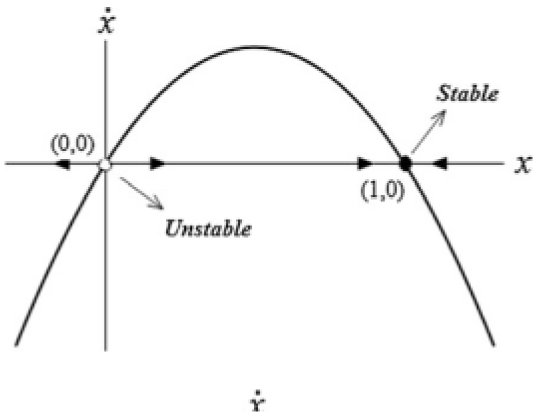

Figure 1.3: Pictorial representation of $f(x) = x(1-x)$

(iii) Here $f(x) = -x - x^3$. Then $f(x) = 0 \Rightarrow -x - x^3 = 0 \Rightarrow x = 0$, as $x \in \mathbb{R}$. So $x = 0$ is the only fixed point of the system. We now see that $\dot{x} > 0$ when $x < 0$ and $\dot{x} < 0$ when $x > 0$. This shows that the fixed point $x = 0$ is stable. The graphical representation of the flow generated by the system is displayed in

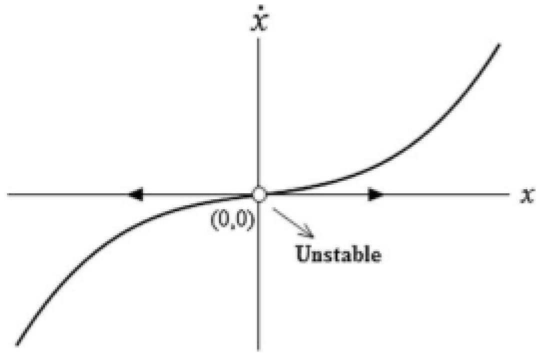

Figure 1.4: Graphical representation of $f(x) = \left(x + x^3\right)$

Fig. 1.6.

Example 1.8 Determine the equilibrium points and sketch the phase diagram in the neighborhood of the equilibrium points for the system represented as $\dot{x} + x\,\mathrm{sgn}(x) = 0$.

Solution Given system is $\dot{x} + x\,\mathrm{sgn}(x) = 0$, that is, $\dot{x} = -x\,\mathrm{sgn}(x)$, where the function $\mathrm{sgn}(x)$ is defined as

$$\mathrm{sgn}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}$$

For equilibrium points, we have

$$\dot{x} = 0 \Rightarrow x\,\mathrm{sgn}\,x = 0 \Rightarrow x = 0$$

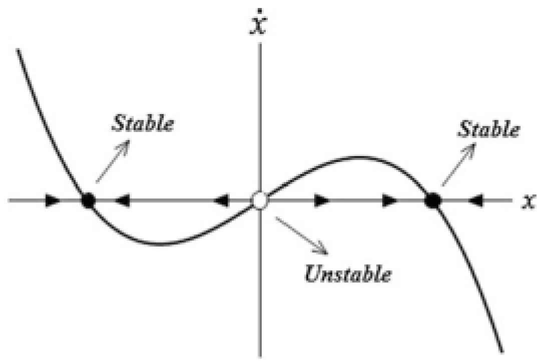

Figure 1.5: raphical representation of the flow generated by $\left(x - x^3\right)$

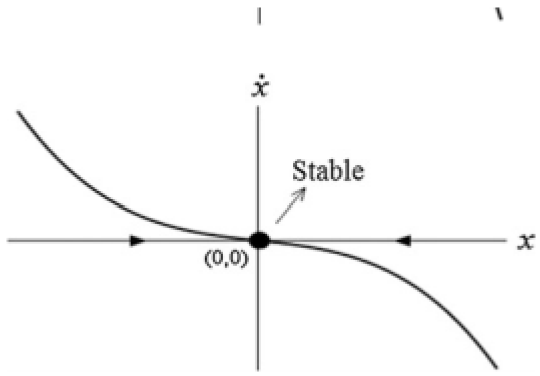

Figure 1.6: Graphical representation of $f(x) = \left(-x - x^3\right)$ versus x

This shows that the system has only one equilibrium point at $x = 0$. In flow analysis we see that the velocity $\dot{x} < 0$ for all $x \neq 0$. The flow is to the right direction, when $\dot{x} > 0$, in the negative x-axis and to the left direction, when $\dot{x} < 0$, in the positive x-axis. This is shown in the phase diagram depicted in Fig. 1.7, which shows that the fixed point origin is semi-stable.

1.7 Conservative and Dissipative Dynamical Systems

The dichotomy of dynamical systems in conservative versus dissipative is very important. They have some fundamental properties. Particularly, conservative systems are the integral part of Hamiltonian mechanics. We give here only the formal definitions of conservative and dissipative systems. Consider an autonomous system represented as

$$\dot{x} = f(x), x \in \mathbb{R}^n. \tag{1.2}$$

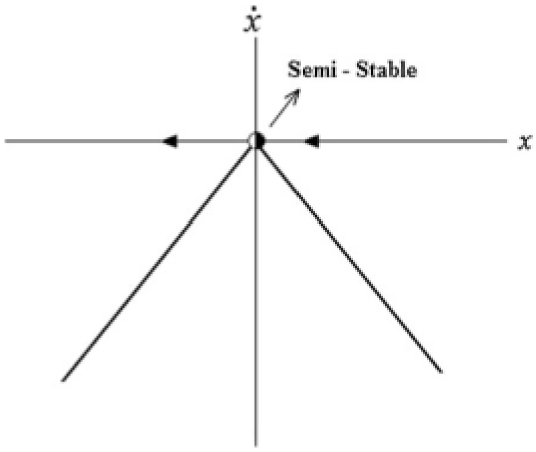

Figure 1.7: Graphical representation of the flow $\dot{x} = -x \operatorname{sgn} x$

The conservative and dissipative systems are defined with respect to the divergence of the corresponding vector field, which in turn refers to the conservation of volume or area in their state space or phase plane, respectively as follows:

A system is said to be **conservative** if the divergence of its vector field is zero. On the other hand, it is said to be **dissipative** if its vector field has negative divergence. The phase volume in a conservative system is constant under the flow while for a dissipative system the phase volume occupied by the system is gradually decreased as the time t increases and shrinks to zero as $t \to \infty$. When divergence of vector field is positive, the phase volume is gradually expanding. We shall discuss it in a later chapter. We state a lemma below which gives the change of volume in a phase space for an autonomous system.

Sometimes, it is useful to find the evolution of volume in the phase space of a system $\dot{x} = f(x), x \in \mathbb{R}^n$. The system generates a flow $\phi(t, x)$. We give Liouville's theorem which describes the time evolution of volume under the flow $\phi(t, x)$. Before this we now give the following lemma.

Lemma Consider an autonomous vector field $\dot{x} = f(x), x \in \mathbb{R}^n$ and generates a flow $\phi_t(x)$. Let D_0 be a domain in $\mathbb{R}^n$ and $\phi_t(D_0)$ be its evolution under the flow. If $V(t)$ is the volume of D_t, then the time rate of change of volume is given as $\left.\frac{\mathrm{d}v}{\mathrm{d}t}\right|_{t=0} = \int_{D_0} \nabla \cdot f dx$.

Proof The volume $V(t)$ can be expressed in the following form using the definition of the Jacobian of a transformation as

$$V(t) = \int_{D_0} \left| \frac{\partial\phi(t,x)}{\partial x} \right| dx$$

Expanding Taylor series of $\phi(t, x)$ in the neighborhood of $t = 0$, we get

$$\begin{aligned}\phi(t,x) &= x + f(x)t + O\left(t^2\right)\\ \Rightarrow \frac{\partial \phi}{\partial x} &= I + \frac{\partial f}{\partial x}t + O\left(t^2\right)\end{aligned}$$

Here I is the $n \times n$ identity matrix and

$$\begin{aligned}\left|\frac{\partial \phi}{\partial x}\right| &= \left|I + \frac{\partial f}{\partial x}t\right| + O\left(t^2\right)\\ &= 1 + \text{trace}\left(\frac{\partial f}{\partial x}\right)t + O\left(t^2\right) \text{ [Using expansion of the determinant]}\end{aligned}$$

Now, trace $\left(\frac{\partial f}{\partial x}\right) = \nabla \cdot f$, so we have

$$V(t) = V(0) + \int_{D_0} t\nabla \cdot f dx + O\left(t^2\right)$$

This gives $\frac{\mathrm{d}v}{\mathrm{d}t}\Big|_{t=0} = \int_{D_0} \nabla \cdot f dx$.

Theorem (Liouville's Theorem) Suppose $\nabla \cdot f = 0$ for a vector field f. Then for any region $D_0 \subseteq \mathbb{R}^n$, the volume $V(t)$ generated by the flow $\phi(t, x)$ is $V(t) = V(0)$, $V(0)$ being the volume of D_0.

Proof Suppose the divergence of the vector field f is everywhere constant, that is, $\nabla \cdot f = c$. For arbitrary time t_0 the evolution equation for the volume is given as $\dot{V} = cV$. This gives $V(t) = V(0)e^{ct}$. When the vector field is divergence free, that is, $c = 0$, we get the result $\dot{V} = 0 \Rightarrow V(t) = V(0) =$ constant. Thus we can say that the flow generated by a time independent system is volume preserving.

Examples of conservative and dissipative systems are presented below.

(a) Consider a linear and undamped pendulum represented as $\ddot{x} + x = 0$. This is an example of a conservative system. Setting $\dot{x} = y$, we can write it as a system of equations

$$\left.\begin{aligned}\dot{x} &= y\\ \dot{y} &= -x\end{aligned}\right\}$$

The system may also be written in the compact form $\dot{x} = fx)$, where $f(x) = \binom{y}{-x}$. The divergence of the vector field f is given by $\nabla \cdot f = \frac{\partial}{\partial x}(y) + \frac{\partial}{\partial y}(-x) = 0$. According to the definition, the system is conservative and the area occupied in the xy-phase plane is constant.

(b) The damped pendulum governed by $\ddot{x}+\alpha\dot{x}+\beta x = 0, \alpha, \beta > 0$ is an example of a dissipative system. Setting $\dot{x} = y$, we can write the system as

$$\left.\begin{aligned} \dot{x} &= y \\ \dot{y} &= -xy - \beta x \end{aligned}\right\}$$

The vector field is then expressed as $f(x) = \binom{y}{-\alpha y-\beta x}$. Now, $\vec{\nabla} \cdot f = \frac{\partial}{\partial x}(y) + \frac{\partial}{\partial y}(-\alpha y - \beta x) = -\alpha < 0$, since $\alpha > 0$.

This shows that the divergence of the vector field is negative.

So the system is dissipative in nature and the area in the phase plane is decreasing as time goes on. This is the simplest linear oscillator with linear damping. It describes a spring-mass system with a damper in parallel. The spring force is proportional to the extension x of the spring and the damping or frictional force is proportional to the velocity $\dot{x}$. The two constants α and β are related to the stiffness of the spring and the degrees of friction in the damper, respectively. According to the above lemma, the change in phase area is given by

$$A(t) = cA(0)e^{-\alpha}, \alpha > 0 \text{ as } t \to \infty, c \text{ being a constant.}$$

Example Find the phase volume element for the systems **(i)** $\dot{x} = -x$, **(ii)** $\dot{x} = ax - bxy, \dot{y} = bxy - cy$ where $x, y \geq 0$ and a, b, c are positive constants.

Solution (i) The flow of the system $\dot{x} = -x$ is attracted toward the point $x = 0$. The time rate of change of volume element $V(t)$ under the flow is given as

$$\left.\frac{\mathrm{d}V}{\mathrm{d}t}\right|_{r=0} = -\int_{D(0)} \mathrm{d}x = -V(0)$$

$$\text{or, } V(t) = V(0)e^{-t} \to 0 \text{ as } t \to \infty.$$

Hence the phase volume element $V(t)$ shrinks exponentially.

(ii) The given system is a Lotka-Volterra predator-prey population model. The rate of change in phase area $A(t)$ is given as

$$\begin{aligned} \frac{\mathrm{d}V}{\mathrm{d}t} &= -\int \vec{\nabla} : f \,\mathrm{d}x\,\mathrm{d}y \\ &= -\int (a - c - by + bx)\mathrm{d}x\,\mathrm{d}y \end{aligned}$$

This shows that a phase area periodically shrinks and expands.

1.8 Some Definitions

In this section we give some important preliminary definitions relating to flow of a system. The definitions given here are elaborately discussed in the later chapters for higher dimensional systems.

Invariant set A set $D \subset \mathbb{R}^n$ is said to be an invariant set under the flow ϕ_t if for any point $p \in D, \phi_t(p) \in D$ for all $t \in \mathbb{R}$. The set D is said to be positively invariant if $\phi_r(p) \in D$ for $t \geq 0$. Trajectories starting in an invariant set remain in the set for all times. An interval is called trapping if it is mapped into itself and is said to be invariant if it is mapped exactly onto itself. Moreover, if a bounded interval is trapping, then all of its trajectories are trapped inside and must converge to a closed, invariant, and bounded limit set. Basically these limit sets are the attractors of a system. So the periodic orbits are examples of invariant sets. We now define two limiting topological concepts which are relevant to the orbits of dynamical systems.

Limit points (ω - and α-limit points)

The asymptotic behavior of a trajectory may be related with limit points/sets or cycles and are termed as ω - and α-limit points/sets or cycles. We now give the definitions.

A point $p \in \mathbb{R}^n$ is called an ω-(resp. a α-) limit point if there exists a sequence $\{t_i\}$ with $t_i \to \infty$ (resp. $t_i \to -\infty$) such that $\phi(t_i, x) \to p$ as $i \to \infty$. The ω-limit set(cycle) is denoted by $\Lambda(x)$ and is defined as

$$\Lambda(x) = \left\{x \in \mathbb{R}^n \mid \exists \{t_i\} \text{ with } t_i \to \infty \text{ and } \phi(t_i, x) \to p \text{ as } i \to \infty\right\}.$$

Similarly, the α-limit set (cycle), $\mu(x)$, is defined as

$$\mu(x) = \left\{x \in \mathbb{R}^n \mid \exists \{t_i\} \text{ with } t_i \to -\infty \text{ and } \phi(t_i, x) \to p \text{ as } i \to \infty\right\}.$$

For example, consider a flow $\phi(t, x)$ on $\mathbb{R}^2$ generated by the system $\dot{r} = cr(1-r), \dot{\theta} = 1, c$ being a positive constant. For $x \neq 0$, let p be any point of the closed orbit C and take $\{t_i\}_{i=1}^{\infty}$ to be the sequence of $t > 0$. The trajectory through x crosses the radial line through p. So, $t_i \to \infty$ as $i \to \infty$ and $\phi(t_i, x) \to p$ as $i \to \infty$. If x lies in the closed orbit C, then $\phi(t_i, x) = p$ for each i. Hence every point of C is a ω-limit point of x and so $\Lambda(x) = C$ for every $x \neq 0$. When $|x| \leq 1$, the sequence $\{t_i\}_{i=1}^{\infty}$ with $t < 0$ gives the α-limit set $\mu(x) = \begin{cases} \{0\} & \text{for } |x| < 1 \\ \text{closed orbit} & \text{for } |x| = 1 \end{cases}$.

When $|x| > 1$, there is no sequence $\{t_i\}_{i=1}^{\infty}$, with $t_i \to \infty$ as $i \to \infty$, such that $\phi(t_i, x)$ exists as $i \to \infty$. So, $\mu(x)$ is empty when $|x| > 1$. The closed orbit C is called a limit cycle of the system.

The trajectory of a system through a point x is the set $\gamma(x) = \bigcup_{t \in \mathbb{R}} \phi(t, x)$ and the corresponding positive semi-trajectory $\gamma^+(x)$ and negative semi-trajectory $\gamma^-(x)$ are defined as follows:

$$\gamma^+(x) = \bigcup_{t \geq 0} \phi(t, x) \text{ and } \gamma^-(x) = \bigcup_{t \leq 0} \phi(t, x).$$

Attracting set A closed invariant set $D \subset \mathbb{R}^n$ for a flow ϕ_r is said to be an attracting set if there exists some neighborhood U in D such that $\forall t \geq 0, \phi(t, U) \subset U$ and $\bigcap_{t>0} \phi(t, U) = D$.

Absorbing set A positive invariant compact subset $B \subseteq \mathbb{R}^n$ is said to be an absorbing set if there exists a bounded subset C of $\mathbb{R}^n$ with $C \supset B$ such that $t_C > 0 \Rightarrow \phi(t, C) \subset B \forall t \geq t_C$ (see the book by Wiggins [7] for details).

Trapping zone An open set U in an invariant set $D \subset \mathbb{R}^n$ in an attracting set for a flow generated by a system is called a trapping zone. Let a set A be closed and invariant. The set A is said to be stable if and only if every neighborhood of A contains a neighborhood U of A which is trapping.

Basin of attraction The domain (called as basin of attraction) of an attracting set D is defined as $\bigcup_{t\leq 0} \phi(t, U)$ where U is any open set in $D \subset \mathbb{R}^n$.

Consider the one-dimensional system $\dot{x} = -x^4 \sin(\pi/x)$. It has countably infinite set of fixed points at $x^* = 0, \pm\frac{1}{n}, n = 1, 2, 3, \ldots$. Now,

$$\begin{aligned} f(x) = -x^4 \sin(\pi/x) &\Rightarrow f'(x) = -4x^3 \sin(\pi/x) + \pi x^2 \cos(\pi/x) \\ &\Rightarrow f'\left(x^*\right)\Big|_{x-\pm\frac{1}{n}} = \frac{\pi}{n^2}\cos(n\pi) = \frac{\pi}{n^2}(-1)^n. \end{aligned}$$

The fixed point $x^* = 0$ is neither attracting nor repelling. The interval $[-1, 1]$ is an attracting set of the given system. The fixed points $x^* = \pm\frac{1}{2n}, n = 1, 2, \ldots$ are repelling while the fixed points $x^* = \pm\frac{1}{(2\pi-1)}, n = 1, 2, \ldots$ are attracting.

CHAPTER 2

LINEAR SYSTEMS

This chapter presents a study of linear systems of ordinary differential equations:

$$\dot{\mathbf{x}} = A\mathbf{x} \tag{2.1}$$

where $\mathbf{x} \in \mathbf{R}^n, A$ is an $n \times n$ matrix and

$$\dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt} = \begin{bmatrix} \frac{dx_1}{dt} \\ \vdots \\ \frac{dx_n}{dt} \end{bmatrix}$$

It is shown that the solution of the linear system 2.1 together with the initial condition $\mathbf{x}(0) = \mathbf{x}_0$ is given by

$$\mathbf{x}(t) = e^{At}\mathbf{x}_0$$

where e^{At} is an $n \times n$ matrix function defined by its Taylor series. A good portion of this chapter is concerned with the computation of the matrix e^{At} in terms of the eigenvalues and eigenvectors of the square matrix A. Throughout this cour all vectors will be written as column vectors and A^T will denote the transpose of the matrix A.

2.1 Uncoupled Linear Systems

The method of separation of variables can be used to solve the first-order linear differential equation

$$\dot{x} = ax$$

The general solution is given by

$$x(t) = ce^{at}$$

where the constant $c = x(0)$, the value of the function $x(t)$ at time $t = 0$.

Now consider the uncoupled linear system

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

This system can be written in matrix form as (2.1). where

$$A = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$$

Note that in this case A is a diagonal matrix, $A = \text{diag}[-1, 2]$, and in general whenever A is a diagonal matrix, the system 2.1 reduces to an uncoupled linear system. The general solution of the above uncoupled linear system can once again be found by the method of separation of variables. It is given by

$$x_1(t) = c_1 e^{-t} \tag{2.2}$$
$$x_2(t) = c_2 e^{2t} \tag{2.3}$$

or equivalently by

$$\mathbf{x}(t) = \begin{bmatrix} \mathrm{e}^{-t} & 0 \\ 0 & \mathrm{e}^{2t} \end{bmatrix} \mathbf{c} \tag{2.4}$$

where $\mathbf{c} = \mathbf{x}(0)$. Note that the solution curves 2.2, 4 lie on the algebraic curves $y = k/x^2$ where the constant $k = c_1^2 c_2$. The solution 2.2, 4 or 2.4 defines a motion along these curves; i.e., each point $\mathbf{c} \in \mathbf{R}^2$ moves to the point $\mathbf{x}(t) \in \mathbf{R}^2$ given by 2.4 after time t. This motion can be described geometrically by drawing the solution curves 2.2, 4 in the x_1, x_2 plane, referred to as the phase plane, and by using arrows to indicate the direction of the motion along these curves with increasing time t; cf. Figure 1. For $c_1 = c_2 = 0, x_1(t) = 0$ and $x_2(t) = 0$ for all $t \in \mathbf{R}$ and the origin is referred to as an equilibrium point in this example. Note that solutions starting on the x_1-axis approach the origin as $t \to \infty$ and that solutions starting on the x_2-axis approach the origin as $t \to -\infty$.

The phase portrait of a system of differential equations such as 2.1 with $\mathbf{x} \in \mathbf{R}^n$ is the set of all solution curves of 2.1 in the phase space $\mathbf{R}^n$. Figure 1 gives a geometrical representation of the phase portrait of the uncoupled linear system considered above. The dynamical system defined by the linear system 2.1 in this example is simply the mapping $\phi : \mathbf{R} \times \mathbf{R}^2 \to \mathbf{R}^2$ defined by the solution $\mathbf{x}(t, \mathbf{c})$ given by 2.4; i.e.,

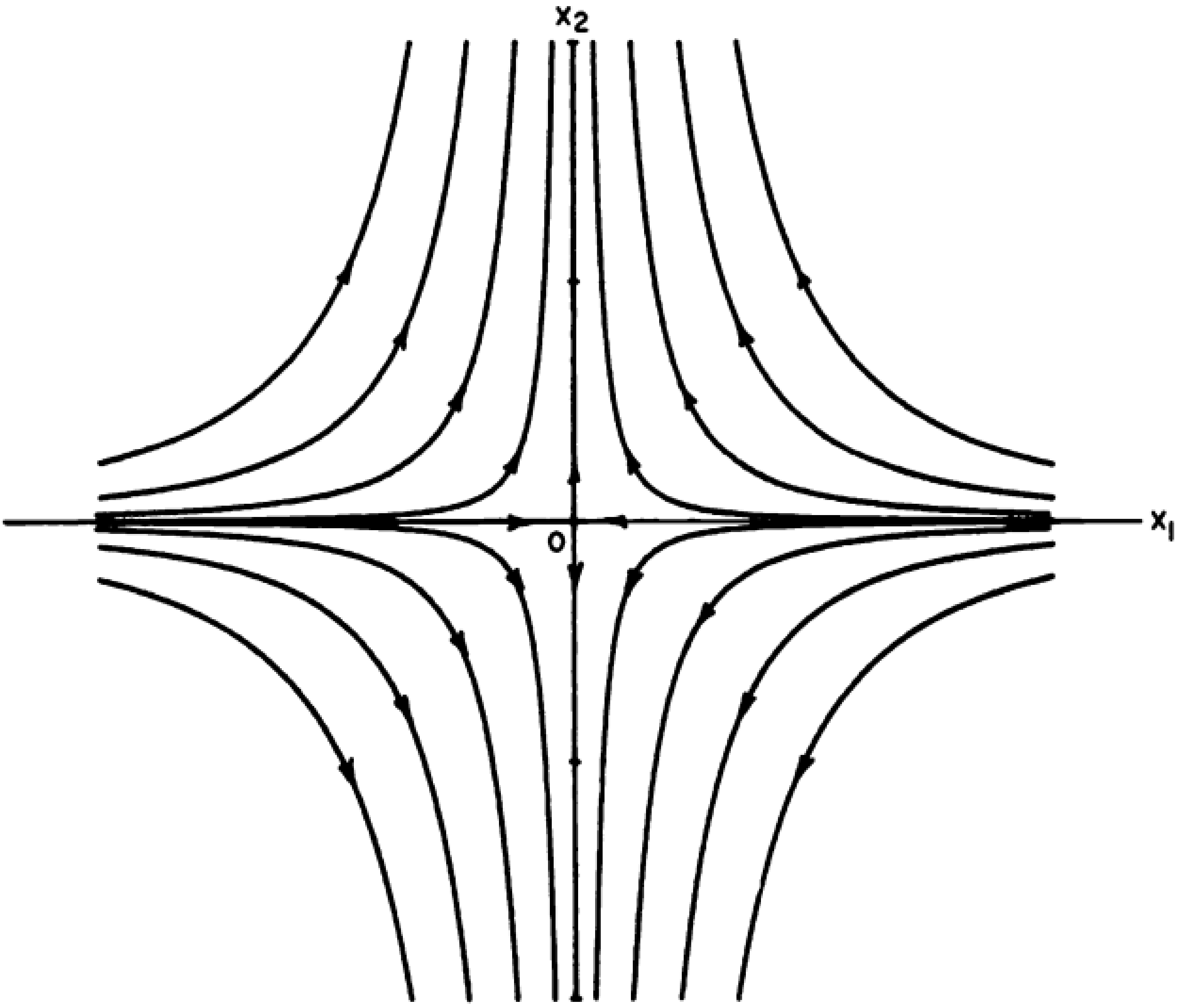

Figure 1

Geometrically, the dynamical system describes the motion of the points in phase space along the solution curves defined by the system of differential equations.

The function

$$\mathbf{f}(\mathbf{x}) = A\mathbf{x}$$

on the right-hand side of 2.1 defines a mapping $\mathbf{f} : \mathbf{R}^2 \to \mathbf{R}^2$ (linear in this case).

This mapping (which need not be linear) defines a vector field on R^2; i.e., to each point $x \in R^2$, the mapping f assigns a vector $f(x)$. If we draw each vector $f(\mathbf{x})$ with its initial point at the point $\mathbf{x} \in R^2$, we obtain a geometrical representation of the vector field as shown in Figure 2.

Note that at each point $\mathbf{x}$ in the phase space $\mathbf{R}^2$, the solution curves 2.2 are tangent to the vectors in the vector field $A\mathbf{x}$. This follows since at time $t = t_0$, the velocity vector $\mathbf{v}_0 = \dot{\mathbf{x}}(t_0)$ is tangent to the curve $\mathbf{x} = \mathbf{x}(t)$ at the point $\mathrm{x}_0 = \mathbf{x}(t_0)$ and since $\dot{\mathrm{x}} = A\mathbf{x}$ along the solution curves. Consider the following uncoupled linear system in $\mathbf{R}^3$:

$$\begin{aligned} \dot{x}_1 &= x_1 \\ \dot{x}_2 &= x_2 \\ \dot{x}_3 &= -x_3 \end{aligned} \tag{2.5}$$

The general solution is given by

$$\begin{aligned} x_1(t) &= c_1 e^t \\ x_2(t) &= c_2 \mathrm{e}^t \\ x_3(t) &= c_3 \mathrm{e}^{-t} \end{aligned}$$

And the phase portrait for this system is shown in Figure 3 above. The x_1, x_2 plane is referred to as the unstable subspace of the system (2.5) and
the x_3 axis is called the stable subspace of the system (2.5). Precise definitions of the stable and unstable subspaces of a linear system will be given in the next section.

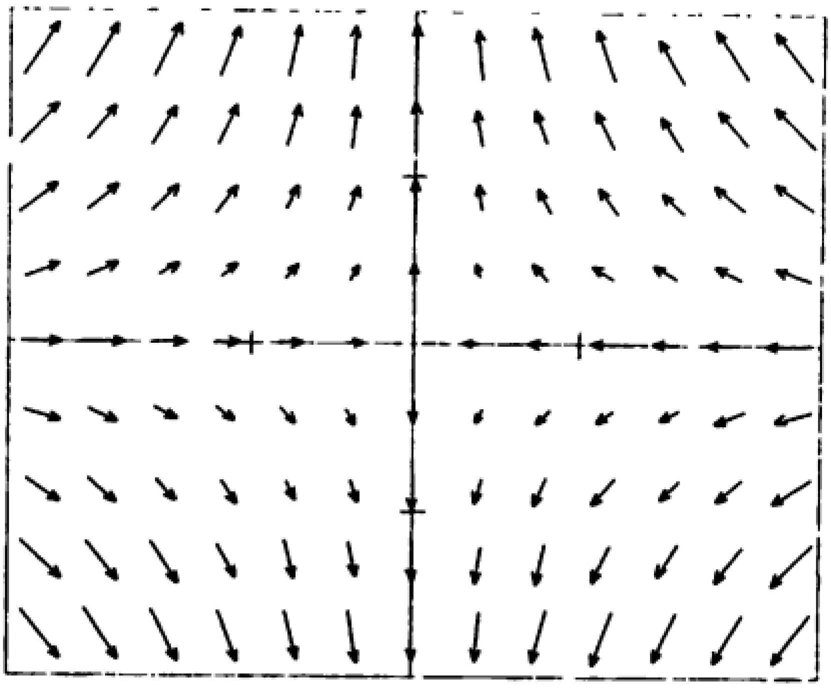

Figure 2

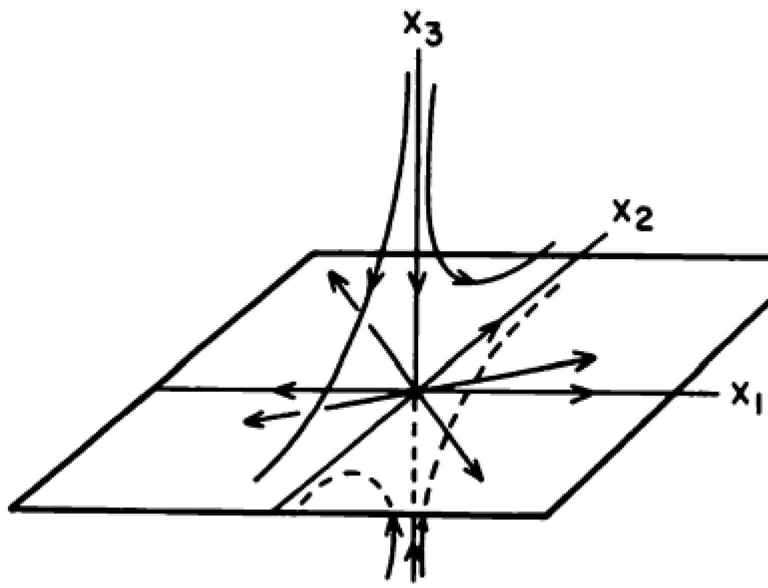

Figure 3

2.2 Diagonalization

The algebraic technique of diagonalizing a square matrix **A** can be used to reduce the linear system (2.1) to an uncoupled linear system. We first consider the case when A has real, distinct eigenvalues. The following theorem from linear algebra then allows us to solve the linear system (2.1).

Theorem: If the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of an $n \times n$ matrix A are real and distinct, then any set of corresponding eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ forms a basis for $\mathbf{R}^n$, the matrix $P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \cdots \mathbf{v}_n \end{bmatrix}$ is invertible and

$$P^{-1}AP = \operatorname{diag}[\lambda_1, \ldots, \lambda_n]$$

This theorem says that if a linear transformation $T : \mathbf{R}^n \to \mathbf{R}^n$ is represented by the $n \times n$ matrix A with respect to the standard basis $\{e_1, e_2, \ldots, e_n\}$ for $\mathbf{R}^n$, then with respect to any basis of eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}, T$ is represented by the diagonal matrix of eigenvalues, $\operatorname{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n]$. A proof of this theorem can be found, for example, in Lowenthal [Lo]. In order to reduce the system (2.1) to an uncoupled linear system using the above theorem, define the linear transformation of coordinates

$$\mathbf{y} = P^{-1}\mathbf{x}$$

where P is the invertible matrix defined in the theorem. Then

$$\mathbf{x} = P\mathbf{y},$$
$$\dot{\mathbf{y}} = P^{-1}\dot{\mathbf{x}} = P^{-1}A\mathbf{x} = P^{-1}AP\mathbf{y}$$

and, according to the above theorem, we obtain the uncoupled linear system

$$\dot{\mathbf{y}} = \operatorname{diag}[\lambda_1, \ldots, \lambda_n]\,\mathbf{y}$$

This uncoupled linear system has the solution

$$\mathbf{y}(t) = \operatorname{diag}\left[e^{\lambda_1 t}, \ldots, e^{\lambda_n t}\right]\mathbf{y}(0)$$

(Cf. problem 4 in Problem Set 1.) And then since $\mathbf{y}(0) = P^{-1}\mathbf{x}(0)$ and $\mathbf{x}(t) = P\mathbf{y}(t)$, it follows that (2.1) has the solution

$$\mathbf{x}(t) = PE(t)P^{-1}\mathbf{x}(0) \tag{2.6}$$

where $E(t)$ is the diagonal matrix

$$E(t) = \operatorname{diag}\left[e^{\lambda_1 t}, \ldots, e^{\lambda_n t}\right]$$

Corollary. Under the hypotheses of the above theorem, the solution of the linear system (2.1) is given by the function $\mathbf{x}(t)$ defined by (2.6).

Example 2.2.1 *Consider the linear system*

$$\dot{x}_1 = -x_1 - 3x_2$$
$$\dot{x}_2 = 2x_2$$

which can be written in the form (2.1) with the matrix

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

The eigenvalues of A are $\lambda_1 = -1$ and $\lambda_2 = 2$. A pair of corresponding eigenvectors is given by

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

The matrix P and its inverse are then given by

$$P = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad P^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

The student should verify that

$$P^{-1}AP = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$$

Then under the coordinate transformation $\mathbf{y} = P^{-1}\mathbf{x}$, *we obtain the uncoupled linear system*

$$\begin{aligned} \dot{y}_1 &= -y_1 \\ \dot{y}_2 &= 2y_2 \end{aligned}$$

which has the general solution $y_1(t) = c_1 e^{-t}$, $y_2(t) = c_2 e^{2t}$. *The phase portrait for this system is given in Figure 1 in Section 0.1 which is reproduced below. And according to the above corollary, the general solution to the original linear system of this example is given by*

$$\mathbf{x}(t) = P \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{2t} \end{bmatrix} P^{-1}\mathbf{c}$$

where $\mathbf{c} = \mathbf{x}(0)$, *or equivalently by*

$$\begin{aligned} x_1(t) &= c_1 e^{-t} + c_2 \left(e^{-t} - e^{2t}\right) \\ x_2(t) &= c_2 e^{2t} \end{aligned} \tag{3}$$

The phase portrait for the linear system of this example can be found by sketching the solution curves defined by (3). It is shown in Figure 2. The phase portrait in Figure 2 can also be obtained from the phase portrait in Figure 1 by applying the linear transformation of coordinates $\mathbf{x} = P\mathbf{y}$. *Note that the subspaces spanned by the eigenvectors* $\mathbf{v}_1$ *and* $\mathbf{v}_2$ *of the matrix A determine the stable and unstable subspaces of the linear system (2.1) according to the following definition: Suppose that the* $n \times n$ *matrix A has k negative eigenvalues* $\lambda_1, \ldots, \lambda_k$ *and* $n - k$ *positive eigenvalues* $\lambda_{k+1}, \ldots, \lambda_n$ *and that these eigenvalues are distinct. Let* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *be a corresponding set of eigenvectors. Then the stable and unstable subspaces of the linear system (2.1),* E^s *and* E^u, *are the linear* *subspaces spanned by* $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ *and* $\{\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ *respectively; i.e.,*

$$E^s = \text{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$$
$$E^u = \text{Span}\{\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$$

If the matrix A has pure imaginary eigenvalues, then there is also a center subspace E^c*; cf. Problem 2(c) in Section 0.1. The stable, unstable and center subspaces are defined for the general case in Section 0.9.*

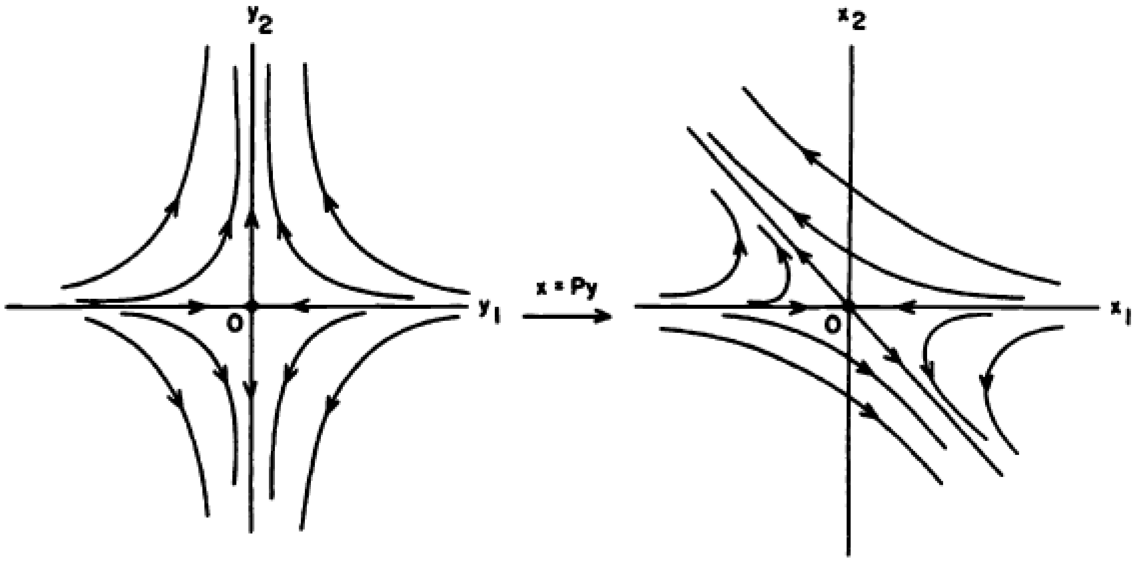

Figure 1 **Figure 2**

2.3 Exponentials of Operators

In order to define the exponential of a linear operator $T : \mathbf{R}^n \to \mathbf{R}^n$, it is necessary to define the concept of convergence in the linear space $L(\mathbf{R}^n)$ of linear operators on $\mathbf{R}^n$. This is done using the operator norm of T defined by

$$\|T\| = \max_{|\mathbf{x}|\leq 1} |T(\mathbf{x})|$$

where $|\mathbf{x}|$ denotes the Euclidean norm of $\mathbf{x} \in \mathbf{R}^n$; i.e.,

$$|\mathbf{x}| = \sqrt{x_1^2 + \cdots + x_n^2}$$

The operator norm has all of the usual properties of a norm, namely, for $S, T \in L(\mathbf{R}^n)$

(a) $\|T\| \geq 0$ and $\|T\| = 0$ iff $T = 0$

(b) $\|kT\| = |k|\|T\|$ for $k \in \mathbf{R}$

(c) $\|S + T\| \leq \|S\| + \|T\|$.

It follows from the Cauchy-Schwarz inequality that if $T \in L(\mathbf{R}^n)$ is represented by the matrix A with respect to the standard basis for $\mathbf{R}^n$, then $\|A\| \leq \sqrt{n}\ell$ where ℓ is the maximum length of the rows of A.

The convergence of a sequence of operators $T_k \in L(\mathbf{R}^n)$ is then defined in terms of the operator norm as follows:

Definition 1. A sequence of linear operators $T_k \in L(\mathbf{R}^n)$ is said to converge to a linear operator $T \in L(\mathbf{R}^n)$ as $k \to \infty$, i.e.,

$$\lim_{k\to\infty} T_k = T$$

if for all $\varepsilon > 0$ there exists an N such that for $k \geq N$, $\|T - T_k\| < \varepsilon$.

Lemma. For $S, T \in L(\mathbf{R}^n)$ and $\mathbf{x} \in \mathbf{R}^n$,

(1) $|T(\mathbf{x})| \leq \|T\||\mathbf{x}|$

(2) $\|TS\| \leq \|T\|\|S\|$

(3) $\left\|T^k\right\| \leq \|T\|^k$ for $k = 0, 1, 2, \ldots$.

Proof. (1) is obviously true for $\mathbf{x} = \mathbf{0}$. For $\mathbf{x} \neq 0$ define the unit vector $\mathbf{y} = \mathbf{x}/|\mathbf{x}|$. Then from the definition of the operator norm,

$$\|T\| \geq |T(\mathbf{y})| = \frac{1}{|\mathbf{x}|}|T(\mathbf{x})|$$

(2) For $|\mathbf{x}| \leq 1$, it follows from (1) that

$$\begin{aligned} |T(S(\mathbf{x}))| &\leq \|T\||S(\mathbf{x})| \\ &\leq \|T\|\|S\||\mathbf{x}| \\ &\leq \|T\|\|S\|. \end{aligned}$$

Therefore,

$$\|TS\| = \max_{|\mathbf{x}|\leq 1} |TS(\mathbf{x})| \leq \|T\|\|S\|$$

and (3) is an immediate consequence of (2).

Theorem. Given $T \in L(\mathbf{R}^n)$ and $t_0 > 0$, the series

$$\sum_{k=0}^{\infty} \frac{T^k t^k}{k!}$$

is absolutely and uniformly convergent for all $|t| \leq t_0$.

Proof. Let $\|T\| = a$. It then follows from the above lemma that for $|t| \leq t_0$,

$$\left\|\frac{T^k t^k}{k!}\right\| \le \frac{\|T\|^k |t|^k}{k!} \le \frac{a^k t_0^k}{k!}$$

But

$$\sum_{k=0}^{\infty} \frac{a^k t_0^k}{k!} = e^{a t_0}$$

It therefore follows from the Weierstrass M-Test that the series

$$\sum_{k=0}^{\infty} \frac{T^k t^k}{k!}$$

is absolutely and uniformly convergent for all $|t| \le t_0$; cf. [R], p.148.

The exponential of the linear operator T is then defined by the absolutely convergent series

$$e^T = \sum_{k=0}^{\infty} \frac{T^k}{k!}$$

It follows from properties of limits that e^T is a linear operator on $\mathbf{R}^n$ and it follows as in the proof of the above theorem that $\left\|e^T\right\| \le e^{\|T\|}$.

Since our main interest in this chapter is the solution of linear systems of the form

$$\dot{\mathbf{x}} = A\mathbf{x}$$

we shall assume that the linear transformation T on $\mathbf{R}^n$ is represented by the $n \times n$ matrix A with respect to the standard basis for $\mathbf{R}^n$ and define the exponential e^{At}.

Definition 2. Let A be an $n \times n$ matrix. Then for $t \in \mathbf{R}$,

$$e^{At} = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!}$$

For an $n \times n$ matrix A, e^{At} is an $n \times n$ matrix which can be computed in terms of the eigenvalues and eigenvectors of A. This will be carried out
in the remainder of this chapter. As in the proof of the above theorem $\left\|e^{At}\right\| \le e^{\|A\||t|}$ where $\|A\| = \|T\|$ and T is the linear transformation $T(\mathbf{x}) = A\mathbf{x}$.

We next establish some basic properties of the linear transformation e^T in order to facilitate the computation of e^T or of the $n \times n$ matrix e^A.

Proposition 1. If P and T are linear transformations on $\mathbf{R}^n$ and $S = PTP^{-1}$, then $e^S = Pe^T P^{-1}$.

Proof. It follows from the definition of e^S that

$$e^S = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{\left(PTP^{-1}\right)^k}{k!} = P \lim_{n\to\infty} \sum_{k=0}^{n} \frac{T^k}{k!} P^{-1} = Pe^T P^{-1}$$

The next result follows directly from Proposition 1 and Definition 2.

Corollary 1. If $P^{-1}AP = \operatorname{diag}\left[\lambda_j\right]$ then $e^{At} = P \operatorname{diag}\left[e^{\lambda_j t}\right] P^{-1}$.

Proposition 2. If S and T are linear transformations on $\mathbf{R}^n$ which commute, i.e., which satisfy $ST = TS$, then $\mathrm{e}^{S+T} = e^S e^T$.

Proof. If $ST = TS$, then by the binomial theorem

$$(S+T)^n = n! \sum_{j+k=n} \frac{S^j T^k}{j!k!}$$

Therefore,

$$e^{S+T} = \sum_{n=0}^{\infty} \sum_{j+k=n} \frac{S^j T^k}{j!k!} = \sum_{j=0}^{\infty} \frac{S^j}{j!} \sum_{k=0}^{\infty} \frac{T^k}{k!} = \mathrm{e}^S \mathrm{e}^T$$

We have used the fact that the product of two absolutely convergent series is an absolutely convergent series which is given by its Cauchy product; cf. [R], p. 74.

Upon setting $S = -T$ in Proposition 2, we obtain

Corollary 2. If T is a linear transformation on $\mathbf{R}^n$, the inverse of the linear transformation e^T is given by $\left(e^T\right)^{-1} = e^{-T}$.

Corollary 3. If

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

then

$$e^A = e^a \begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix}$$

Proof. If $\lambda = a + ib$, it follows by induction that

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}^k = \begin{bmatrix} \operatorname{Re}\left(\lambda^k\right) & -\operatorname{Im}\left(\lambda^k\right) \\ \operatorname{Im}\left(\lambda^k\right) & \operatorname{Re}\left(\lambda^k\right) \end{bmatrix}$$

where Re and Im denote the real and imaginary parts of the complex number λ respectively. Thus,

$$\begin{aligned}
\mathrm{e}^{A} &= \sum_{k=0}^{\infty} \begin{bmatrix} \mathrm{Re}\left(\frac{\lambda_k}{k!}\right) & -\mathrm{Im}\left(\frac{\lambda^k}{k!}\right) \\ \mathrm{Im}\left(\frac{\lambda^k}{k!}\right) & \mathrm{Re}\left(\frac{\lambda^k}{k!}\right) \end{bmatrix} \\
&= \begin{bmatrix} \mathrm{Re}\left(\mathrm{e}^{\lambda}\right) & -\mathrm{Im}\left(e^{\lambda}\right) \\ \mathrm{Im}\left(e^{\lambda}\right) & \mathrm{Re}\left(e^{\lambda}\right) \end{bmatrix} \\
&= e^{a} \begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix}
\end{aligned}$$

Note that if $a = 0$ in Corollary 3, then e^{A} is simply a rotation through b radians.

Corollary 4. If

$$A = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$$

then

$$e^{A} = e^{a} \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$$

Proof. Write $A = aI + B$ where

$$B = \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix}$$

Then aI commutes with B and by Proposition 2,

$$e^{A} = e^{aI} e^{B} = e^{a} e^{B}$$

And from the definition

$$e^{B} = I + B + B^{2}/2! + \cdots = I + B$$

since by direct computation $B^{2} = B^{3} = \cdots = 0$.

We can now compute the matrix e^{At} for any 2×2 matrix A. In Section 1.8 of this chapter it is shown that there is an invertible 2×2 matrix P (whose columns consist of generalized eigenvectors of A) such that the matrix

$$B = P^{-1} A P$$

has one of the following forms

$$B = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}, \quad B = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \quad \text{or} \quad B = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

It then follows from the above corollaries and Definition 2 that

$$e^{Bt} = \begin{bmatrix} e^{\lambda t} & 0 \\ 0 & e^{\mu t} \end{bmatrix}, \quad e^{Bt} = e^{\lambda t} \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \quad \text{or} \quad e^{Bt} = e^{at} \begin{bmatrix} \cos bt & -\sin bt \\ \sin bt & \cos bt \end{bmatrix}$$

respectively. And by Proposition 1, the matrix e^{At} is then given by

$$e^{At} = Pe^{Bt}P^{-1}$$

As we shall see in Section 1.4, finding the matrix e^{At} is equivalent to solving the linear system (1) in Section 1.1.

2.4 The Fundamental Theorem for Linear Systems

Let A be an $n \times n$ matrix. In this section we establish the fundamental fact that for $x_0 \in \mathbb{R}^n$ the initial value problem

$$\begin{aligned} \dot{\mathbf{x}} &= A\mathbf{x} \\ \mathbf{x}(0) &= \mathbf{x}_0. \end{aligned} \tag{1}$$

has a unique solution for all $t \in R^n$ which is given by

$$x(t) = \exp(At)X_0. \tag{2}$$

Notice the similarity in the form of the solution (2) and the solution $x(t) = \exp(At)X_0$ of the elementary first-order differential equation $x' = ax$ and initial condition $\mathbf{x}(0) = \mathbf{x}_0$

In order to prove this theorem, we first compute the derivative of the exponential function e^{At} using the basic fact from analysis that two convergent limit processes can be interchanged if one of them converges uniformly.

Lemma. Let A be a square matrix, then

$$\frac{d}{dt}e^{At} = Ae^{At}.$$

Proof. Since A commutes with itself, it follows from Proposition 2 and Definition 2 in Section 3 that

$$\begin{aligned}
\frac{d}{dt}e^{At} &= \lim_{h\to 0}\frac{e^{A(t+h)} - e^{At}}{h} \\
&= \lim_{h\to 0} e^{At}\frac{\left(e^{Ah} - I\right)}{h} \\
&= e^{At}\lim_{h\to 0}\lim_{k\to\infty}\left(A + \frac{A^2h}{2!} + \cdots + \frac{A^kh^{k-1}}{k!}\right) \\
&= Ae^{At}.
\end{aligned}$$

The last equality follows since by the theorem in Section 1.3 the series defining e^{Ah} converges uniformly for $|h| \leq 1$ and we can therefore interchange the two limits.

Theorem(The Fundamental Theorem for Linear Systems).
Let A be an $n \times n$ matrix. Then for a given $\mathbf{x}_0 \in \mathbf{R}^n$, the initial value problem

$$\begin{aligned}
\dot{\mathbf{x}} &= A\mathbf{x} \\
\mathbf{x}(0) &= \mathbf{x}_0
\end{aligned} \tag{1}$$

has a unique solution given by

$$\mathbf{x}(t) = e^{At}\mathbf{x}_0. \tag{2}$$

Proof. By the preceding lemma, if $\mathbf{x}(t) = e^{At}\mathbf{x}_0$, then

$$\mathbf{x}'(t) = \frac{d}{dt}e^{At}\mathbf{x}_0 = Ae^{At}\mathbf{x}_0 = A\mathbf{x}(t)$$

for all $t \in \mathbf{R}$. Also, $\mathbf{x}(0) = I\mathbf{x}_0 = \mathbf{x}_0$. Thus $\mathbf{x}(t) = e^{At}\mathbf{x}_0$ is a solution. To see that this is the only solution, let $\mathbf{x}(t)$ be any solution of the initial value problem (1) and set

$$\mathbf{y}(t) = e^{-At}\mathbf{x}(t).$$

Then from the above lemma and the fact that $\mathbf{x}(t)$ is a solution of (1)

$$\begin{aligned}
\mathbf{y}'(t) &= -Ae^{-At}\mathbf{x}(t) + e^{-At}\mathbf{x}'(t) \\
&= -Ae^{-At}\mathbf{x}(t) + e^{-At}A\mathbf{x}(t) \\
&= 0
\end{aligned}$$

for all $t \in \mathbf{R}$ since e^{-At} and A commute. Thus, $\mathbf{y}(t)$ is a constant. Setting $t = 0$ shows that $y(t) = x_0$ and therefore any solution of the initial value problem (1) is given by $\mathbf{x}(t) = e^{At}\mathbf{y}(t) = e^{At}\mathbf{x}_0$. This completes the proof of the theorem.

Example 2.4.1 *Solve the initial value problem*

$$\dot{\mathbf{x}} = A\mathbf{x}$$

$$\mathbf{x}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

for

$$A = \begin{bmatrix} -2 & -1 \\ 1 & -2 \end{bmatrix}$$

and sketch the solution curve in the phase plane $\mathbf{R}^2$*. By the above theorem and Corollary 3 of the last section, the solution is given by*

$$\mathbf{x}(t) = e^{At}\mathbf{x}_0 = e^{-2t}\begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = e^{-2t}\begin{bmatrix} \cos t \\ \sin t \end{bmatrix}.$$

It follows that $|\mathbf{x}(t)| = e^{-2t}$ *and that the angle* $\theta(t) = \tan^{-1} x_2(t)/x_1(t) = t$*. The solution curve therefore spirals into the origin as shown in Figure 1 below.*

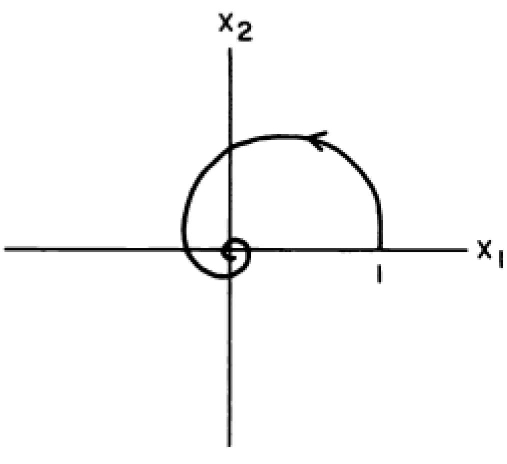

Figure 1

2.5 Linear Systems in R^2

In this section we discuss the various phase portraits that are possible for the linear system

$$\dot{\mathbf{x}} = A\mathbf{x} \tag{1}$$

when $\mathbf{x} \in \mathbf{R}^2$ and A is a 2×2 matrix. We begin by describing the phase portraits for the linear system

$$\dot{\mathbf{x}} = B\mathbf{x} \tag{2}$$

where the matrix $B = P^{-1}AP$ has one of the forms given at the end of Section 1.3. The phase portrait for the linear system (1) above is then obtained from the phase portrait for (2) under the linear transformation of coordinates $\mathbf{x} = P\mathbf{y}$ as in Figures 1 and 2 in Section 1.2. First of all, if

$$B = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}, \quad B = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \quad \text{or} \quad B = \begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

it follows from the fundamental theorem in Section 1.4 and the form of the matrix e^{Bt} computed in Section 1.3 that the solution of the initial value problem (2) with $\mathbf{x}(0) = x_0$ is given by

$$\mathbf{x}(t) = \begin{bmatrix} e^{\lambda t} & 0 \\ 0 & e^{\mu t} \end{bmatrix}\mathbf{x}_0, \quad \mathbf{x}(t) = e^{\lambda t}\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}\mathbf{x}_0,$$

or

$$\mathbf{x}(t) = e^{at}\begin{bmatrix} \cos bt & -\sin bt \\ \sin bt & \cos bt \end{bmatrix}\mathbf{x}_0$$

respectively. We now list the various phase portraits that result from these solutions, grouped according to their topological type with a finer classification of sources and sinks into various types of unstable and stable nodes and foci:

Case I. $B = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$ with $\lambda < 0 < \mu$.

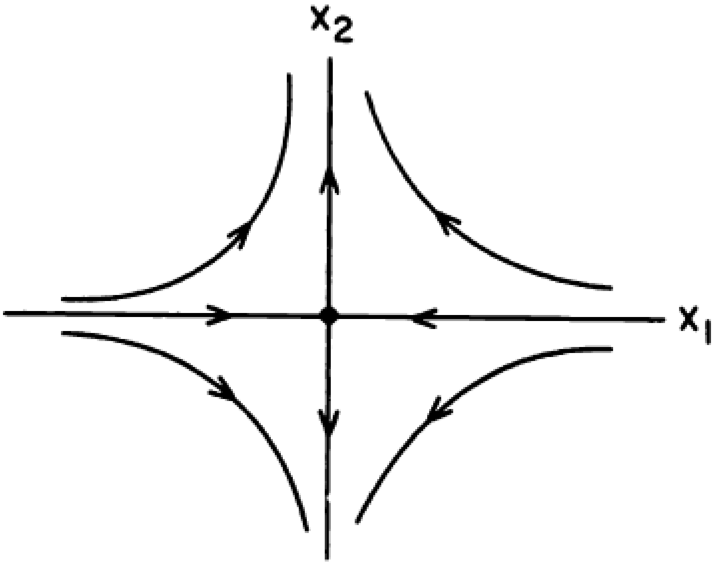

Figure 1. A saddle at the origin.

The phase portrait for the linear system (2) in this case is given in Figure 1. See the first example in Section 1.1. The system (2) is said to have a saddle at the origin in this case. If $\mu < 0 < \lambda$, the arrows in Figure 1 are reversed. Whenever A has two real eigenvalues of opposite sign, $\lambda < 0 < \mu$, the phase portrait for the linear system (1) is linearly equivalent to the phase portrait shown in Figure 1; i.e., it is obtained from Figure 1 by a linear transformation of coordinates; and the stable and unstable subspaces of (1) are determined by the eigenvectors of A as in the Example in Section 1.2. The four non-zero trajectories or solution curves that approach the equilibrium point at the origin as $t \to \pm\infty$ are called separatrices of the system.

Case II. $B = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$ with $\lambda \le \mu < 0$ or $B = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$ with $\lambda < 0$. The phase portraits for the linear system (2) in these cases are given in Figure 2. Cf. the phase portraits in Problems 1(a), (b) and (c) of Problem Set 1 respectively. The origin is referred to as a stable node in each of these cases. It is called a proper node in the first case with $\lambda = \mu$ and an improper node in the other two cases. If $\lambda \ge \mu > 0$ or if $\lambda > 0$ in Case II, the arrows in Figure 2 are reversed and the origin is referred to as an unstable node. Whenever A has two negative eigenvalues $\lambda \le \mu < 0$, the phase portrait of the linear system (1) is linearly equivalent to one of the phase portraits shown in Figure 2. The stability of the node is determined by the sign of the eigenvalues: stable if $\lambda \le \mu < 0$ and unstable if $\lambda \ge \mu > 0$. Note that each trajectory in Figure 2 approaches the equilibrium point at the origin along a well-defined tangent line $\theta = \theta_0$, determined by an eigenvector of A, as $t \to \infty$.

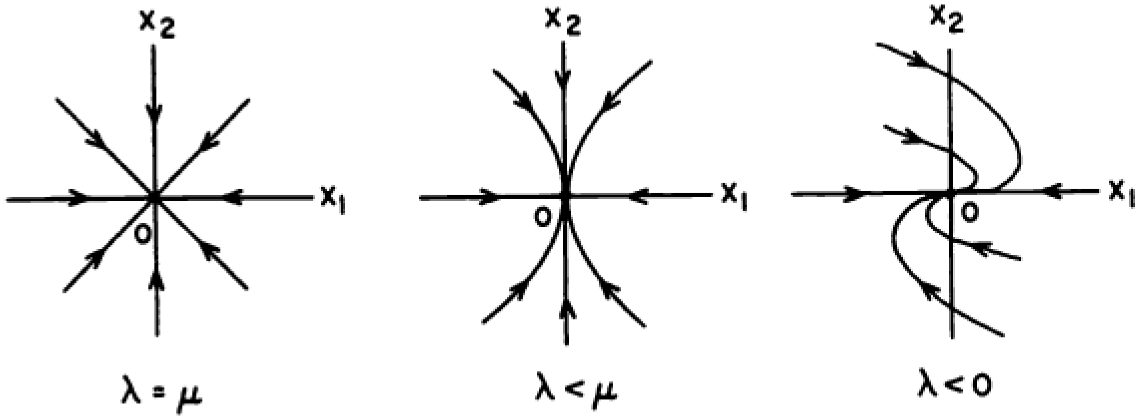

Figure 2. A stable node at the origin.

Case III. $B = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ with $a < 0$. $b > 0$ $b < 0$

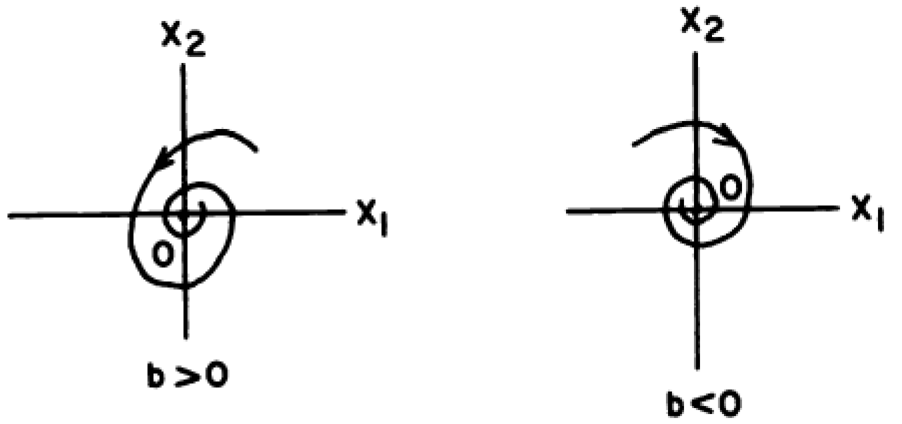

Figure 3. A stable focus at the origin.

The phase portrait for the linear system (2) in this case is given in Figure 3. Cf. Problem 9. The origin is referred to as a stable focus in these cases. If $a > 0$, the trajectories spiral away from the origin with increasing t and the origin is called an unstable focus. Whenever A has a pair of complex conjugate eigenvalues with nonzero real part, $a \pm ib$, with $a < 0$, the phase portraits for the system (1) is linearly equivalent to one of the phase portraits shown in Figure 3. Note that the trajectories in Figure 3 do not approach the origin along well-defined tangent lines; i.e., the angle $\theta(t)$ that the vector $\mathbf{x}(t)$ makes with the x_1-axis does not approach a constant θ_0 as $t \to \infty$, but rather $|\theta(t)| \to \infty$ as $t \to \infty$ and $|\mathbf{x}(t)| \to 0$ as $t \to \infty$ in this case.

Case IV. $B = \begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix}$ The phase portrait for the linear system (2) in this case is given in Figure 4. Cf. Problem 1(d) in Problem Set 1. The system (2) is said to have a center at the origin in this case. Whenever A has a pair of pure imaginary complex conjugate eigenvalues, $\pm ib$, the phase portrait of the linear system (1) is linearly equivalent to one of the phase portraits shown in Figure 4. Note that the trajectories or solution curves in Figure 4 lie on circles $|\mathbf{x}(t)| =$ constant. In general, the trajectories of the system (1) will lie on ellipses and the solution $\mathbf{x}(t)$ of (1) will satisfy $m \le |\mathbf{x}(t)| \le M$ for all $t \in \mathbf{R}$; cf. the following Example. The angle $\theta(t)$ also satisfies $|\theta(t)| \to \infty$ as $t \to \infty$ in this case.

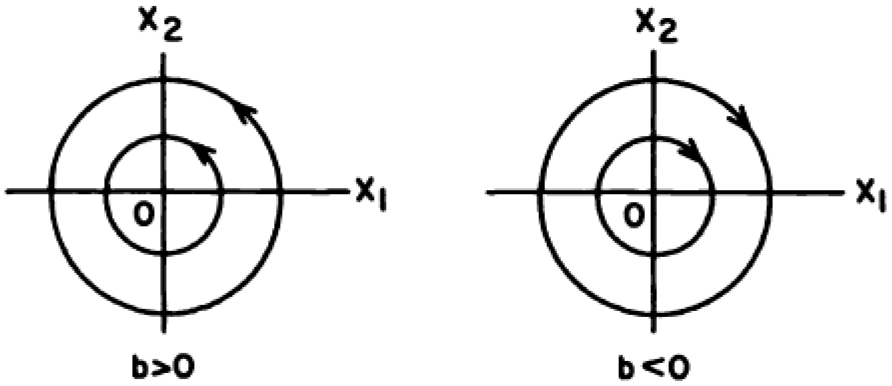

Figure 4. A center at the origin.

If one (or both) of the eigenvalues of A is zero, i.e., if $\det A = 0$, the origin is called a degenerate equilibrium point of (1). The various portraits for the linear system (1) are determined in Problem 4 in this case.

Example 2.5.1 ***(A linear system with a center at the origin)***.
The linear system

$$\dot{\text{x}} = A\text{x}$$

with

$$A = \begin{bmatrix} 0 & -4 \\ 1 & 0 \end{bmatrix}$$

has a center at the origin since the matrix A has eigenvalues $\lambda = \pm 2i$. According to the theorem in Section 1.6, the invertible matrix

$$P = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \quad \textit{with} \quad P^{-1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix}$$

reduces A to the matrix

$$B = P^{-1}AP = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$$

The student should verify the calculation. The solution to the linear system $\dot{\mathbf{x}} = A\mathbf{x}$, as determined by Sections 1.3 and 1.4 , is then given by

$$\mathbf{x}(t) = P \begin{bmatrix} \cos 2t & -\sin 2t \\ \sin 2t & \cos 2t \end{bmatrix} P^{-1}\mathbf{c} = \begin{bmatrix} \cos 2t & -2\sin 2t \\ 1/2\sin 2t & \cos 2t \end{bmatrix} \mathbf{c}$$

where $\mathbf{c} = \mathbf{x}(0)$, or equivalently by

$$x_1(t) = c_1 \cos 2t - 2c_2 \sin 2t$$

$$x_2(t) = 1/2 c_1 \sin 2t + c_2 \cos 2t.$$

It is then easily shown that the solutions satisfy

$$x_1^2(t) + 4x_2^2(t) = c_1^2 + 4c_2^2$$

for all $t \in \mathbf{R}$; i.e., the trajectories of this system lie on ellipses as shown in Figure 5.

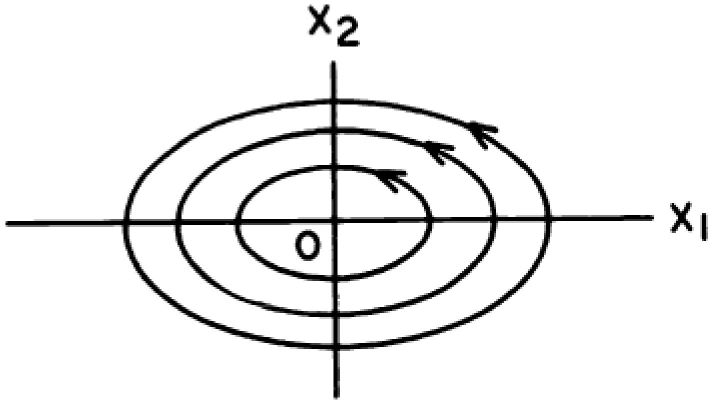

Figure 5. A center at the origin.

Definition 1. The linear system (1) is said to have a saddle, a node, a focus or a center at the origin if the matrix A is similar to one of the matrices B in Cases I, II, III or IV respectively, i.e., if its phase portrait is linearly equivalent to one of the phase portraits in Figures 1, 2, 3 or 4 respectively.

Remark If the matrix A is similar to the matrix B, i.e., if there is a nonsingular matrix P such that $P^{-1}AP = B$, then the system (1) is transformed into the system (2) by the linear transformation of coordinates $\mathbf{x} = P\mathbf{y}$. If B has the form III, then the phase portrait for the system (2) consists of either a counterclockwise motion (if $b > 0$) or a clockwise motion (if $b < 0$) on either circles (if $a = 0$) or spirals (if $a \neq 0$). Furthermore, the direction of rotation of trajectories in the phase portraits for the systems (1) and (2) will be the same if $\det P > 0$ (i.e., if P is orientation preserving) and it will be opposite if $\det P < 0$ (i.e., if P is orientation reversing).

For $\det A \neq 0$ there is an easy method for determining if the linear system has a saddle, node, focus or center at the origin. This is given in the next theorem. Note that if $\det A \neq 0$ then $A\mathbf{x} = 0$ if $\mathbf{x} = 0$; i.e., the origin is the only equilibrium point of the linear system (1) when $\det A \neq 0$. If the origin is a focus or a center, the sign σ of $\dot{x}_2$ for $x_2 = 0$ (and for small $x_1 > 0$) can be used to determine whether the motion is counterclockwise (if $\sigma > 0$) or clockwise (if $\sigma < 0$).

Theorem. Let $\delta = \det A$ and $\tau = \operatorname{trace} A$ and consider the linear system

$$\dot{\mathbf{x}} = A\mathbf{x}. \tag{1}$$

(a) If $\delta < 0$ then (1) has a saddle at the origin.

(b) If $\delta > 0$ and $\tau^2 - 4\delta \geq 0$ then (1) has a node at the origin; it is stable if $\tau < 0$ and unstable if $\tau > 0$.

(c) If $\delta > 0, \tau^2 - 4\delta < 0$, and $\tau \neq 0$ then (1) has a focus at the origin; it is stable if $\tau < 0$ and unstable if $\tau > 0$.
(d) If $\delta > 0$ and $\tau = 0$ then (1) has a center at the origin.

Note that in case (b), $\tau^2 \geq 4|\delta| > 0$; i.e., $\tau \neq 0$.
Proof The eigenvalues of the matrix A are given by

$$\lambda = \frac{\tau \pm \sqrt{\tau^2 - 4\delta}}{2}$$

Thus (a) if $\delta < 0$ there are two real eigenvalues of opposite sign.
(b) If $\delta > 0$ and $\tau^2 - 4\delta \geq 0$ then there are two real eigenvalues of the same sign as τ;
(c) if $\delta > 0, \tau^2 - 4\delta < 0$ and $\tau \neq 0$ then there are two complex conjugate eigenvalues $\lambda = a \pm ib$ and, as will be shown in Section 1.6, A is similar to the matrix B in Case III above with $a = \tau/2$; and
(d) if $\delta > 0$ and $\tau = 0$ then there are two pure imaginary complex conjugate eigenvalues. Thus, cases a, b, c and d correspond to the Cases I, II, III and IV discussed above and we have a saddle, node, focus or center respectively.
Definition 2. A stable node or focus of (1) is called a sink of the linear system and an unstable node or focus of (1) is called a source of the linear system.

The above results can be summarized in a "bifurcation diagram," shown in Figure 6, which separates the (τ, δ)-plane into three components in which the solutions of the linear system (1) have the same "qualitative structure". In describing the topological behavior or qualitative structure of the solution set of a linear system, we do not distinguish between nodes and foci, but only if they are stable or unstable. There are eight different topological types of behavior that are possible for a linear system according to whether $\delta \neq 0$ and it has a source, a sink, a center or a saddle or whether $\delta = 0$ and it has one of the four types of behavior determined in Problem 4.

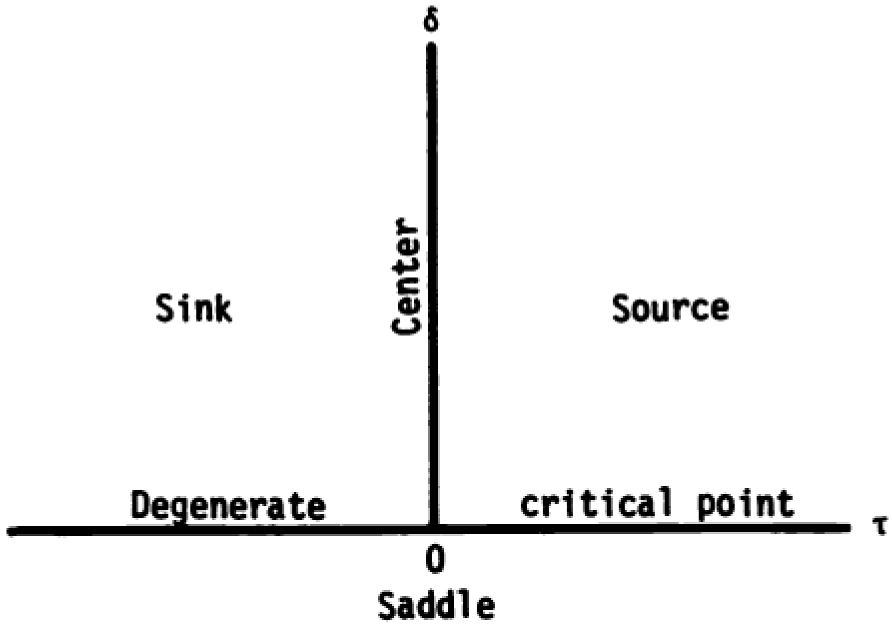

Figure 6. A bifurcation diagram for the linear system (1).

2.6 Complex Eigenvalues

If the $2n \times 2n$ real matrix A has complex eigenvalues, then they occur in complex conjugate pairs and if A has $2n$ distinct complex eigenvalues, the following theorem from linear algebra proved in Hirsch and Smale [H/S] allows us to solve the linear system (2.1).

Theorem. If the $2n \times 2n$ real matrix A has $2n$ distinct complex eigenualues $\lambda_j = a_j + ib_j$ and $\bar{\lambda}_j = a_j - ib_j$ and corresponding complex eigenvectors

$w_j = u_j + iv_j$ and $\bar{w}_j = u_j - iv_j$, $j = 1, \ldots, n$, then $\{u_1, v_1, \ldots, u_n, v_n\}$ is a basis for $\mathbf{R}^{2n}$, the matrix

$$P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{u}_1 & \mathbf{v}_2 & \mathbf{u}_2 & \cdots & \mathbf{v}_\mathrm{n} & \mathbf{u}_n \end{bmatrix}$$

is invertible and

$$P^{-1}AP = \mathrm{ding}\begin{bmatrix} a_j & -b_j \\ b_j & a_j \end{bmatrix},$$

a real $2n \times 2n$ matrix with 2×2 blocks along the diagonal.

Remark. Note that if instead of the matrix P we use the invertible matrix

$$Q = \begin{bmatrix} \mathbf{u}_1 & \mathbf{v}_1 & \mathbf{u}_2 & \mathbf{v}_2 & \cdots & \mathbf{u}_n & \mathbf{v}_n \end{bmatrix}$$

then

$$Q^{-1}AQ = \operatorname{diag}\begin{bmatrix} a_j & b_j \\ -b_j & a_j \end{bmatrix}.$$

The next corollary then follows from the above theorem and the fundamental theorem in Section 1.4.

Corollary. Under the hypotheses of the above theorem, the solution of the instial value problem

$$\begin{aligned} \dot{\mathbf{x}} &= A\mathbf{x} \\ \mathbf{x}(0) &= x_0 \end{aligned} \tag{1}$$

is given by

$$\mathbf{x}(t) = P \operatorname{diag} e^{a_j t} \begin{bmatrix} \cos b_j t & -\sin b_j t \\ \sin b_j t & \cos b_j t \end{bmatrix} P^{-1} x_0.$$

Note that the matrix

$$R = \begin{bmatrix} \cos bt & -\sin bt \\ \sin bt & \cos bt \end{bmatrix}$$

represents a rotation through bt radians.

Example 2.6.1 *Solve the initial value problem (1) for*

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

The matrix A has the complex eigenvalues $\lambda_1 = 1 + i$ and $\lambda_2 = 2 + i$ (as well as $\bar{\lambda}_1 = 1 - i$ and $\bar{\lambda}_2 = 2 - i$). A corresponding pair of complex eigenvectors is

$$w_1 = u_1 + iv_1 = \begin{bmatrix} i \\ 1 \\ 0 \\ 0 \end{bmatrix} \text{ and } w_2 = u_2 + iv_2 = \begin{bmatrix} 0 \\ 0 \\ 1+i \\ 1 \end{bmatrix}.$$

The matrix

$$P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{u}_1 & \mathbf{v}_2 & \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

is invertible,

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

and

$$P^{-1}AP = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

The solution to the initial value problem (1) is given by

$$\begin{aligned} \mathbf{x}(t) &= P \begin{bmatrix} e^t \cos t & -e^t \sin t & 0 & 0 \\ e^t \sin t & e^t \cos t & 0 & 0 \\ 0 & 0 & e^{2t} \cos t & -e^{2t} \sin t \\ 0 & 0 & e^{2t} \sin t & e^{2t} \cos t \end{bmatrix} P^{-1}\mathbf{x}_0 \\ &= \begin{bmatrix} e^t \cos t & -e^t \sin t & 0 & 0 \\ e^t \sin t & e^t \cos t & 0 & 0 \\ 0 & 0 & e^{2t}(\cos t + \sin t) & -2e^{2t} \sin t \\ 0 & 0 & e^{2t} \sin t & e^{2t}(\cos t - \sin t) \end{bmatrix} \mathbf{x}_0 \end{aligned}$$

In case A has both real and complex eigenvalues and they are distinct, we have the following result: If A has distinct real eigenvalues λ_j and corresponding eigenvectors $\mathbf{v}_j$, $j = 1, \ldots, k$ and distinct complex eigenvalues $\lambda_j = a_j + ib_j$ and $\bar{\lambda}_j = a_j - ib_j$ and corresponding eigenvectors $\mathbf{w}_j = u_j + i\mathbf{v}_j$ and $\bar{w}_j = u_j - i\mathbf{v}_j$, $j = k+1, \ldots, n$, then the matrix

$$P = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_k & \mathbf{v}_{k+1} & \mathbf{u}_{k+1} & \cdots & \mathbf{v}_n & \mathbf{u}_n \end{bmatrix}$$

is invertible and

$$P^{-1}AP = \operatorname{diag}[\lambda_1, \ldots, \lambda_k, B_{k+1}, \ldots, B_n]$$

where the 2×2 blocks

$$B_j = \begin{bmatrix} a_j & -b_j \\ b_j & a_j \end{bmatrix}$$

for $j = k+1, \dots, n$. We illustrate this result with an example.

Example 2.6.2 *The matrix*

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

has eigenvalues $\lambda_1 = -3, \lambda_2 = 2+i$ (and $\bar{\lambda}_2 = 2-i$). The corresponding eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{w}_2 = u_2 + i\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1+i \\ 1 \end{bmatrix}$$

Thus

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

and

$$P^{-1}AP = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{bmatrix}$$

The solution of the initial value problem (1) is given by

$$\begin{aligned} \mathbf{x}(t) &= P \begin{bmatrix} e^{-3t} & 0 & 0 \\ 0 & e^{2t}\cos t & -e^{2t}\sin t \\ 0 & e^{2t}\sin t & e^{2t}\cos t \end{bmatrix} P^{-1}x_0 \\ &= \begin{bmatrix} e^{-3t} & 0 & 0 \\ 0 & e^{2t}(\cos t + \sin t) & -2e^{2t}\sin t \\ 0 & e^{2t}\sin t & e^{2t}(\cos t - \sin t) \end{bmatrix} x_0 \end{aligned}$$

The stable subspace E^s is the x_1-axis and the unstable subspace E^u is the x_2, x_3 plane. The phase portrait is given in Figure 1.

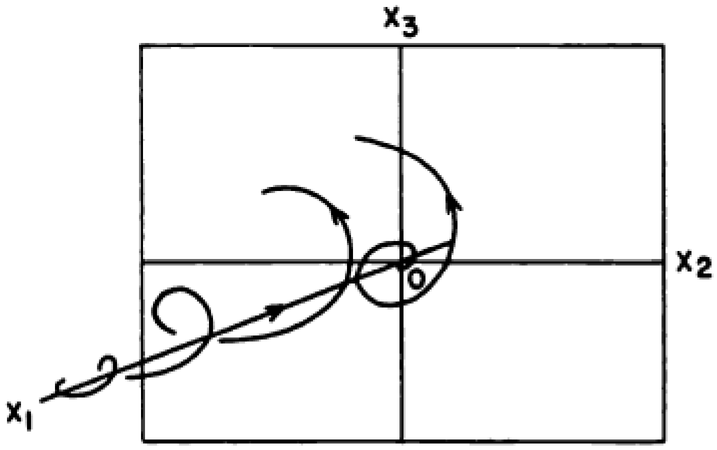

Figure 1

2.7 Multiple Eigenvalues

The fundamental theorem for linear systems in Section 1.4 tells us that the solution of the linear system (2.1) together with the initial condition $\mathbf{x}(0) = x_0$ is given by

$$\mathbf{x}(t) = e^{At}\mathbf{x}_0$$

We have seen how to find the $n \times n$ matrix e^{At} when A has distinct eigenvalues. We now complete the picture by showing how to find e^{At}, i.e., how to solve the linear system (1), when A has multiple eigenvalues.

Definition 1 Let λ be an eigenvalue of the $n \times n$ matrix A of multiplicity $m \leq n$. Then for $k = 1, \ldots, m$, any nonzero solution $\mathbf{v}$ of

$$(A - \lambda I)^k \mathbf{v} = 0$$

is called a generalized eigenvector of A.

Definition 2 An $n \times n$ matrix N is said to be nilpotent of order k if $N^{k-1} \neq 0$ and $N^k = 0$.

The following theorem is proved, for example, in Appendix III of Hirsch and Smale [H/S].

Theorem 1. Let A be a real $n \times n$ matrix with real eigenvalues $\lambda_1, \ldots, \lambda_n$ repeated according to their multiplicity. Then there exists a basis of generalized eigenvectors for $\mathbf{R}^n$. And if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is any basis of generalized eigenvectors for $\mathbf{R}^n$, the matrix $P = [\mathbf{v}_1 \cdots \mathbf{v}_n]$ is invertible,

$$A = S + N$$

where

$$P^{-1}SP = \text{diag}\left[\lambda_j\right]$$

the matrix $N = A - S$ is nilpotent of order $k \le n$, and S and N commute, i.e., $SN = NS$.

This theorem together with the propositions in Section 1.3 and the fundamental theorem in Section 1.4 then lead to the following result.

Corollary 1. Under the hypotheses of the above theorem, the linear system (2.1), together with the initial condition $\mathbf{x}(0) = \mathbf{x}_0$, has the solution

$$\mathbf{x}(t) = P\,\text{diag}\left[e^{\lambda_j t}\right]P^{-1}\left[I + Nt + \cdots + \frac{N^{k-1}t^{k-1}}{(k-1)!}\right]\mathbf{x}_0$$

If λ is an eigenvalue of multiplicity n of an $n \times n$ matrix A, then the above results are particularly easy to apply since in this case

$$S = \text{diag}[\lambda]$$

with respect to the usual basis for $\mathbf{R}^n$ and

$$N = A - S$$

The solution to the initial value problem (1) together with $\mathbf{x}(0) = \mathbf{x}_0$ is therefore given by

$$\mathbf{x}(t) = e^{\lambda t}\left[I + Nt + \cdots + \frac{N^k t^k}{k!}\right]\mathbf{x}_0$$

Let us consider two examples where the $n \times n$ matrix A has an eigenvalue of multiplicity n. In these examples, we do not need to compute a basis of generalized eigenvectors to solve the initial value problem!

Theorem 2. Let A be a real $2n \times 2n$ matrix with complex eigenvalues $\lambda_j = a_j + ib_j$ and $\bar{\lambda}_j = a_j - ib_j$, $j = 1, \ldots, n$. Then there exists generalized complex eigenvectors $\mathbf{w}_j = \mathbf{u}_j + i\mathbf{v}_j$ and $\overline{\mathbf{w}}_j = \mathbf{u}_j - i\mathbf{v}_j, i = 1, \ldots, n$ such that $\{\mathbf{u}_1, \mathbf{v}_1, \ldots, \mathbf{u}_n, \mathbf{v}_n\}$ is a basis for $\mathbf{R}^{2n}$. For any such basis, the matrix $P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{u}_1 & \cdots & \mathbf{v}_n & \mathbf{u}_n \end{bmatrix}$ is invertible,

$$A = S + N$$

where

$$P^{-1}SP = \text{diag}\begin{bmatrix} a_j & -b_j \\ b_j & a_j \end{bmatrix}$$

the matrix $N = A - S$ is nilpotent of order $k \le 2n$, and S and N commute.

The next corollary follows from the fundamental theorem in Section 1.4 and the results in Section 1.3:

Corollary 2. Under the hypotheses of the above theorem, the solution of the initial value problem (2.1), together with $\mathbf{x}(0) = \mathbf{x}_0$, is given by

$$\mathbf{x}(t) = P \operatorname{diag} e^{a_j t} \begin{bmatrix} \cos b_j t & -\sin b_j t \\ \sin b_j t & \cos b_j t \end{bmatrix} P^{-1} \left[I + \cdots + \frac{N^k t^k}{k!} \right] \mathbf{x}_0$$

2.8 Stability Theory

In this section we define the stable, unstable and center subspace, E^s, E^u and E^c respectively, of a linear system (2.1). Recall that E^s and E^u were defined in Section 1.2 in the case when A had distinct eigenvalues. We also establish some important properties of these subspaces in this section. Let $\mathbf{w}_j = \mathbf{u}_j + i\mathbf{v}_j$; be a generalized eigenvector of the (real) matrix A corresponding to an eigenvalue $\lambda_j = a_j + ib_j$. Note that if $b_j = 0$ then $\mathbf{v}_j = \mathbf{0}$. And let

$$B = \{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \mathbf{v}_{k+1}, \ldots, \mathbf{u}_m, \mathbf{v}_m\}$$

be a basis of $\mathbf{R}^n$ (with $n = 2m - k$) as established by Theorems 1 and 2 and the Remark in Section 1.7.

Definition 1. Let $\lambda_j = a_j + ib_j, w_j = u_j + i\mathbf{v}_j$ and B be as described above. Then

$$E^a = \operatorname{Span}\left\{\mathbf{u}_j, \mathbf{v}_j \mid a_j < 0\right\}$$
$$E^c = \operatorname{Span}\left\{\mathbf{u}_j, \mathbf{v}_j \mid a_j = 0\right\}$$

and

$$E^u = \operatorname{Span}\left\{\mathbf{u}_j, \mathbf{v}_j \mid a_j > 0\right\};$$

i.e., E^s, E^c and E^u are the subspaces of $\mathbf{R}^n$ spanned by the real and imaginary parts of the generalized eigenvectors $\mathbf{w}_j$ corresponding to eigenvalues λ_j with negative, zero and positive real parts respectively.

Definition 2. If all eigenvalues of the $n \times n$ matrix A have nonzero real part, then the flow $e^{At}; \mathbf{R^n} \to \mathbf{R^n}$ is called a hyperbolic flow and (2.1) is called a hyperbolic linear system.

Definition 3. A subspace $E \subset \mathbf{R}^n$ is said to be invariant with respect to the flow $e^{At} : \mathbf{R}^n \to \mathbf{R}^n$ if $e^{At}E \subset E$ for all $t \in \mathbf{R}$.

We next show that the stable, unstable and center subspaces, E^s, E^u and E^c of (2.1) are invariant under the flow e^{At} of the linear system (2.1); i.e., any solution starting in E^u, E^u or E^c at time $t = 0$ remains in E^u, E^u or E^c respectively for all $t \in \mathbf{R}$.

Lemma. Let E be the generalized eigenspace of A corresponding to an eigenvalue λ. Then $AE \subset E$.

Proof. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a basis of generalized eigenvectors for E. Then given $\mathbf{v} \in E$,

$$\mathbf{v} = \sum_{j=1}^{k} c_j \mathbf{v}_j$$

and by linearity

$$A\mathbf{v} = \sum_{j=1}^{k} c_j A v_j$$

Now since each $\mathbf{v}_j$ satisfies

$$(A - \lambda I)^{k_j} \mathbf{v}_j = \mathbf{0}$$

for some minimal k_j, we have

$$(A - \lambda I)\mathbf{v}_j = \mathbf{V}_j$$

where $\mathbf{V}_j \in \mathrm{Ker}(A - \lambda I)^{k_j - 1} \subset E$. Thus, it follows by induction that $A\mathbf{v}_j = \lambda \mathbf{v}_j + \mathbf{V}_j \in E$ and since E is a subspace of $\mathbf{R}^\mathbf{n}$, it follows that

$$\sum_{j=1}^{k} c_j A v_j \in E$$

i.e., $A\mathbf{v} \in E$ and therefore $AE \subset$ E.

Theorem 1. Let A be a real $n \times n$ matrix. Then

$$\mathbf{R}^n = E^s \oplus E^u \oplus E^c$$

where E^s, E^u and E^c are the stable, unstable and center subspaces of (2.1) respectively; furthermore, E^a, E^u and E^c are invariant with respect to the flow e^{At} of (2.1) respectively.

Proof. Since $B = \{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \mathbf{v}_{k+1}, \ldots, \mathbf{u}_m, \mathbf{v}_m\}$ described at the beginning of this section is a basis for $\mathbf{R}^{\boldsymbol{n}}$, it follows from the definition of E^s, E^u and E^c that

$$\mathbf{R}^n = E^s \oplus E^u \oplus E^c$$

If $x_0 \in E^s$ then

$$\mathbf{x}_0 = \sum_{j=1}^{n_4} c_j \mathbf{V}_j$$

where $\mathbf{V}_j = \mathbf{v}_j$ or $\mathbf{u}_j$ and $\left\{\mathbf{V}_j\right\}_{j=1}^{\mathbf{n}_*} \subset B$ is a basis for the stable subspace E^s as described in Definition 1 .

Then by the linearity of e^{At}, it follows that

$$e^{At}\mathrm{x}_0 = \sum_{j=1}^{n_4} c_j e^{At}\mathbf{V}_j$$

But

$$e^{At}\mathbf{V}_j = \lim_{k\to\infty}\left[I + At + \cdots + \frac{A^k t^k}{k!}\right]\mathbf{V}_j \in E^*$$

since for $j = 1, \ldots, n_s$ by the above lemma $A^k\mathbf{V}_j \in E^s$ and since E^s is complete. Thus, for all $t \in \mathbf{R}$, $e^{At}x_0 \in E^s$ and therefore $e^{At}E^s \subset E^s$; i.e., E^s is invariant under the flow e^{At}. It can similarly be shown that E^w and E^c are invariant under the flow e^{At}.

We next generalize the definition of sinks and sources of two-dimensional systems given in Section 1.5.

Definition 4. If all of the eigenvalues of A have negative (positive) real parts, the origin is called a sink (source) for the linear system (2.1).

Theorem 2. The following statements are equivalent:

(a) For all $\mathbf{x}_0 \in \mathbf{R}^n$, $\lim_{t\to\infty} e^{At}\mathbf{x}_0 = 0$ and for $\mathbf{x}_0 \neq 0$, $\lim_{t\to-\infty}\left|e^{At}\mathbf{x}_0\right| = \infty$.

(b) All eigenvalues of A have negative real part.

(c) There are positive constants a, c, m and M such that for all $\mathbf{x}_0 \in \mathbf{R}^n$

$$\left|e^{At}\mathbf{x}_0\right| \leq Me^{-ct}\left|\mathbf{x}_0\right|$$

for $t \geq 0$ and

$$\left|e^{At}\mathbf{x}_0\right| \geq me^{-at}\left|\mathbf{x}_0\right|$$

for $t \leq 0$.

Proof (a $\Rightarrow$ b): If one of the eigenvalues $\lambda = a + ib$ has positive real part, $a > 0$, then by the theorem and corollary in Section 1.8, there exists an $\mathbf{x}_0 \in \mathbf{R}^n, \mathbf{x}_0 \neq \mathbf{0}$, such that $\left|e^{At}\mathbf{x}_0\right| \geq e^{at}\left|\mathbf{x}_0\right|$. Therefore $\left|e^{At}\mathbf{x}_0\right| \to \infty$ as $t \to \infty$ i.e.,

$$\lim_{t\to\infty} e^{At}\mathbf{x}_0 \neq 0.$$

And if one of the eigenvalues of A has zero real part, say $\lambda = ib$, then by the corollary in Section 1.8, there exists $\mathrm{x}_0 \in \mathbf{R}^n, \mathbf{x}_0 \neq 0$ such that at least one component of the solution is of the form $ct^k\cos bt$ or $ct^k\sin bt$ with $k \geq 0$. And once again

$$\lim_{t\to\infty} e^{At}\mathbf{x}_0 \neq 0.$$

Thus, if not all of the eigenvalues of A have negative real part, there exists $\mathbf{x}_0 \in \mathbf{R}^n$ such that $e^{A1}\mathbf{x}_0 \nrightarrow \mathbf{0}$ as $t \to \infty$; i.e., a $\Rightarrow$ b. (b $\Rightarrow$ c) : If all of the eigenvalues of A have negative real part, then it follows from the Jordan canonical form theorem and its corollary in Section 1.8 that there exist positive constants a, c, m and M such that for all $\mathbf{x}_0 \in \mathbf{R}^n$ $\left|e^{At}\mathbf{x}_0\right| \leq Me^{-ct}\left|\mathbf{x}_0\right|$ for $t \geq 0$ and $\left|e^{At}\mathbf{x}_0\right| \geq me^{-at}\left|\mathbf{x}_0\right|$ for $t \leq 0$. (c $\Rightarrow$ a):

If this last pair of inequalities is satisfied for all $x_0 \in \mathbf{R}^n$, it follows by taking the limit as $t \to \pm\infty$ on each side of the above inequalities that

$$\lim_{t\to\infty} \left|e^{At}\mathbf{x}_0\right| = 0 \text{ and that } \quad \lim_{t\to-\infty} \left|e^{At}\mathbf{x}_0\right| = \infty$$

for $x_0 \neq 0$. This completes the proof of Theorem 2.
The next theorem is proved in exactly the same manner as Theorem 2 above using the theorem and its corollary in Section 1.8.

Theorem 3. The following statements are equivalent:
(a) For all $\mathbf{x}_0 \in \mathbf{R}^n \cdot \lim_{t\to-\infty} e^{At}\mathbf{x}_0 = 0$ and for $\mathbf{x}_0 \neq 0, \lim_{t\to\infty} \left|e^{At}\mathbf{x}_0\right| = \infty$.
(b) All eigenvalues of A have positive real part.
(c) There are positive constants a, c, m and M such that for all $\mathbf{x}_0 \in \mathbf{R}^n$

$$\left|e^{At}\mathbf{x}_0\right| \leq Me^{ct} \left|\mathbf{x}_0\right|$$

for $t \leq 0$ and

$$\left|e^{At}x_0\right| \geq me^{at} \left|\mathbf{x}_0\right|$$

for $t \geq 0$.
Corollary. If $\mathbf{x}_0 \in E^s$, then $e^{At}\mathbf{x}_0 \in E^s$ for all $t \in \mathbf{R}$ and

$$\lim_{t\to\infty} e^{At}x_0 = 0.$$

And if $\mathrm{x}_0 \in E^u$, then $e^{At}\mathrm{x}_0 \in E^u$ for all $t \in \mathbf{R}$ and

$$\lim_{t\to-\infty} e^{At}x_0 = 0.$$

Thus, we see that all solutions of (1) which start in the stable manifold E^s of (1) remain in E^* for all t and approach the origin exponentially fast as $t \to \infty$; and all solutions of (1) which start in the unstable manifold E^u of (1) remain in E^u for all t and approach the origin exponentially fast as $t \to -\infty$.

2.9 Nonhomogeneous Linear Systems

In this section we solve the nonhomogeneous linear system

$$\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{b}(t) \tag{2.7}$$

where A is an $n \times n$ matrix and $\mathbf{b}(t)$ is a continuous vector valued function.

Definition. A fundamental matrix solution of (2.1) is any nonsingular $n \times n$ matrix function $\Phi(t)$ that seatisfies

$$\Phi'(t) = A\Phi(t) \text{ for all } t \in \mathbf{R}.$$

Note that according to the lemma in Section 1.4, $\Phi(t) = e^{At}$ is a fundamental matrix solution which satisfies $\Phi(0) = I$, the $n \times n$ identity matrix. Furthermore, any fundamental matrix solution $\Phi(t)$ of (2.7) is given by $\Phi(t) = e^{At}C$ for some nonsingular matrix C. Once we have found a fundamental matrix solution of (2.7), it is easy to solve the nonhomogeneous system (2.1). The result is given in the following theorem.

Theorem 1. If $\Phi(t)$ is any fundamental matrix solution of (2.7), then the solution of the nonhomogeneous linear system (2.1) and the initial condition $\mathbf{x}(0) = \mathbf{x}_0$ is unique and is given by

$$\mathbf{x}(t) = \Phi(t)\Phi^{-1}(0)\mathbf{x}_0 + \int_0^t \Phi(t)\Phi^{-1}(\tau)\mathbf{b}(\tau)d\tau. \tag{2.8}$$

Proof. For the function $\mathbf{x}(t)$ defined above,

$$\begin{aligned}\mathbf{x}'(t) =& \Phi'(t)\Phi^{-1}(0)\mathbf{x}_0 + \Phi(t)\Phi^{-1}(t)\mathbf{b}(t)\\ &+ \int_0^t \Phi'(t)\Phi^{-1}(\tau)\mathbf{b}(\tau)d\tau\end{aligned}$$

And since $\Phi(t)$ is a fundamental matrix solution of (2.1), it follows that

$$\begin{aligned}\mathbf{x}'(t) &= A\left[\Phi(t)\Phi^{-1}(0)\mathbf{x}_0 + \int_0^t \Phi(t)\Phi^{-1}(\tau)\mathbf{b}(\tau)d\tau\right] + \mathbf{b}(t)\\ &= A\mathbf{x}(t) + \mathbf{b}(t)\end{aligned}$$

for all $t \in \mathbf{R}$. And this completes the proof of the theorem.

Remark 1. If the matrix A in (2.7) is time dependent, $A = A(t)$, then exactly the same proof shows that the solution of the nonhomogenous linear system (2.7) and the initial condition $x(0) = x_0$ is given by (2.8) provided that $\Phi(t)$ is a fundamental matrix solution of (2.1) with a variable coefficient matrix $A = A(t)$. For the most part, we do not consider solutions of (2.1) with $A = A(t)$ in this book. The reader should consult [C/L], [H] or [W] for a discussion of this topic which requires series methods and the theory of special functions.

Remark 2. With $\Phi(t) = e^{At}$, the solution of the nonhomogeneous linear system (2.7), as given in the above theorem, has the form

$$\mathbf{x}(t) = e^{At}\mathbf{x}_0 + e^{At}\int_0^t e^{-A\tau}\mathbf{b}(\tau)d\tau.$$

CHAPTER 3

NONLINEAR SYSTEMS: LOCAL THEORY

In Chapter 1 we saw that any linear system (2.1) has a unique solution through each point x_0 in the phase space $\mathbf{R}^n$; the solution is given by $\mathbf{x}(t) = e^{At}\mathbf{x}_0$ and it is defined for all $t \in \mathbf{R}$. In this chapter we begin our study of nonlinear systems of differential equations

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{2}$$

where $\mathbf{f} : E \to \mathbf{R}^n$ and E is an open subset of $\mathbf{R}^n$. We show that under certain conditions on the function f, the nonlinear system (2) has a unique solution through each point $\mathrm{x}_0 \in E$ defined on a maximal interval of existence $(\alpha, \beta) \subset \mathbf{R}$. In general, it is not possible to solve the nonlinear system (2); however, a great deal of qualitative information about the local behavior of the solution is determined in this chapter. In particular, we establish the Hartman-Grobman Theorem and the Stable Manifold Theorem which show that topologically the local behavior of the nonlinear system (2) near an equilibrium point x_0 where $f(x_0) = 0$ is typically determined by the behavior of the linear system (2.1) near the origin when the matrix $A = D\mathrm{f}(\mathrm{x}_0)$, the derivative of $\mathbf{f}$ at $\mathbf{x}_0$. We also discuss some of the ramifications of these theorems for two-dimensional systems when $\det D\mathbf{f}(\mathrm{x}_0) \neq 0$ and cite some of the local results of Andronov et al. [A-I] for planar systems (2) with $\det D\mathbf{f}(\mathbf{x}_0) = 0$.

3.1 Some Preliminary Concepts and Definitions

Before beginning our discussion of the fundamental theory of nonlinear systems of differential equations, we present some preliminary concepts and definitions. First of all, in this book we shall only consider autonomous systems of ordinary differential equations (2) as opposed to nonautonomous systems

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \tag{3}$$

where the function f can depend on the independent variable t; however, any nonautonomous system (3) with $\mathbf{x} \in \mathbf{R}^n$ can be written as an autonomous system (2) with $\mathbf{x} \in \mathbf{R}^{n+1}$ simply by letting $x_{n+1} = t$ and $\dot{x}_{n+1} = 1$. The fundamental theory for (2) and (3) does not differ significantly although it is possible to obtain the existence and uniqueness of solutions of (3) under slightly weaker hypotheses on $\mathbf{f}$ as a function of t; cf. for example Coddington and Levinson [C/L].

Notice that the existence of the solution of the elementary differential equation (2) is given by

$$x(t) = x(0) + \int_0^t f(s)ds$$

if $f(t)$ is integrable. And in general, the differential equations (2) or (3) will have a solution if the function f is continuous; cf. $[C/L]$, p. 6. However, continuity of the function $\mathbf{f}$ in (2) is not sufficient to guarantee uniqueness of the solution as the next example shows.

Example 1. The initial value problem

$$\dot{x} = 3x^{2/3}$$
$$x(0) = 0$$

has two different solutions through the point $(0, 0)$, namely

$$u(t) = t^3$$

and

$$v(t) \equiv 0$$

for all $t \in \mathbf{R}$. Clearly, each of these functions satisfies the differential equation for all $t \in \mathbf{R}$ as well as the initial condition $x(0) = 0$. (The first solution $u(t) = t^3$ can be obtained by the method of separation of variables). Notice that the function $f(x) = 3x^{2/3}$ is continuous at $x = 0$ but that it is not differentiable there.

Another feature of nonlinear systems that differs from linear systems is that even when the function f in (2) is defined and continuous for all $\mathbf{x} \in \mathbf{R^n}$, the solution $\mathbf{x}(t)$ may become unbounded at some finite time $t = \beta$; i.e., the solution may only exist on some proper subinterval $(\alpha, \beta) \subset \mathbf{R}$. This is illustrated by the next example.

Example 2. Consider the initial value problem

$$\dot{x} = x^2$$
$$x(0) = 1.$$

The solution, which can be found by the method of separation of variables, is given by

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

This solution is only defined for $t \in (-\infty, 1)$ and

$$\lim_{t \to 1^-} x(t) = \infty.$$

The interval $(-\infty, 1)$ is called the maximal interval of existence of the solution of this initial value problem. Notice that the function $x(t) = (1-\ t)^{-1}$ has another branch defined on the interval $(1, \infty)$; however, this branch is not considered as part of the solution of the initial value problem since the initial time $t = 0 \notin (1, \infty)$.

Before stating and proving the fundamental existence-uniqueness theorem for the nonlinear system (1), it is first necessary to define some terminology and notation concerning the derivative $D\mathbf{f}$ of a function $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^n$.

Definition 1. The function $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^n$ is differentiable at $\mathbf{x}_0 \in \mathbf{R}^n$ if there is a linear transformation $D\mathbf{f}(\mathrm{x}_0) \in L(\mathbf{R^n})$ that satisfies

$$\lim_{|\mathbf{h}| \to 0} \frac{|\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{x}_0) - D\mathbf{f}(\mathbf{x}_0)\,\mathbf{h}|}{|\mathbf{h}|} = 0$$

The linear transformation $D\mathbf{f}(\mathbf{x}_0)$ is called the derivative of $\mathbf{f}$ at $\mathbf{x}_0$.

Theorem 1. If $\mathbf{f} : \mathbf{R}^{\boldsymbol{n}} \to \mathbf{R^n}$ is differentiable at $\mathbf{x}_0$, then the partial derivatives $\dfrac{\partial f_i}{\partial x_j}, i, j = 1, \ldots, n$, all exist at x_0 and for all $\mathbf{x} \in \mathbf{R}^n$,

$$D\mathbf{f}(\mathbf{x}_0)\,\mathbf{x} = \sum_{j=1}^{n} \frac{\partial \mathbf{f}}{\partial x_j}(\mathbf{x}_0)\, x_j.$$

Thus, if $\mathbf{f}$ is a differentiable function, the derivative $D\mathbf{f}$ is given by the $n \times n$ Jacobian matrix

$$D\mathbf{f} = \left[\frac{\partial f_i}{\partial x_j}\right]$$

Definition 2. Suppose that V_1 and V_2 are two normed linear spaces with respective norms $\|\cdot\|_1$ and $\|\cdot\|_2$; i.e., V_1 and V_2 are linear spaces with norms $\|\cdot\|_1$ and $\|\cdot\|_2$ satisfying a-c in Section 1.3 of Chapter 1 . Then

$$\mathrm{F}\colon V_1 \to V_2$$

is continuous at $\mathrm{x}_0 \in V_1$ if for all $\varepsilon > 0$ there exists a $\delta > 0$ such that $\mathbf{x} \in V_1$ and $\|\mathbf{x} - \mathbf{x}_0\|_1 < \delta$ implies that

$$\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{x}_0)\|_2 < \varepsilon.$$

And $\mathbf{F}$ is said to be continuous on the set $E \subset V_1$ if it is continuous at each point $\mathbf{x} \in E$. If $\mathbf{F}$ is continuous on $E \subset V_1$, we write $\mathbf{F} \in C(E)$.

Definition 3. Suppose that $\mathbf{f} : E \to \mathbf{R}^n$ is differentiable on E. Then $\mathbf{f} \in C^1(E)$ if the derivative $D\mathbf{f} : E \to L(\mathbf{R}^n)$ is continuous on E.

The next theorem, gives a simple test for deciding whether or not a function $\mathrm{f} : E \to \mathbf{R}^n$ belongs to $C^1(E)$.

Theorem 2. Suppose that E is an open subset of $\mathbf{R^n}$ and that $\mathbf{f} : E \to \mathbf{R}^n$. Then $\mathrm{f} \in C^1(E)$ iff the partial derivatives $\dfrac{\partial f_i}{\partial x_j}, i, j = 1, \ldots, n$, exist and are continuous on E.

Remark 1. For E an open subset of $\mathbf{R}^n$, the higher order derivatives $D^k\mathbf{f}(\mathbf{x}_0)$ of a function $\mathbf{f} : E \to \mathbf{R}^n$ are defined in a similar way and it can be shown that $\mathrm{f} \in C^k(E)$ if and only if the partial derivatives

$$\frac{\partial^k f_i}{\partial x_{j_1} \cdots \partial x_{j_k}}$$

with $i, j_1, \ldots, j_k = 1, \ldots, n$, exist and are continuous on E. Furthermore, $D^2\mathbf{f}(\mathbf{x}_0) : E \times E \to \mathbf{R}^n$ and for $(\mathbf{x}, \mathbf{y}) \in E \times E$ we have

$$D^2\mathbf{f}(\mathbf{x}_0)(\mathbf{x}, \mathbf{y}) = \sum_{j_1, j_2 = 1}^{n} \frac{\partial^2 \mathbf{f}(\mathbf{x}_0)}{\partial x_{j_1} \partial x_{j_2}} x_{j_1} y_{j_2}.$$

Similar formulas hold for $D^k\mathbf{f}(\mathbf{x}_0) : (E \times \cdots \times E) \to \mathbf{R}^n$.

A function $\mathbf{f} : E \to \mathbf{R}^n$ is said to be analytic in the open set $E \subset \mathbf{R}^n$ if each component $f_j(\mathbf{x})$, $j = 1, \ldots, n$, is analytic in E, i.e., if for $j = 1, \ldots, n$ and $\mathbf{x}_0 \in E$, $f_j(\mathbf{x})$ has a Taylor series which converges to $f_j(\mathbf{x})$ in some neighborhood of $\mathbf{x}_0$ in E.

3.2 The Fundamental Existence-Uniqueness Theorem

In this section, we establish the fundamental existence-uniqueness theorem for a nonlinear autonomous system of ordinary differential equations (2) under the hypothesis that $\mathbf{f} \in C^1(E)$ where E is an open subset of $\mathbf{R^n}$. Picard's classical method of successive approximations is used to prove this theorem. The more modern approach based on the contraction mapping principle is relegated to the problems at the end of this section. The method of successive approximations not only allows us to establish the existence and uniqueness of the solution of the initial value problem associated with (2), but it also allows us to establish the continuity and differentiability of the solution with respect to initial conditions and parameters. This is done in the next section. The method is also used in the proof of the Stable Manifold and in the proof of the Hartman-Grobman. The method of successive approximations is one of the basic tools used in the qualitative theory of ordinary differential equations.

Definition 1. Suppose that $\mathbf{f} \in C(E)$ where E is an open subset of $\mathbf{R}^n$. Then $\mathbf{x}(t)$ is a solution of the differential equation (2) on an interval I if $\mathbf{x}(t)$ is differentiable on I and if for all $t \in I, \mathbf{x}(t) \in E$ and

$$\mathbf{x}'(t) = \mathbf{f}(\mathbf{x}(t))$$

And given $x_0 \in E, x(t)$ is a solution of the initial value problem

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) \\ \mathbf{x}(t_0) &= \mathbf{x}_0 \end{aligned}$$

on an interval I if $t_0 \in I, \mathbf{x}(t_0) = \mathbf{x}_0$ and $\mathbf{x}(t)$ is a solution of the differential equation (2) on the interval I.

In order to apply the method of successive approximations to establish the existence of a solution of (2), we need to define the concept of a Lipschitz condition and show that C^1 functions are locally Lipschitz.

Definition 2. Let E be an open subset of $\mathbf{R}^n$. A function $\mathrm{f} : E \to \mathbf{R}^n$ is said to satisfy a Lipschitz condition on E if there is a positive constant K such that for all $x, y \in E$

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| \leq K|\mathbf{x} - \mathbf{y}|.$$

The function $\mathbf{f}$ is said to be locally Lipschitz on E if for each point $\mathbf{x}_0 \in E$ there is an ε-neighborhood of $\mathrm{x}_0, N_\varepsilon(\mathrm{x}_0) \subset E$ and a constant $K_0 > 0$ such that for all $\mathbf{x}, \mathbf{y} \in N_\varepsilon(x_0)$

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| \leq K_0|\mathbf{x} - \mathbf{y}|.$$

By an ε-neighborhood of a point $\mathbf{x}_0 \in \mathbf{R}^n$, we mean an open ball of positive radius ε; i.e.,

$$N_\varepsilon(x_0) = \{\mathbf{x} \in \mathbf{R}^n \| \mathbf{x} - \mathrm{x}_0 \mid < \varepsilon\}.$$

Lemma. Let E be an open subset of $\mathbf{R}^\mathbf{n}$ and let $\mathbf{f} : E \to \mathbf{R}^n$. Then, if $\mathbf{f} \in C^1(E)$, $\mathbf{f}$ is locally Lipschitz on E.

Proof. Since E is an open subset of $\mathbf{R}^{\boldsymbol{n}}$, given $\mathrm{x}_0 \in E$, there is an $\varepsilon > 0$ such that $N_\varepsilon(\mathrm{x}_0) \subset E$. Let

$$K = \max_{|\mathrm{x}-\mathrm{x}_0| \le \varepsilon/2} \|D\mathbf{f}(\mathrm{x})\|,$$

the maximum of the continuous function $D\mathbf{f}(\mathbf{x})$ on the compact set $\mid x - \mathrm{x}_0 \mid \le \varepsilon/2$. Let N_0 denote the $\varepsilon/2$-neighborhood of $\mathrm{x}_0, N_{\varepsilon/2}(\mathrm{x}_0)$. Then for $\mathbf{x}, \mathbf{y} \in N_0$, set $\mathbf{u} = \mathbf{y} - \mathbf{x}$. It follows that $\mathbf{x} + s\mathbf{u} \in N_0$ for $0 \le s \le 1$ since N_0 is a convex set. Define the function $F : [0, 1] \to \mathbf{R}^n$ by

$$\mathbf{F}(s) = \mathbf{f}(\mathbf{x} + s\mathbf{u}).$$

Then by the chain rule,

$$\mathbf{F}'(s) = D\mathbf{f}(\mathbf{x} + s\mathbf{u})\mathbf{u}$$

and therefore

$$\begin{aligned}
\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) &= \mathbf{F}(1) - \mathbf{F}(0) \\
&= \int_0^1 \mathbf{F}'(s)ds = \int_0^1 D\mathbf{f}(\mathbf{x} + s\mathbf{u})\mathbf{u}ds.
\end{aligned}$$

It then follows from the lemma that

$$\begin{aligned}
|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})| &\le \int_0^1 |D\mathbf{f}(\mathbf{x} + s\mathrm{u})\mathrm{u}|ds \\
&\le \int_0^1 \|D\mathbf{f}(\mathbf{x} + s\mathrm{u})\||\mathbf{u}|ds \\
&\le K|\mathbf{u}| = K|\mathbf{y} - \mathbf{x}|.
\end{aligned}$$

And this proves the lemma. Picard's method of successive approximations is based on the fact that $\mathbf{x}(t)$ is a solution of the initial value problem

$$\begin{aligned}
\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) \\
\mathbf{x}(0) &= \mathbf{x}_0 \qquad (4)
\end{aligned}$$

if and only if $\mathbf{x}(t)$ is a continuous function that satisfies the integral equation

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_0^t \mathbf{f}(\mathbf{x}(s))ds.$$

The successive approximations to the solution of this integral equation are defined by the sequence of functions

$$\begin{aligned} \mathbf{u}_0(t) &= \mathbf{x}_0 \\ \mathbf{u}_{k+1}(t) &= \mathbf{x}_0 + \int_0^t \mathbf{f}\left(\mathbf{u}_k(s)\right) ds \end{aligned} \tag{5}$$

for $k = 0, 1, 2, \ldots$. In order to illustrate the mechanics involved in the method of successive approximations, we use the method to solve an elementary linear differential equation.

Definition 3. Let V be a normed linear space. Then a sequence $\{u_k\} \subset V$ is called a Cauchy sequence if for all $\varepsilon > 0$ there is an N such that $k, m \geq N$ implies that

$$\|u_k - \mathbf{u}_m\| < \varepsilon.$$

The space V is called complete if every Cauchy sequence in V converges to an element in V.

The following theorem, establishes the completeness of the normed linear space $C(I)$ with $I = [-a, a]$.

Theorem. For $I = [-a, a], C(I)$ is a complete normed linear space.
We can now prove the fundamental existence-uniqueness theorem for nonlinear systems.

Theorem (The Fundamental Existence-Uniqueness Theorem). Let E be an open subset of $\mathbf{R}^n$ containing $\mathbf{x}_0$ and assume that $\mathbf{f} \in C^1(E)$. Then there exists an $a > 0$ such that the initial value problem (4) has a unique solution $\mathbf{x}(t)$ on the interval $[-a, a]$.

Proof. Since $\mathrm{f} \in C^1(E)$, it follows from the lemma that there is an ε neighborhood $N_\varepsilon\left(\mathrm{x}_0\right) \subset E$ and a constant $K > 0$ such that for all $\mathbf{x}, \mathbf{y} \in N_\varepsilon\left(\mathbf{x}_0\right)$,

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| \leq K|\mathbf{x} - \mathbf{y}|.$$

Let $b = \varepsilon/2$. Then the continuous function $\mathbf{f}(\mathbf{x})$ is bounded on the compact set

$$N_0 = \left\{\mathbf{x} \in \mathbf{R}^n \| \mathbf{x} - \mathbf{x}_0 \mid \leq b\right\}.$$

Let

$$M = \max_{\mathbf{x} \in N_0} |\mathbf{f}(\mathbf{x})|.$$

Let the successive approximations $u_k(t)$ be defined by (5). Then assuming that there exists an $a > 0$ such that $u_k(t)$ is defined and continuous on $[-a, a]$ and satisfies

$$\max_{[-a,a]} |u_k(t) - x_0| \le b, \tag{6}$$

it follows that $\mathrm{f}(\mathrm{u}_k(t))$ is defined and continuous on $[-a, a]$ and therefore that

$$\mathbf{u}_{k+1}(t) = \mathrm{x}_0 + \int_0^t \mathrm{f}(\mathbf{u}_k(s))\, ds$$

is defined and continuous on $[-a, a]$ and satisfies

$$|\mathrm{u}_{k+1}(t) - \mathrm{x}_0| \le \int_0^t |\mathbf{f}(\mathrm{u}_k(s))|\, ds \le Ma$$

for all $t \in [-a, a]$. Thus, choosing $0 < a \le b/M$, it follows by induction that $u_k(t)$ is defined and continuous and satisfies (6) for all $t \in [-a, a]$ and $k = 1, 2, 3, \ldots$.

Next, since for all $t \in [-a, a]$ and $k = 0, 1, 2, 3, \ldots, \mathbf{u}_k(t) \in N_0$, it follows from the Lipschitz condition satisfied by f that for all $t \in [-a, a]$

$$\begin{aligned}
|\mathbf{u}_2(t) - \mathbf{u}_1(t)| &\le \int_0^t |\mathbf{f}(\mathbf{u}_1(s)) - \mathbf{f}(\mathbf{u}_0(s))|\, ds \\
&\le K \int_0^t |\mathbf{u}_1(s) - \mathbf{u}_0(s)|\, ds \\
&\le Ka \max_{[-a,a]} |\mathbf{u}_1(t) - \mathbf{x}_0| \\
&\le Kab.
\end{aligned}$$

And then assuming that

$$\max_{[-a,a]} \left|\mathbf{u}_j(t) - \mathbf{u}_{j-1}(t)\right| \le (Ka)^{j-1} b \tag{7}$$

for some integer $j \ge 2$, it follows that for all $t \in [-a, a]$

$$\begin{aligned}
\left|\mathbf{u}_{j+1}(t) - \mathbf{u}_j(t)\right| &\le \int_0^t \left|\mathbf{f}\left(\mathbf{u}_j(s)\right) - \mathbf{f}\left(\mathbf{u}_{j-1}(s)\right)\right| ds \\
&\le K \int_0^t \left|\mathbf{u}_j(s) - \mathbf{u}_{j-1}(s)\right| ds \\
&\le Ka \max_{[-a,a|} \left|\mathbf{u}_j(t) - \mathbf{u}_{j-1}(t)\right| \\
&\le (Ka)^j b.
\end{aligned}$$

Thus, it follows by induction that (7) holds for $j = 2, 3, \ldots$. Setting $\alpha =$ and choosing $0 < a < 1/K$, we see that for $m > k \geq N$ and $t \in [-a, a$

$$\begin{aligned}
|\mathbf{u}_m(t) - \mathbf{u}_k(t)| &\leq \sum_{j=k}^{m-1} \left|\mathbf{u}_{j+1}(t) - \mathbf{u}_j(t)\right| \\
&\leq \sum_{j=N}^{\infty} \left|\mathbf{u}_{j+1}(t) - \mathbf{u}_j(t)\right| \\
&\leq \sum_{j=N}^{\infty} \alpha^j b = \frac{\alpha^N}{1-\alpha} b.
\end{aligned}$$

This last quantity approaches zero as $N \to \infty$. Therefore, for all ε there exists an N such that $m, k \geq N$ implies that

$$\|\mathbf{u}_m - \mathbf{u}_k\| = \max_{|-a,a|} |\mathbf{u}_m(t) - \mathbf{u}_k(t)| < \varepsilon;$$

i.e., $\{u_k\}$ is a Cauchy sequence of continuous functions in $C([-a, a])$. I lows from the above theorem that $u_k(t)$ converges to a continuous funs $\mathbf{u}(t)$ uniformly for all $t \in [-a, a]$ as $k \to \infty$. And then taking the lim both sides of equation (3) defining the successive approximations, w that the continuous function

$$\mathbf{u}(t) = \lim_{k \to \infty} \mathbf{u}_k(t) \tag{8}$$

satisfies the integral equation

$$\mathbf{u}(t) = \mathbf{x}_0 + \int_0^t \mathbf{f}(\mathbf{u}(s))ds \tag{9}$$

for all $t \in [-a, a]$. We have used the fact that the integral and the limit can be interchanged since the limit in (8) is uniform for all $t \in [-a, a]$. Then since $\mathbf{u}(t)$ is continuous, $\mathbf{f}(\mathbf{u}(t))$ is continuous and by the fundamental theorem of calculus, the right-hand side of the integral equation (9) is differentiable and

$$\mathbf{u}'(t) = \mathbf{f}(\mathbf{u}(t))$$

for all $t \in [-a, a]$. Furthermore, $\mathbf{u}(0) = \mathbf{x}_0$ and from (6) it follows that $\mathbf{u}(t) \in N_c(\mathbf{x}_0) \subset E$ for all $t \in [-a, a]$. Thus $\mathbf{u}(t)$ is a solution of the initial value problem (4) on $[-a, a]$. It remains to show that it is the only solution.

Let $\mathbf{u}(t)$ and $\mathbf{v}(t)$ be two solutions of the initial value problem (4) on $[-a, a]$. Then the continuous function $|\mathbf{u}(t) - \mathbf{v}(t)|$ achieves its maximum at some point $t_1 \in [-a, a]$. It follows that

$$\begin{aligned}
\|\mathbf{u} - \mathbf{v}\| &= \max_{[-a,a]} |\mathbf{u}(t) - \mathbf{v}(t)| \\
&= \left| \int_0^{t_1} \mathbf{f}(\mathbf{u}(s)) - \mathbf{f}(\mathbf{v}(s)) ds \right| \\
&\leq \int_0^{|t_1|} |\mathbf{f}(\mathbf{u}(s)) - \mathbf{f}(\mathbf{v}(s))| ds \\
&\leq K \int_0^{|t_1|} |\mathbf{u}(s) - \mathbf{v}(s)| ds \\
&\leq Ka \max |\mathbf{u}(t) - \mathbf{v}(t)| \\
&\leq Ka \|\mathbf{u} - \mathbf{v}\|
\end{aligned}$$

But $Ka < 1$ and this last inequality can only be satisfied if $\|\mathbf{u} - \mathbf{v}\| = 0$. Thus, $\mathbf{u}(t) = \mathbf{v}(t)$ on $[-a, a]$. We have shown that the successive approximations (5) converge uniformly to a unique solution of the initial value problem (4) on the interval $[-a, a]$ where a is any number satisfying $0 < a < \min\left(\frac{b}{M}, \frac{1}{K}\right)$

Remark. Exactly the same method of proof shows that the initial value problem

$$\begin{aligned}
\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) \\
\mathbf{x}(t_0) &= \mathbf{x}_0
\end{aligned}$$

has a unique solution on some interval $[t_0 - a, t_0 + a]$.

3.3 Dependence on Initial Conditions and Parameters

In this section we investigate the dependence of the solution of the initial value problem

$$\begin{aligned}
\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) \\
\mathbf{x}(0) &= \mathbf{y}
\end{aligned} \tag{1}$$

on the initial condition $\mathbf{y}$. If the differential equation depends on a parameter $\mu \in R^m$, i.e., if the function $f(\mathbf{x})$ in (1) is replaced by $f(\mathbf{x}, \mu)$, then the solution $\mathbf{u}(t, \mathbf{y}, \boldsymbol{\mu})$ will also depend on the parameter $\boldsymbol{\mu}$. Roughly speaking, the dependence of the solution $\mathbf{u}(t, \mathbf{y}, \boldsymbol{\mu})$ on the initial condition $\mathbf{y}$ and the parameter $\boldsymbol{\mu}$ is as continuous as the function f. In order to establish this type of continuous dependence of the solution on initial conditions and parameters, we first establish a result due to T.H. Gronwall.

Lemma (Gronwall). Suppose that $g(t)$ is a continuous real valued function that satisfies $g(t) \geq 0$ and

$$g(t) \leq C + K \int_0^t g(s)ds$$

for all $t \in [0, a]$ where C and K are positive constants. It then follows that for all $t \in [0, a]$,

$$g(t) \leq Ce^{Kt}$$

Proof

Let $G(t) = C + K \int_0^t g(s)ds$ for $t \in [0, a]$. Then $G(t) \geq g(t)$ and $G(t) > 0$ for all $t \in [0, a]$. It follows from the fundamental theorem of calculus that

$$G'(t) = Kg(t)$$

and therefore that

$$\frac{G'(t)}{G(t)} = \frac{Kg(t)}{G(t)} \leq \frac{KG(t)}{G(t)} = K$$

for all $t \in [0, a]$. And this is equivalent to saying that

$$\frac{d}{dt}(\log G(t)) \leq K$$

or

$$\log G(t) \leq Kt + \log G(0)$$

or

$$G(t) \leq G(0)e^{Kt} = Ce^{Kt}$$

for all $t \in [0, a]$, which implies that $g(t) \leq Ce^{Kt}$ for all $t \in [0, a]$.

Theorem (Dependence on Initial Conditions).

Let E be an open subset of $\mathbf{R}^n$ containing $\mathbf{x}_0$ and assume that $\mathbf{f} \in C^1(E)$. Then there exists an $a > 0$ and $a\delta > 0$ such that for all $\mathbf{y} \in N_\delta(\mathbf{x}_0)$ the initial value problem

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

$$\mathbf{x}(0) = \mathbf{y}$$

has a unique solution $\mathbf{u}(t, \mathbf{y})$ with $\mathbf{u} \in C^1(G)$ where $G = [-a, a] \times N_6(\mathbf{x}_0) \subset \mathbf{R}^{n+1}$; furthermore, for each

$\mathbf{y} \in N_\delta(\mathbf{x}_0)$, $\mathbf{u}(t, \mathbf{y})$ is a twice continuously differentiable function of t for $t \in [-a, a]$.

Proof.

Since $\mathrm{f} \in C^1(E)$, it follows from the lemma in Section 2.2 that there is an ε-neighborhood $N_\varepsilon(\mathbf{x}_0) \subset E$ and a constant $K > 0$ such that for all $\mathbf{x}$ and $\mathbf{y} \in N_\varepsilon(\mathbf{x}_0)$,

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| \leq K|\mathbf{x} - \mathbf{y}|$$

As in the proof of the fundamental existence theorem, let $N_0 = \{\mathbf{x} \in \mathbf{R}^n \mid |\mathbf{x} - \mathbf{x}_0| \leq \varepsilon/2\}$, let M_0 be the maximum of $|\mathbf{f}(\mathbf{x})|$ on N_0 and let M_1 be the maximum of $\|D\mathbf{f}(\mathbf{x})\|$ on N_0. Let $\delta = \varepsilon/4$, and for $\mathbf{y} \in N_\delta(\mathbf{x}_0)$ define the successive approximations $\mathbf{u}_k(t, \mathbf{y})$ as

$$\begin{aligned} \mathbf{u}_0(t, \mathbf{y}) &= \mathbf{y} \\ \mathbf{u}_{k+1}(t, \mathbf{y}) &= \mathbf{y} + \int_0^t \mathbf{f}(\mathbf{u}_k(s, \mathbf{y}))\, ds \end{aligned} \tag{2}$$

Assume that $\mathbf{u}_k(t, \mathbf{y})$ is defined and continuous for all $(t, \mathbf{y}) \in G = [-a, a] \times N_\delta(\mathbf{x}_0)$ and that for all $\mathbf{y} \in N_\delta(\mathbf{x}_0)$

$$\left\|\mathbf{u}_k(t, \mathbf{y}) - \mathbf{x}_0\right\| < \varepsilon/2 \tag{3}$$

where $\|\cdot\|$ denotes the maximum over all $t \in [-a, a]$. This is clearly satisfied for $k = 0$. And assuming this is true for k, it follows that $\mathbf{u}_{k+1}(t, \mathbf{y})$, defined by the above successive approximations, is continuous on G. This follows since a continuous function of a continuous function is continuous and since the above integral of the continuous function $\mathrm{f}(\mathbf{u}_k(s, y))$ is continuous in t by the fundamental theorem of calculus and also in $\mathbf{y}$; cf. Rudin [R] or Carslaw [C]. We also have

$$\left\|\mathbf{u}_{k+1}(t, \mathbf{y}) - \mathbf{y}\right\| \leq \int_0^t \left|\mathbf{f}(\mathbf{u}_k(s, \mathbf{y}))\right| ds \leq M_0 a$$

for $t \in [-a, a]$ and $\mathbf{y} \in N_\delta(\mathbf{x}_0) \subset N_0$. Thus, for $t \in [-a, a]$ and $\mathbf{y} \in N_\delta(\mathbf{x}_0)$ with $\delta = \varepsilon/4$, we have

$$\begin{aligned} \left\|\mathbf{u}_{k+1}(t, \mathbf{y}) - \mathbf{x}_0\right\| &\leq \left\|\mathbf{u}_{k+1}(t, \mathbf{y}) - \mathbf{y}\right\| + \left\|\mathbf{y} - \mathbf{x}_0\right\| \\ &\leq M_0 a + \varepsilon/4 < \varepsilon/2 \end{aligned}$$

provided $M_0 a < \varepsilon/4$, i.e., provided $a < \varepsilon/(4M_0)$. Thus, the above induction hypothesis holds for all $k = 1, 2, 3, \ldots$ and $(t, y) \in G$ provided $a < \varepsilon/(4M_0)$.

We next show that the successive approximations $\mathbf{u}_k(t, y)$ converge uniformly to a continuous function $\mathbf{u}(t, \mathbf{y})$ for all $(t, \mathbf{y}) \in G$ as $k \to \infty$. As in the proof of the fundamental existence theorem,

$$\begin{aligned}\left\|\mathbf{u}_2(t,\mathbf{y}) - \mathbf{u}_1(t,\mathbf{y})\right\| &\leq Ka\left\|\mathbf{u}_1(t,\mathbf{y}) - \mathbf{y}\right\| \\ &\leq Ka\left\|\mathbf{u}_1(t,\mathbf{y}) - \mathbf{x}_0\right\| + Ka\left\|\mathbf{y} - \mathbf{x}_0\right\| \\ &\leq Ka(\varepsilon/2 + \varepsilon/4) \leq Ka\varepsilon\end{aligned}$$

for $(t,\mathbf{y}) \in G$. And then it follows exactly as in the proof of the fundamental existence theorem in Section 2.2 that

$$\left\|\mathbf{u}_{k+1}(t,\mathbf{y}) - \mathbf{u}_k(t,\mathbf{y})\right\| \leq (Ka)^k\varepsilon$$

for $(t,\mathbf{y}) \in G$ and consequently that the successive approximations converge uniformly to a continuous function $\mathbf{u}(t,\mathbf{y})$ for $(t,\mathbf{y}) \in G$ as $k \to \infty$ provided $a < 1/K$. Furthermore, the function $\mathbf{u}(t,\mathbf{y})$ satisfies

$$\mathbf{u}(t,\mathbf{y}) = \mathbf{y} + \int_0^t \mathbf{f}(\mathbf{u}(s,\mathbf{y}))ds$$

for $(t,\mathbf{y}) \in G$ and also $\mathbf{u}(0,\mathbf{y}) = \mathbf{y}$. And it follows from the inequality (3) that $\mathbf{u}(t,\mathbf{y}) \in N_{\varepsilon/2}(\mathbf{x}_0)$ for all $(t,\mathbf{y}) \in G$. Thus, by the fundamental theorem of calculus and the chain rule, it follows that

$$\dot{\mathbf{u}}(t,\mathbf{y}) = \mathbf{f}(\mathbf{u}(t,\mathbf{y}))$$

and that

$$\ddot{\mathbf{u}}(t,\mathbf{y}) = D\mathbf{f}(\mathbf{u}(t,\mathbf{y}))\dot{\mathbf{u}}(t,\mathbf{y})$$

for all $(t,y) \in G$; i.e., $\mathbf{u}(t,y)$ is a twice continuously differentiable function of t which satisfies the initial value problem (1) for all $(t,y) \in G$. The uniqueness of the solution $\mathbf{u}(t,y)$ follows from the fundamental theorem in Section 2.2.

We now show that $\mathbf{u}(t,y)$ is a continuously differentiable function of $\mathbf{y}$ for all $(t,y) \in [-a,a] \times N_{\delta/2}(x_0)$. In order to do this, fix $\mathbf{y}_0 \in N_{\delta/2}(\mathbf{x}_0)$ and choose $\mathbf{h} \in \mathbf{R^n}$ such that $|\mathbf{h}| < \delta/2$. Then $\mathbf{y}_0 + \mathbf{h} \in N_\delta(\mathbf{x}_0)$. Let $\mathbf{u}(t,\mathbf{y}_0)$ and $\mathbf{u}(t,\mathbf{y}_0+\mathbf{h})$ be the solutions of the initial value problem (1) with $\mathbf{y} = \mathbf{y}_0$ and with $y = y_0 + \mathbf{h}$ respectively. It then follows that

$$\begin{aligned}\left|\mathbf{u}(t,\mathbf{y}_0+\mathbf{h}) - \mathbf{u}(t,\mathbf{y}_0)\right| &\leq |\mathbf{h}| + \int_0^t \left|\mathbf{f}(\mathbf{u}(s,\mathbf{y}_0+\mathbf{h})) - \mathbf{f}(\mathbf{u}(s,\mathbf{y}_0))\right| ds \\ &\leq |\mathbf{h}| + K\int_0^t \left|\mathbf{u}(s,\mathbf{y}_0+\mathbf{h}) - \mathbf{u}(s,\mathbf{y}_0)\right| ds\end{aligned}$$

for all $t \in [-a,a]$. Thus, it follows from Gronwall's Lemma that

$$\left|\mathbf{u}(t,\mathbf{y}_0+\mathbf{h}) - \mathbf{u}(t,\mathbf{y}_0)\right| \leq |\mathbf{h}|e^{K|t|} \tag{4}$$

for all $t \in [-a,a]$. We next define $\Phi(t,y_0)$ to be the fundamental matrix solution of the initial value problem

$$\dot{\Phi} = A\left(t, \mathbf{y}_0\right)\Phi$$
$$\Phi\left(0, \mathbf{y}_0\right) = I \tag{5}$$

with $A\left(t, y_0\right) = D\mathbf{f}\left(\mathbf{u}\left(t, \mathbf{y}_0\right)\right)$ and I the $n \times n$ identity matrix. The existence and continuity of $\Phi\left(t, y_0\right)$ on some interval $[-a, a]$ follow from the method of successive approximations as in problem 4 of Problem Set 2 and problem 4 in Problem Set 3. It then follows from the initial value problems for $\mathbf{u}\left(t, y_0\right)$, $\mathbf{u}\left(t, \mathbf{y}_0 + \mathbf{h}\right)$ and $\Phi\left(t, \mathbf{y}_0\right)$ and Taylor's Theorem,

$$\mathbf{f}(\mathbf{u}) - \mathbf{f}\left(\mathbf{u}_0\right) = D\mathbf{f}\left(\mathbf{u}_0\right)\left(\mathbf{u} - \mathbf{u}_0\right) + \mathbf{R}\left(\mathbf{u}, \mathbf{u}_0\right)$$

where $|\mathbf{R}\left(\mathbf{u}, \mathbf{u}_0\right)| / |\mathbf{u} - \mathbf{u}_0| \to 0$ as $|\mathbf{u} - \mathbf{u}_0| \to 0$, that

$$\begin{aligned}
| \mathbf{u}\left(t, \mathbf{y}_0\right) - \mathbf{u}\left(t, \mathbf{y}_0 + \mathbf{h}\right) + \Phi\left(t, \mathbf{y}_0\right)\mathbf{h} \Big| &\leq \int_0^t \Big| \mathbf{f}\left(\mathbf{u}\left(s, \mathbf{y}_0\right)\right) \\
&- \mathbf{f}\left(\mathbf{u}\left(s, \mathbf{y}_0 + \mathbf{h}\right)\right) + D\mathbf{f}\left(\mathbf{u}\left(s, \mathbf{y}_0\right)\right)\Phi\left(s, \mathbf{y}_0\right)\mathbf{h} \mid ds \\
&\leq \int_0^t \left\| D\mathbf{f}\left(\mathbf{u}\left(s, \mathbf{y}_0\right)\right)\right\| \left|\mathbf{u}\left(s, \mathbf{y}_0\right) - \mathbf{u}\left(s, \mathbf{y}_0 + \mathbf{h}\right) + \Phi\left(s, \mathbf{y}_0\right)\mathbf{h}\right| ds \\
&+ \int_0^t \left|\mathbf{R}\left(\mathbf{u}\left(s, \mathbf{y}_0 + \mathbf{h}\right), \mathbf{u}\left(s, \mathbf{y}_0\right)\right)\right| ds
\end{aligned} \tag{6}$$

Since $|\mathbf{R}\left(u, u_0\right)| / |u - u_0| \to 0$ as $|\mathbf{u} - \mathbf{u}_0| \to 0$ and since $\mathbf{u}(s, \mathbf{y})$ is continuous on G, it follows that given any $\varepsilon_0 > 0$, there exists a $\delta_0 > 0$ such that if $|\mathbf{h}| < \delta_0$ then $\left|\mathbf{R}\left(\mathbf{u}\left(s, \mathbf{y}_0\right), \mathbf{u}\left(s, \mathbf{y}_0 + \mathbf{h}\right)\right)\right| < \varepsilon_0 \left|\mathbf{u}\left(s, \mathbf{y}_0\right) - \mathbf{u}\left(s, \mathbf{y}_0 + \mathbf{h}\right)\right|$ for all $s \in [-a, a]$. Thus, if we let

$$g(t) = \left|\mathbf{u}\left(t, \mathbf{y}_0\right) - \mathbf{u}\left(t, \mathbf{y}_0 + \mathbf{h}\right) + \Phi\left(t, \mathbf{y}_0\right)\mathbf{h}\right|$$

it then follows from (4) and (6) that for all $t \in [-a, a]$, $y_0 \in N_{\delta/2}\left(x_0\right)$ and $|\mathbf{h}| < \min\left(\delta_0, \delta/2\right)$ we have

$$g(t) \leq M_1 \int_0^t g(s) ds + \varepsilon_0 |\mathbf{h}| a e^{Ka}.$$

Hence, it follows from Gronwall's Lemma that for any given $\varepsilon_0 > 0$

$$g(t) \leq \varepsilon_0 |\mathbf{h}| a \mathrm{e}^{Ka} e^{M_1 \mathrm{a}}$$

for all $t \in [-a, a]$ provided $|\mathbf{h}| < \min(\delta_0, \delta/2)$. Thus,

$$\lim_{|\mathbf{h}| \to 0} \frac{\left|\mathbf{u}(t, \mathbf{y}_0) - \mathbf{u}(t, \mathbf{y}_0 + \mathbf{h}) + \Phi(t, \mathbf{y}_0)\mathbf{h}\right|}{|\mathbf{h}|} = 0$$

uniformly for all $t \in [-a, a]$. Therefore, according to Definition 1 in Section 2.1,

$$\frac{\partial \mathbf{u}}{\partial \mathbf{y}}(t, \mathbf{y}_0) = \Phi(t, \mathbf{y}_0)$$

for all $t \in [-a, a]$ where $\Phi(t, y_0)$ is the fundamental matrix solution of the initial value problem (5) which is continuous in t and in y_0 for all $t \in [-a, a]$ and $y_0 \in N_{\delta/2}(x_0)$. This completes the proof of the theorem.

Corollary. Under the hypothesis of the above theorem,

$$\Phi(t, \mathbf{y}) = \frac{\partial \mathbf{u}}{\partial \mathbf{y}}(t, \mathbf{y})$$

for $t \in [-a, a]$ and $\mathbf{y} \in N_\delta(\mathrm{x}_0)$ if and only if $\Phi(t, y)$ is the fundamental matrix solution of

$$\Phi = D\mathbf{f}[\mathbf{u}(t, \mathbf{y})]\Phi$$
$$\Phi(0, \mathbf{y}) = I$$

for $t \in [-a, a]$ and $\mathrm{y} \in N_\delta(\mathrm{x}_0)$.

Remark 1. A similar proof shows that if $f \in C^r(E)$ then the solution $\mathbf{u}(t, \mathbf{y})$ of the initial value problem (1) is in $C^r(G)$ where G is defined as in the above theorem. And if $\mathbf{f}(\mathbf{x})$ is a (real) analytic function for $\mathbf{x} \in E$ then $\mathbf{u}(t, \mathbf{y})$ is analytic in the interior of G; cf. [C/L].

Remark 2. If x_0 is an equilibrium point of (1), i.e., if $f(x_0) = 0$ so that $\mathbf{u}(t, \mathbf{x}_0) = \mathbf{x}_0$ for all $t \in \mathbf{R}$, then

$$\Phi(t, \mathbf{x}_0) = \frac{\partial \mathrm{u}}{\partial \mathbf{x}_0}(t, \mathbf{x}_0)$$

satisfies

$$\dot{\Phi} = D\mathbf{f}(\mathbf{x}_0)\Phi$$
$$\Phi(0, \mathbf{x}_0) = I.$$

And according to the Fundamental Theorem for Linear Systems

$$\Phi(t, \mathrm{x}_0) = e^{D\mathbf{f}(\mathrm{x}_0)t}.$$

Remark 3. It follows from the continuity of the solution $\mathbf{u}(t, \mathbf{y})$ of the initial value problem (1) that for each $t \in [-a, a]$

$$\lim_{\mathbf{y}\to\mathbf{x}_0} \mathbf{u}(t, \mathbf{y}) = \mathbf{u}\left(t, \mathbf{x}_0\right).$$

It follows from the inequality (4) that this limit is uniform for all $t \in [-a, a]$. We prove a slightly stronger version of this result in Theorem 4 of the next section.

Theorem 2 (Dependence on Parameters). Let E be an open subset of $\mathbf{R}^{n+m}$ containing the point $(\mathbf{x}_0, \mu_0)$ where $\mathbf{x}_0 \in \mathbf{R}^n$ and $\mu_0 \in \mathbf{R}^m$ and assume that $\mathbf{f} \in C^1(E)$. It then follows that there exists an $a > 0$ and a $\delta > 0$ such that for all $\mathbf{y} \in N_\delta\left(\mathbf{x}_0\right)$ and $\mu \in N_\delta\left(\mu_0\right)$, the initial value problem

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \boldsymbol{\mu})$$
$$\mathbf{x}(0) = \mathbf{y}$$

has a unique solution $\mathbf{u}(t, \mathbf{y}, \boldsymbol{\mu})$ with $\mathbf{u} \in C^1(G)$ where $G = [-a, a] \times N_\delta\left(\mathbf{x}_0\right) \times N_\delta\left(\mu_0\right)$.

This theorem follows immediately from the previous theorem by replacing the vectors $\mathbf{x}_0, \mathbf{x}, \dot{\mathbf{x}}$ and $\mathbf{y}$ by the vectors $(\mathbf{x}_0, \mu_0), (\mathbf{x}, \mu), (\dot{\mathbf{x}}, \mathbf{0})$ and $(\mathbf{y}, \boldsymbol{\mu})$ or it can be proved directly using Gronwall's Lemma and the method of successive approximations.

Lemma 1. Let E be an open subset of $\mathbf{R}^n$ containing $\mathbf{x}_0$ and suppose $\mathbf{f} \in C^1(E)$. Let $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$ be solutions of the initial value problem (1) on the intervals I_1 and I_2. Then $0 \in I_1 \cap I_2$ and if I is any open interval containing 0 and contained in $I_1 \cap I_2$, it follows that $\mathbf{u}_1(t) = \mathbf{u}_2(t)$ for all $t \in I$. Proof. Since $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$ are solutions of the initial value problem (1) on I_1 and I_2 respectively, it follows from Definition 1 in Section 2.2 that $0 \in I_1 \cap I_2$. And if I is an open interval containing 0 and contained in $I_1 \cap I_2$, then the fundamental existence-uniqueness theorem in Section 2.2 implies that $\mathbf{u}_1(t) = \mathbf{u}_2(t)$ on some open interval $(-a, a) \subset I$. Let I^* be the union of all such open intervals contained in I. Then I^* is the largest open interval contained in I on which $\mathbf{u}_1(t) = \mathbf{u}_2(t)$. Clearly, $I^* \subset I$ and if I^* is a proper subset of I, then one of the endpoints t_0 of I^* is contained in $I \subset I_1 \cap I_2$. It follows from the continuity of $u_1(t)$ and $u_2(t)$ on I that

$$\lim_{t\to t_0} \mathbf{u}_1(t) = \lim_{t\to t_0} \mathbf{u}_2(t).$$

Call this common limit $\mathbf{u}_0$. It then follows from the uniqueness of solutions that $\mathbf{u}_1(t) = \mathbf{u}_2(t)$ on some interval $I_0 = (t_0 - a, t_0 + a) \subset I$. Thus, $\mathbf{u}_1(t) = \mathbf{u}_2(t)$ on the interval $I^* \cup I_0 \subset I$ and I^* is a proper subset of $I^* \cup I_0$. But this contradicts the fact that I^* is the largest open interval contained in I on which $\mathbf{u}_1(t) = \mathbf{u}_2(t)$. Therefore, $I^* = I$ and we have $\mathbf{u}_1(t) = \mathbf{u}_2(t)$ for all $t \in I$.

Theorem 1. Let E be an open subset of $\mathbf{R}^n$ and assume that $\mathbf{f} \in C^1(E)$. Then for each point $x_0 \in E$, there is a maximal interval J on which the initial value problem (1) has a unique solution, $\mathbf{x}(t)$; i.e., if the initial value problem has a solution $\mathbf{y}(t)$ on an interval I then $I \subset J$ and $\mathbf{y}(t) = \mathbf{x}(t)$ for all $t \in I$. Furthermore, the maximal interval J is open; i.e., $J = (\alpha, \beta)$.

Proof. By the fundamental existence-uniqueness theorem in Section 2.2, the initial value problem (1) has a unique solution on some open interval $(-a, a)$. Let (α, β) be the union of all open intervals I such that (1) has a solution on I. We define a function $\mathbf{x}(t)$ on (α, β) as follows: Given $t \in (\alpha, \beta)$, t belongs to some open interval I such that (1) has a solution $\mathbf{u}(t)$ on I; for this given $t \in (\alpha, \beta)$, define $\mathbf{x}(t) = \mathbf{u}(t)$. Then $\mathbf{x}(t)$ is a well-defined function of t since if $t \in I_1 \cap I_2$ where I_1 and I_2 are any two open intervals such that (1) has solutions $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$ on I_1 and I_2 respectively, then by the lemma $\mathbf{u}_1(t) = \mathbf{u}_2(t)$ on the open interval $I_1 \cap I_2$. Also, $\mathbf{x}(t)$ is a solution of (1) on (α, β) since each point $t \in (\alpha, \beta)$ is contained in some open interval I on which the initial value problem (1) has a unique solution $\mathbf{u}(t)$ and since $\mathbf{x}(t)$ agrees with $\mathbf{u}(t)$ on I. The fact that J is open follows from the fact that any solution of (1) on an interval $(\alpha, \beta]$ can be uniquely continued to a solution on an interval $(\alpha, \beta + a)$ with $a > 0$ as in the proof of Theorem 2 below.

Definition. The interval (α, β) in Theorem 1 is called the maximal interval of existence of the solution $\mathbf{x}(t)$ of the initial value problem (1) or simply the maximal interval of existence of the initial value problem (1).

Theorem 2. Let E be an open subset of $\mathbf{R}^n$ containing $\mathbf{x}_0$, let $\mathbf{f} \in C^1(E)$, and let (α, β) be the maximal interval of existence of the solution $\mathbf{x}(t)$ of the initial value problem (1). Assume that $\beta < \infty$. Then given any compact set $K \subset E$, there exists a $t \in (\alpha, \beta)$ such that $\mathbf{x}(t) \notin K$.

Proof. Since f is continuous on the compact set K, there is a positive number M such that $|\mathbf{f}(\mathbf{x})| \leq M$ for all $\mathbf{x} \in K$. Let $\mathbf{x}(t)$ be the solution of the initial value problem (1) on its maximal interval of existence (α, β) and assume that $\beta < \infty$ and that $\mathbf{x}(t) \in K$ for all $t \in (\alpha, \beta)$. We first show that $\lim_{t \to \beta^-} x(t)$ exists. If $\alpha < t_1 < t_2 < \beta$ then

$$|\mathbf{x}(t_1) - \mathbf{x}(t_2)| \leq \int_{t_1}^{t_2} |\mathbf{f}(\mathbf{x}(s))| ds \leq M |t_2 - t_1| .$$

Thus as t_1 and t_2 approach β from the left, $|\mathbf{x}(t_2) - \mathbf{x}(t_1)| \to 0$ which, by the Cauchy criterion for convergence in $\mathbf{R}^n$ (i.e., the completeness of $\mathbf{R}^n$) implies that $\lim_{t \to \beta^-} \mathbf{x}(t)$ exists. Let $\mathbf{x}_1 = \lim_{t \to \beta^-} \mathbf{x}(t)$. Then $\mathbf{x}_1 \in K \subset E$ since K is compact. Next define the function $\mathbf{u}(t)$ on $(\alpha, \beta]$ by

$$\mathbf{u}(t) = \begin{cases} \mathbf{x}(t) & \text{for} \quad t \in (\alpha, \beta) \\ \mathrm{x}_1 & \text{for} \quad t = \beta. \end{cases}$$

Then $\mathbf{u}(t)$ is differentiable on $(\alpha, \beta]$. Indeed,

$$\mathrm{u}(t) = \mathrm{x}_0 + \int_0^t \mathrm{f}(\mathrm{u}(s)) ds$$

which implies that

$$\mathbf{u}'(\beta) = \mathbf{f}(\mathbf{u}(\beta));$$

i.e., $\mathbf{u}(t)$ is a solution of the initial value problem (1) on $(\alpha, \beta]$. The function $\mathbf{u}(t)$ is called the continuation of the solution $\mathbf{x}(t)$ to $(\alpha, \beta]$. Since $\mathbf{x}_1 \in E$, it follows from the fundamental existence-uniqueness theorem in Section 2.2 that the initial value problem $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ together with $\mathbf{x}(\beta) = \mathbf{x}_1$ has a unique solution $\mathbf{x}_1(t)$ on some interval $(\beta - a, \beta + a)$. By the above lemma, $\mathbf{x}_1(t) = u(t)$ on $(\beta - a, \beta)$ and $\mathbf{x}_1(\beta) = u(\beta) = \mathbf{x}_1$. So if we define

$$\mathbf{v}(t) = \begin{cases} \mathbf{u}(t) & \text{for} \quad t \in (\alpha, \beta] \\ \mathbf{x}_1(t) & \text{for} \quad t \in [\beta, \beta + a), \end{cases}$$

then $\mathbf{v}(t)$ is a solution of the initial value problem (1) on $(\alpha, \beta + a)$. But this contradicts the fact that (α, β) is the maximal interval of existence for the initial value problem (1). Hence, if $\beta < \infty$, it follows that there exists a $t \in (\alpha, \beta)$ such that $\mathbf{x}(t) \notin K$. If (α, β) is the maximal interval of existence for the initial value problem (1) then $0 \in (\alpha, \beta)$ and the intervals $[0, \beta)$ and $(\alpha, 0]$ are called the right and left maximal intervals of existence respectively. Essentially the same proof yields the following result.

Theorem 3. Let E be an open subset of $\mathbf{R}^n$ containing $\mathbf{x}_0$, let $\mathbf{f} \in C^1(E)$, and let $[0, \beta)$ be the right maximal interval of existence of the solution $\mathbf{x}(t)$ of the initial value problem (1). Assume that $\beta < \infty$. Then given any compact set $K \subset E$, there exists a $t \in (0, \beta)$ such that $\mathbf{x}(t) \notin K$.

Corollary 1. Under the hypothesis of the above theorem, if $\beta < \infty$ and if $\lim_{t \to \beta^-} \mathbf{x}(t)$ exists then $\lim_{t \to \beta^-} \mathbf{x}(t) \in \dot{E}$.

Proof. If $\mathbf{x}_1 = \lim_{t \to \beta^-} \mathbf{x}(t)$, then the function

$$\mathbf{u}(t) = \begin{cases} \mathbf{x}(t) & \text{for} \quad t \in [0, \beta) \\ \mathbf{x}_1 & \text{for} \quad t = \beta \end{cases}$$

is continuous on $[0, \beta]$. Let K be the image of the compact set $[0, \beta]$ under the continuous map $\mathbf{u}(t)$; i.e.,

$$K = \{\mathbf{x} \in \mathbf{R^n} \mid \mathbf{x} = \mathbf{u}(t) \text{ for some } t \in [0, \beta]\}.$$

Then K is compact. Assume that $\mathbf{x}_1 \in E$. Then $K \subset E$ and it follows from Theorem 3 that there exists a $t \in (0, \beta)$ such that $\mathbf{x}(t) \notin K$. This is a contradiction and therefore $\mathbf{x}_1 \notin E$. But since $\mathbf{x}(t) \in E$ for all $t \in [0, \beta)$, it follows that $\mathbf{x}_1 = \lim_{t \to \beta^-} \mathbf{x}(t) \in \bar{E}$. Therefore $\mathbf{x}_1 \in \bar{E} \sim E$; i.e., $\mathbf{x}_1 \in \dot{E}$.

Corollary 2. Let E be an open subset of $\mathbf{R}^n$ containing $\mathbf{x}_0$, let $\mathbf{f} \in C^1(E)$, and let $[0, \beta)$ be the right maximal interval of existence of the solution $\mathbf{x}(t)$ of the initial value problem (1). Assume that there exists a compact set $K \subset E$ such that

$$\{\mathbf{y} \in \mathbf{R^n} \mid \mathbf{y} = \mathbf{x}(t) \text{ for some } t \in [0, \beta)\} \subset K.$$

It then follows that $\beta = \infty$; i.e. the initial value problem (1) has a solution $\mathbf{x}(t)$ on $[0, \infty)$.

Proof. This corollary is just the contrapositive of the statement in Theorem 3.

We next prove the following theorem which strengthens the result on uniform convergence with respect to initial conditions in Remark 3 of Section 2.3.

Theorem 4. Let E be an open subset of $\mathbf{R}^n$ containing $\mathbf{x}_0$ and let $\mathbf{f} \in C^1(E)$. Suppose that the initial value problem (1) has a solution $\mathbf{x}\,(t, \mathbf{x}_0)$ defined on a closed interval $[a, b]$. Then there exists a $\delta > 0$ and a positive constant K such that for all $\mathbf{y} \in N_\delta\,(\mathbf{x}_0)$ the initial value problem

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$
$$\dot{\mathbf{x}}(0) = \mathbf{y} \tag{2}$$

has a unique solution $\mathbf{x}(t, \mathbf{y})$ defined on $[a, b]$ which satisfies

$$\left|\mathbf{x}(t, \mathbf{y}) - \mathbf{x}\,(t, \mathbf{x}_0)\right| \leq \left|\mathbf{y} - \mathbf{x}_0\right| e^{K|t|}$$

and

$$\lim_{\mathbf{y} \to \mathbf{x}_0} \mathbf{x}(t, \mathbf{y}) = \mathbf{x}\,(t, \mathbf{x}_0)$$

uniformly for all $t \in [a, b]$.

Remark 1. If in Theorem 4 we have a function $f(x, \mu)$ depending on a parameter $\mu \in \mathbf{R}^m$ which satisfies f $\in C^1(E)$ where E is an open subset of $\mathbf{R}^{n+m}$ containing ($\mathbf{x}_0, \mu_0$), it can be shown that if for $\boldsymbol{\mu} = \mu_0$ the initial value problem (1) has a solution $\mathbf{x}\,(t, \mathbf{x}_0, \mu_0)$ defined on a closed interval $a \leq t \leq b$, then there is a $\delta > 0$ and a $K > 0$ such that for all y $\in N_\delta$ (x$_0$) and $\mu \in N_\delta\,(\mu_0)$ the initial value problem

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \boldsymbol{\mu})$$
$$\mathbf{x}(0) = \mathbf{y}$$

has a unique solution $\mathbf{x}(t, \mathbf{y}, \boldsymbol{\mu})$ defined for $a \leq t \leq b$ which satisfies

$$\left|\mathbf{x}(t, \mathbf{y}, \boldsymbol{\mu}) - \mathbf{x}\,(t, \mathbf{x}_0, \mu_0)\right| \leq [\![\left|\mathbf{y} - \mathbf{x}_0\right| + \left|\boldsymbol{\mu} - \boldsymbol{\mu}_0\right|]\!]\, e^{K|t|}$$

and

$$\lim_{(\mathbf{y}, \boldsymbol{\mu}) \to (\mathbf{x}_0, \mu_0)} \mathbf{x}(t, \mathbf{y}, \boldsymbol{\mu}) = \mathbf{x}\,(t, \mathbf{x}_0, \mu_0)$$

uniformly for all $t \in [a, b]$. In order to prove this theorem, we first establish the following lemma.

Lemma 2. Let E be an open subset of $\mathbf{R}^n$ and let A be a compact subset of E. Then if $\mathrm{f} : E \to \mathbf{R}^n$ is locally Lipschitz on E, it follows that $\mathbf{f}$ satisfies a Lipschitz condition on A.

Proof. Let M be the maximal value of the continuous function $\mathbf{f}$ on the compact set A. Suppose that f does not satisfy a Lipschitz condition on A. Then for every $K > 0$, we can find $\mathbf{x}, \mathbf{y} \in A$ such that

$$|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})| > K|\mathbf{y} - \mathbf{x}|.$$

In particular, there exist sequences $\mathbf{x}_n$ and $\mathbf{y}_n$ in A such that

$$\left|\mathbf{f}(\mathbf{y}_n) - \mathbf{f}(\mathbf{x}_n)\right| > n\left|\mathbf{y}_n - \mathbf{x}_n\right| \tag{*}$$

for $n = 1, 2, 3, \ldots$. Since A is compact, there are convergent subsequences, call them $\mathbf{x}_n$ and $\mathbf{y}_n$ for simplicity in notation, such that $\mathbf{x}_n \to \mathbf{x}^*$ and $\mathbf{y}_n \to \mathbf{y}^*$ with $\mathbf{x}^*$ and $\mathbf{y}^*$ in A. It follows that $\mathbf{x}^* = \mathbf{y}^*$ since for all $n = 1, 2, 3, \ldots$

$$\left|\mathbf{y}^* - \mathbf{x}^*\right| = \lim_{n\to\infty}\left|\mathbf{y}_n - \mathbf{x}_n\right| \le \frac{1}{n}\left|\mathbf{f}(\mathbf{y}_n) - \mathbf{f}(\mathbf{x}_n)\right| \le \frac{2M}{n}.$$

Now, by hypotheses, there exists a neighborhood N_0 of $\mathbf{x}^*$ and a constant K_0 such that f satisfies a Lipschitz condition with Lipschitz constant K_0 for all $\mathbf{x}$ and $\mathbf{y} \in N_0$. But since $\mathbf{x}_n$ and $\mathbf{y}_n$ approach $\mathbf{x}^*$ as $n \to \infty$, it follows that $\mathbf{x}_n$ and $\mathbf{y}_n$ are in N_0 for n sufficiently large; i.e., for n sufficiently large

$$\left|\mathbf{f}(\mathbf{y}_n) - \mathbf{f}(\mathbf{x}_n)\right| \le K\left|\mathbf{y}_n - \mathbf{x}_n\right|.$$

But for $n \ge K$, this contradicts the above inequality (*) and this establishes the lemma.

Proof (of Theorem 4). Since $[a, b]$ is compact and $\mathbf{x}(t, x_0)$ is a continuous function of t, $\{\mathbf{x} \in \mathbf{R^n} \mid \mathbf{x} = \mathbf{x}(t, \mathbf{x}_0)$ and $a \le t \le b\}$ is a compact subset of E. And since E is open, there exists an $\varepsilon > 0$ such that the compact set

$$A = \{\mathbf{x} \in \mathbf{R}^n \| \mathbf{x} - \mathbf{x}(t, \mathbf{x}_0)| \le \varepsilon \text{ and } a \le t \le b\}$$

is a subset of E. Since $\mathrm{f} \in C^1(E)$, it follows from the lemma in Section 2.2 that f is locally Lipschitz in E; and then by the above lemma, f satisfies a Lipschitz condition

$$|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})| \le K|\mathbf{y} - \mathbf{x}|$$

for all $\mathbf{x}, \mathbf{y} \in A$. Choose $\delta > 0$ so small that $\delta \le \varepsilon$ and $\delta \le \varepsilon e^{-K(b-a)}$. Let $\mathbf{y} \in N_\delta(\mathbf{x}_0)$ and let $\mathbf{x}(t, \mathbf{y})$ be the solution of the initial value problem (2) on its maximal interval of existence (α, β). We shall show that $[a, b] \subset (\alpha, \beta)$. Suppose that $\beta \le b$. It then follows that $\mathbf{x}(t, \mathbf{y}) \in A$ for all $t \in (\alpha, \beta)$ because if this were not true then there would exist a $t^* \in (\alpha, \beta)$ such that $\mathbf{x}(t, \mathbf{x}_0) \in A$ for $t \in (\alpha, t^*]$ and $\mathbf{x}(t^*, \mathbf{y}) \in \dot{A}$. But then

$$\begin{aligned}\left|\mathbf{x}(t, \mathbf{y}) - \mathbf{x}(t, \mathbf{x}_0)\right| &\leq \left|\mathbf{y} - \mathbf{x}_0\right| + \int_0^t \mid \mathbf{f}(\mathbf{x}(s, \mathbf{y})) - f(\mathbf{x}(s, \mathbf{x}_0) \mid ds \\ &\leq \left|\mathbf{y} - \mathbf{x}_0\right| + K \int_0^t \left|\mathbf{x}(s, \mathbf{y}) - \mathbf{x}(s, \mathbf{x}_0)\right| ds\end{aligned}$$

for all $t \in (\alpha, t^*]$. And then by Gronwall's Lemma in Section 2.3, it follows that

$$\left|\mathbf{x}(t^*, \mathbf{y}) - \mathbf{x}(t^*, \mathbf{x}_0)\right| \leq \left|\mathbf{y} - \mathbf{x}_0\right| e^{K|t^*|} < \delta e^{K(b-a)} < \varepsilon$$

since $t^* < \beta \leq b$. Thus $\mathbf{x}(t^*, \mathbf{y})$ is an interior point of A, a contradiction. Thus, $\mathbf{x}(t, y) \in A$ for all $t \in (\alpha, \beta)$. But then by Theorem 2, (α, β) is not the maximal interval of existence of $\mathbf{x}(t, \mathbf{y})$, a contradiction. Thus $b < \beta$. It is similarly proved that $\alpha < a$. Hence, for all $y \in N_\delta(x_0)$, the initial value problem (2) has a unique solution defined on $[a, b]$. Furthermore, if we assume that there is a $t^* \in [a, b)$ such that $\mathbf{x}(t, \mathbf{y}) \in A$ for all $t \in [a, t^*)$ and $\mathbf{x}(t^*, \mathbf{y}) \in \dot{A}$, a repeat of the above argument based on Gronwall's Lemma leads to a contradiction and shows that $\mathbf{x}(t, \mathbf{y}) \in A$ for all $t \in [a, b]$ and hence that

$$\left|\mathbf{x}(t, \mathbf{y}) - \mathbf{x}(t, \mathbf{x}_0)\right| \leq \left|\mathbf{y} - \mathbf{x}_0\right| e^{K|t|}$$

for all $t \in [a, b]$. It then follows that

$$\lim_{\mathbf{y} \to \mathbf{x}_0} \mathbf{x}(t, \mathbf{y}) = \mathbf{x}(t, \mathbf{x}_0)$$

uniformly for all $t \in [a, b]$.

3.4 The Flow Defined by a Differential Equation

In Section 1.9 of Chapter 1, we defined the flow, $e^{At} : \mathbf{R}^n \to \mathbf{R}^n$, of the linear system

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

The mapping $\phi_t = e^{\mathrm{At}}$ satisfies the following basic properties for all $x \in \mathbf{R}^n$:

(i) $\phi_0(\mathbf{x}) = \mathbf{x}$

(ii) $\phi_s\left(\phi_t(\mathbf{x})\right) = \phi_{s+t}(\mathbf{x})$ for all $s, t \in \mathbf{R}$

(iii) $\phi_{-t}\left(\phi_t(\mathbf{x})\right) = \phi_t\left(\phi_{-t}(\mathbf{x})\right) = \mathbf{x}$ for all $t \in \mathbf{R}$.

Property (i) follows from the definition of e^{At}, property (ii) follows from Proposition 2 in Section 1.3 of Chapter 1, and property (iii) follows from Corollary 2 in Section 1.3 of Chapter 1. In this section, we define the flow, ϕ_t, of the nonlinear system

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

and show that it satisfies these same basic properties. In the following definition, we denote the maximal interval of existence (α, β) of the solution $\phi(t, \mathbf{x}_0)$ of the initial value problem

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) \\ \mathbf{x}(0) &= \mathbf{x}_0 \end{aligned} \tag{2}$$

by $I(\mathrm{x}_0)$ since the endpoints α and β of the maximal interval generally depend on x_0; cf. problems 1(a) and (d) in Section 2.4.

Definition 1. Let E be an open subset of $\mathbf{R}^n$ and let $\mathbf{f} \in C^1(E)$. For $\mathbf{x}_0 \in E$, let $\phi(t, \mathbf{x}_0)$ be the solution of the initial value problem (2) defined on its maximal interval of existence $I(\mathrm{x}_0)$. Then for $t \in I(\mathrm{x}_0)$, the set of mappings ϕ_t defined by

$$\phi_t(\mathbf{x}_0) = \phi(t, \mathbf{x}_0)$$

is called the flow of the differential equation (1) or the flow defined by the differential equation (1); ϕ_t is also referred to as the flow of the vector field $\mathbf{f}(\mathbf{x})$.

If we think of the initial point x_0 as being fixed and let $I = I(x_0)$, then the mapping $\phi(\cdot, \mathbf{x}_0) : I \to E$ defines a solution curve or trajectory of the system (1) through the point $x_0 \in E$. As usual, the mapping $\phi(\cdot, x_0)$ is identified with its graph in $I \times E$ and a trajectory is visualized as a motion along a curve Γ through the point x_0 in the subset E of the phase space $\mathbf{R}^n$; cf. Figure 1. On the other hand, if we think of the point x_0 as varying throughout $K \subset E$, then the flow of the differential equation (1), $\phi_t : K \to E$ can be viewed as the motion of all the points in the set K; cf. Figure 2.

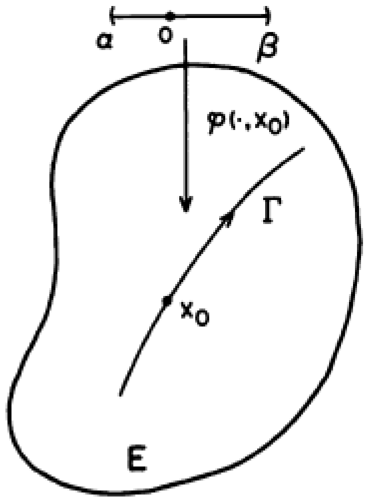

Figure 1. A trajectory Γ of the system (1).

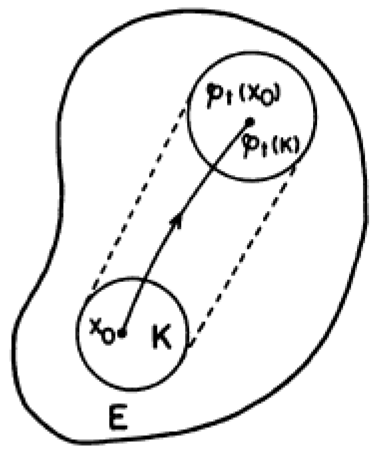

Figure 2. The flow ϕ_t of the system (1).

If we think of the differential equation (1) as describing the motion of a fluid, then a trajectory of (1) describes the motion of an individual particle in the fluid while the flow of the differential equation (1) describes the motion of the entire fluid. We now show that the basic properties (i)-(iii) of linear flows are also satisfied by nonlinear flows. But first we extend Theorem 1 of Section 2.3, establishing that $\phi(t, x_0)$ is a locally smooth function, to a global result. Using the same notation as in Definition 1, let us define the set $\Omega \subset \mathbf{R} \times E$ as

$$\Omega = \{(t, x_0) \in \mathbf{R} \times E \mid t \in I(\mathbf{x}_0)\}.$$

Theorem 1. Let E be an open subset of $\mathbf{R}^n$ and let $\mathbf{f} \in C^1(E)$. Then Ω is an open subset of $\mathbf{R} \times E$ and $\phi \in C^1(\Omega)$.

Proof. If $(t_0, x_0) \in \Omega$ and $t_0 > 0$, then according to the definition of the set Ω, the solution $\mathbf{x}(t) = \phi(t, x_0)$ of the initial value problem (2) is defined on $[0, t_0]$. Thus, as in the proof of Theorem 2 in Section 2.4, the solution $\mathbf{x}(t)$ can be extended to an interval $[0, t_0 + \varepsilon]$ for some $\varepsilon > 0$; i.e., $\phi(t, x_0)$ is defined on the closed interval $[t_0 - \varepsilon, t_0 + \varepsilon]$. It then follows from Theorem 4 in Section 2.4 that there exists a neighborhood of $\mathrm{x}_0, N_\delta(\mathrm{x}_0)$, such that

$\phi(t, \mathbf{y})$ is defined on $[t_0 - \varepsilon, t_0 + \varepsilon] \times N_\delta(\mathbf{x}_0)$; i.e., $(t_0 - \varepsilon, t_0 + \varepsilon) \times N_\delta(\mathbf{x}_0) \subset \Omega$. Therefore, Ω is open in $R \times E$. It follows from Theorem 4 in Section 2.4 that $\phi \in C^1(G)$ where $G = (t_0 - \varepsilon, t_0 + \varepsilon) \times N_\delta(\mathbf{x}_0)$. A similar proof holds for $t_0 \leq 0$, and since (t_0, x_0) is an arbitrary point in Ω, it follows that $\phi \in C^1(\Omega)$.

Remark. Theorem 1 can be generalized to show that if $\mathrm{f} \in C^r(E)$ with $r \geq 1$, then $\phi \in C^r(\Omega)$ and that if f is analytic in E, then ϕ is analytic in Ω.

Theorem 2. Let E be an open set of $\mathbf{R}^n$ and let $\mathbf{f} \in C^1(E)$. Then for all $x_0 \in E$, if $t \in I(x_0)$ and $s \in I\big(\phi_t(x_0)\big)$, it follows that $s + t \in I(x_0)$ and

$$\phi_{s+t}(x_0) = \phi_s\left(\phi_t(x_0)\right).$$

Proof. Suppose that $s > 0, t \in I(x_0)$ and $s \in I\big(\phi_t(x_0)\big)$. Let the maximal interval $I(\mathbf{x}_0) = (\alpha, \beta)$ and define the function $\mathbf{x} : (\alpha, s+t] \to E$ by

$$\mathbf{x}(r) = \begin{cases} \phi(r, x_0) & \text{if } \alpha < r \le t \\ \phi\left(r - t, \phi_t(x_0)\right) & \text{if } t \le r \le s + t. \end{cases}$$

Then $\mathbf{x}(r)$ is a solution of the initial value problem (2) on $(\alpha, s+t)$. Hence $s+t \in I(\mathbf{x}_0)$ and by uniqueness of solutions

$$\phi_{s+t}(x_0) = \mathbf{x}(s+t) = \phi\left(s, \phi_t(x_0)\right) = \phi_s\left(\phi_t(x_0)\right).$$

If $s = 0$ the statement of the theorem follows immediately. And if $s < 0$, then we define the function $\mathbf{x} : [s+t, \beta) \to E$ by

$$\mathbf{x}(t) = \begin{cases} \phi(r, \mathbf{x}_0) & \text{if } t \le r < \beta \\ \phi\left(r - t, \phi_t(x_0)\right) & \text{if } s + t \le r \le t. \end{cases}$$

Then $\mathbf{x}(r)$ is a solution of the initial value problem (2) on $[s+t, \beta)$ and the last statement of the theorem follows from the uniqueness of solutions as above.

Theorem 3. Under the hypotheses of Theorem 1, if $(t, x_0) \in \Omega$ then there exists a neighborhood U of x_0 such that $\{t\} \times U \subset \Omega$. It then follows that the set $V = \phi_t(U)$ is open in E and that

$$\phi_{-t}\left(\phi_t(\mathbf{x})\right) = \mathbf{x} \text{ for all } \mathbf{x} \in U$$

and

$$\phi_t\left(\phi_{-t}(\mathbf{y})\right) = \mathbf{y} \text{ for all } \mathbf{y} \in V.$$

Proof. If $(t, x_0) \in \Omega$ then if follows as in the proof of Theorem 1 that there exists a neighborhood of $x_0, U = N_\delta(x_0)$, such that $(t - \varepsilon, t + \varepsilon) \times U \subset \Omega$; thus, $\{t\} \times U \subset \Omega$. For $\mathbf{x} \in U$, let $\mathbf{y} = \phi_t(\mathbf{x})$ for all $t \in I(\mathbf{x})$. Then $-t \in I(\mathbf{y})$ since the function $\mathrm{h}(s) = \phi(s + t, \mathbf{y})$ is a solution of (1) on $[-t, 0]$ that satisfies $\mathbf{h}(-t) = \mathbf{y}$; i.e., ϕ_{-t} is defined on the set $V = \phi_t(U)$. It then follows from Theorem 2 that $\phi_{-t}\left(\phi_t(\mathbf{x})\right) = \phi_0(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in U$ and that $\phi_t\left(\phi_{-t}(\mathbf{y})\right) = \phi_0(\mathbf{y}) = \mathbf{y}$ for all $\mathbf{y} \in V$. It remains to prove that V is open. Let $V^* \supset V$ be the maximal subset of E on which ϕ_{-t} is defined. V^* is open because Ω is open and $\phi_{-t} : V^* \to E$ is continuous because by Theorem 1, ϕ is continuous. Therefore, the inverse image of the open set U under the continuous map ϕ_{-t}, i.e., $\phi_t(U)$, is open in E. Thus, V is open in E.

In Chapter 3 we show that the time along each trajectory of (1) can be rescaled, without affecting the phase portrait of (1), so that for all $\mathbf{x}_0 \in E$, the solution $\phi(t, x_0)$ of the initial value problem (2) is defined for all $t \in \mathbf{R}$; i.e., for all $x_0 \in E, I(x_0) = (-\infty, \infty)$. This rescaling avoids some of the complications found in stating the above theorems. Once this rescaling has been made, it follows that $\Omega = \mathbf{R} \times E, \phi \in C^1(\mathbf{R} \times E), \phi_t \in C^1(E)$ for all $t \in \mathbf{R}$, and properties (i)-(iii) for the flow of the nonlinear system (1) hold for all $t \in \mathbf{R}$ and $\mathbf{x} \in E$ just as for the linear flow e^{At}. In the remainder of this chapter, and in particular in Sections 2.7 and 2.8 of this chapter, it will be assumed that this rescaling has been made so that for all $x_0 \in E, \phi(t, x_0)$ is defined for all $t \in \mathbf{R}$; i.e., we shall assume throughout the remainder of this chapter that the flow of the nonlinear system (1) $\phi_t \in C^1(E)$ for all $t \in \mathbf{R}$.

Definition 2. Let E be an open subset of $\mathbf{R}^n$, let $f \in C^1(E)$, and let $\phi_t : E \to E$ be the flow of the differential equation (1) defined for all $t \in \mathbf{R}$. Then a set $S \subset E$ is called invariant with respect to the flow ϕ_t if $\phi_t(S) \subset S$ for all $t \in \mathbf{R}$ and S is called positively (or negatively) invariant with respect to the flow ϕ_t if $\phi_t(S) \subset S$ for all $t \geq 0$ (or $t \leq 0$). In Section 1.9 of Chapter 1 we showed that the stable, unstable and center subspaces of the linear system $\dot{\mathbf{x}} = A\mathbf{x}$ are invariant under the linear flow $\phi_t = e^{At}$. A similar result is established in Section 2.7 for the nonlinear flow ϕ_t of (1).

3.5 Linearization

A good place to start analyzing the nonlinear system

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

is to determine the equilibrium points of (1) and to describe the behavior of (1) near its equilibrium points. In the next two sections it is shown that the local behavior of the nonlinear system (1) near a hyperbolic equilibrium point x_0 is qualitatively determined by the behavior of the linear system

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

with the matrix $A = D\mathbf{f}(\mathbf{x}_0)$, near the origin. The linear function $A\mathbf{x} = \mathrm{D}\,\mathbf{f}(\mathbf{x}_0)\,\mathbf{x}$ is called the linear part of $\mathbf{f}$ at $\mathbf{x}_0$.

Definition 1. A point $x_0 \in \mathbf{R}^n$ is called an equilibrium point or critical point of (1) if $\mathbf{f}(\mathbf{x}_0) = \mathbf{0}$. An equilibrium point $\mathbf{x}_0$ is called a hyperbolic equilibrium point of (1) if none of the eigenvalues of the matrix $D\mathbf{f}(x_0)$ have zero real part. The linear system (2) with the matrix $A = Df(x_0)$ is called the linearization of

(1) at $\mathbf{x}_0$.

If $x_0 = 0$ is an equilibrium point of (1), then $\mathbf{f}(0) = 0$ and, by Taylor's Theorem,

$$\mathbf{f}(\mathbf{x}) = D\mathbf{f}(\mathbf{0})\mathbf{x} + \frac{1}{2}D^2\mathbf{f}(\mathbf{0})(\mathbf{x}, \mathbf{x}) + \cdots .$$

It follows that the linear function $D\mathbf{f}(\mathbf{0})\mathbf{x}$ is a good first approximation to the nonlinear function $\mathbf{f}(\mathbf{x})$ near $\mathbf{x} = \mathbf{0}$ and it is reasonable to expect that the behavior of the nonlinear system (1) near the point $\mathbf{x} = 0$ will be approximated by the behavior of its linearization at $\mathbf{x} = \mathbf{0}$. In Section 2.7 it is shown that this is indeed the case if the matrix $Df(0)$ has no zero or pure imaginary eigenvalues.

Note that if x_0 is an equilibrium point of (1) and $\phi_t : E \to \mathbf{R}^n$ is the flow of the differential equation (1), then $\phi_t(\mathbf{x}_0) = \mathbf{x}_0$ for all $t \in \mathbf{R}$. Thus, $\mathbf{x}_0$ is called a fixed point of the flow ϕ_t; it is also called a zero, a critical point, or a singular point of the vector field $\mathbf{f} : E \to \mathbf{R}^n$. We next give a rough classification of the equilibrium points of (1) according to the signs of the real parts of the eigenvalues of the matrix $Df(\mathrm{x}_0)$. A finer classification is given in Section 2.10 for planar vector fields.

Definition 2. An equilibrium point x_0 of (1) is called a sink if all of the eigenvalues of the matrix $D\mathbf{f}(\mathrm{x}_0)$ have negative real part; it is called a source if all of the eigenvalues of $D\mathbf{f}(\mathrm{x}_0)$ have positive real part; and it is called a saddle if it is a hyperbolic equilibrium point and $Df(\mathrm{x}_0)$ has at least one eigenvalue with a positive real part and at least one with a negative real part.

3.6 The Stable Manifold Theorem

The stable manifold theorem is one of the most important results in the local qualitative theory of ordinary differential equations. The theorem shows that near a hyperbolic equilibrium point x_0, the nonlinear system (1) has stable and unstable manifolds S and U tangent at x_0 to the stable and unstable subspaces E^s and E^u of the linearized system (2) where $A = D\mathbf{f}(\mathbf{x}_0)$. Furthermore, S and U are of the same dimensions as E^a and E^u, and if ϕ_t is the flow of the nonlinear system (1), then S and U are positively and negatively invariant under ϕ_t respectively and satisfy

$$\lim_{t\to\infty} \phi_t(\mathbf{c}) = \mathrm{x}_0 \text{ for all } \mathbf{c} \in S$$

and

$$\lim_{t\to-\infty} \phi_t(\mathbf{c}) = \mathbf{x}_0 \text{ for all } \mathbf{c} \in U.$$

We first illustrate these ideas with an example and then make them more precise by proving the stable manifold theorem. It is assumed that the equilibrium point $\mathbf{x}_0$ is located at the origin throughout the remainder of this section. If this is not the case, then the equilibrium point x_0 can be translated to the origin by the affine transformation of coordinates $x \to \mathbf{x} - \mathbf{x_0}$.

Example 1. Consider the nonlinear system

$$\begin{aligned}
\dot{x}_1 &= -x_1 \\
\dot{x}_2 &= -x_2 + x_1^2 \\
\dot{x}_3 &= x_3 + x_1^2.
\end{aligned}$$

The only equilibrium point of this system is at the origin. The matrix

$$A = D\mathbf{f}(0) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus, the stable and unstable subspaces E^s and E^u of (2) are the x_1, x_2 plane and the x_3-axis respectively. After solving the first differential equation, $\dot{x}_1 = -x_1$, the nonlinear system reduces to two uncoupled first-order linear differential equations which are easily solved. The solution is given by

$$\begin{aligned}
x_1(t) &= c_1 \mathrm{e}^{-t} \\
x_2(t) &= c_2 \mathrm{e}^{-t} + c_1^2 \left(\mathrm{e}^{-t} - \mathrm{e}^{-2t} \right) \\
x_3(t) &= c_3 \mathrm{e}^{t} + \frac{c_1^2}{3} \left(\mathrm{e}^{t} - \mathrm{e}^{-2t} \right)
\end{aligned}$$

where $\mathbf{c} = \mathbf{x}(0)$. Clearly, $\lim_{t\to\infty} \phi_t(\mathbf{c}) = \mathbf{0}$ iff $c_3 + c_1^2/3 = 0$. Thus,

$$S = \left\{ \mathbf{c} \in \mathbf{R}^3 \mid c_3 = -c_1^2/3 \right\}.$$

Similarly, $\lim_{t\to-\infty} \phi_t(c) = 0$ iff $c_1 = c_2 = 0$ and therefore

$$U = \left\{ c \in \mathbf{R}^3 \mid c_1 = c_2 = 0 \right\}$$

The stable and unstable manifolds for this system are shown in Figure 1. Note that the surface S is tangent to E^s, i.e., to the x_1, x_2 plane at the origin and that $U = E^u$.

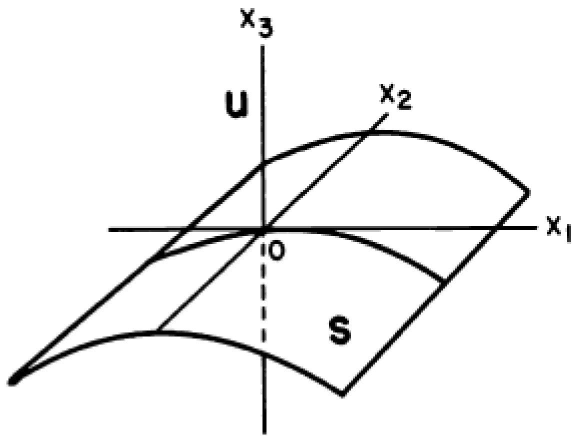

Figure 1

Before proving the stable manifold theorem, we first define the concept of a smooth surface or differentiable manifold.

Definition 1. Let X be a metric space and let A and B be subsets of X. A homeomorphism of A onto B is a continuous one-to-one map of A onto $B, h : A \to B$, such that $h^{-1} : B \to A$ is continuous. The sets A and B are called homeomorphic or topologically equivalent if there is a homeomorphism of A onto B. If we wish to emphasize that h maps A onto B.

Definition 2. An n-dimensional differentiable manifold, M (or a manifold of class C^k), is a connected metric space with an open covering $\{U_\alpha\}$, i.e., $M = U_\alpha U_\alpha$, such that

(1) for all α, U_α is homeomorphic to the open unit ball in $\mathbf{R}^n, B = \{\mathbf{x} \in \mathbf{R}^n \| \mathbf{x} |< 1\}$, i.e., for all α there exists a homeomorphism of U_α onto $B, \mathbf{h}_\alpha : U_\alpha \to B$, and

(2) if $U_\alpha \cap U_\beta \neq \emptyset$ and $\mathbf{h}_\alpha : U_\alpha \to B, \mathbf{h}_\beta : U_\beta \to B$ are homeomorphisms, then $\mathbf{h}_\alpha\left(U_\alpha \cap U_\beta\right)$ and $\mathbf{h}_\beta\left(U_\alpha \cap U_\beta\right)$ are subsets of $\mathbf{R}^n$ and the map

$$\mathbf{h} = \mathbf{h}_\alpha \circ \mathbf{h}_\beta^{-1} : \mathbf{h}_\beta\left(U_\alpha \cap U_\beta\right) \to \mathbf{h}_\alpha\left(U_\alpha \cap U_\beta\right)$$

is differentiable (or of class C^k) and for all $\mathbf{x} \in \mathbf{h}_\beta\left(U_\alpha \cap U_\beta\right)$, the Jacobian determinant $\det D\mathbf{h}(\mathbf{x}) \neq 0$. The manifold M is said to be analytic if the maps $h = h_\alpha \circ h_\beta^{-1}$ are analytic.

The cylindrical surface S in the above example is a two-dimensional differentiable manifold. The projection of the x_1, x_2 plane onto S maps the unit disks centered at the points (m, n) in the $x_1 x_2$ plane onto homeomorphic images of the unit disk $B = \left\{\mathbf{x} \in \mathbf{R}^2 \mid x_1^2 + x_2^2 < 1\right\}$. These sets $U_{mn} \subset S$ then form a countable open cover of S in this case.

The pair ($U_\alpha, \mathbf{h}_\alpha$) is called a chart for the manifold M and the set of all charts is called an atlas for M. The differentiable manifold M is called orientable if there is an atlas with $\det D\mathbf{h}_\alpha \circ \mathbf{h}_\beta^{-1}(\mathbf{x}) > 0$ for all α, β and $\mathbf{x} \in \mathbf{h}_\beta\left(U_\alpha \cap U_\beta\right)$.

Theorem (The Stable Manifold Theorem). Let E be an open subset of $\mathbf{R}^n$ containing the origin, let $\mathbf{f} \in C^1(E)$, and let ϕ_t be the flow of the nonlinear system (1). Suppose that $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ and that $D\mathbf{f}(\mathbf{0})$ has k eigenvalues with negative real part and $n-k$ eigenvalues with positive real part. Then there exists a k-dimensional differentiable manifold S tangent to the stable subspace E^s of the linear system (2) at 0 such that for all $t \geq 0, \phi_t(S) \subset S$ and for all $\mathrm{x}_0 \in S$.

$$\lim_{t\to\infty} \phi_t\left(\mathbf{x}_0\right) = \mathbf{0}$$

and there exists an $n-k$ dimensional differentiable manifold U tangent to the unstable subspace E^u of (2) at 0 such that for all $t \leq 0, \phi_t(U) \subset U$ and for all $\mathbf{x}_0 \in U$,

$$\lim_{t\to-\infty} \phi_t\left(\mathbf{x}_0\right) = \mathbf{0}.$$

Before proving this theorem, we remark that if f $\in C^1(E)$ and $f(0) = 0$, then the system (1) can be written as

$$\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{F}(\mathbf{x}) \tag{3}$$

where $A = D\mathbf{f}(\mathbf{0}), \mathbf{F}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) - A\mathbf{x}, \mathbf{F} \in C^1(E), \mathbf{F}(\mathbf{0}) = \mathbf{0}$ and $D\mathbf{F}(0) = 0$. This in turn implies that for all $\varepsilon > 0$ there is a $\delta > 0$ such that $|x| \leq \delta$ and $|\mathbf{y}| \leq \delta$ imply that

$$|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})| \leq \varepsilon|\mathbf{x} - \mathbf{y}|. \tag{4}$$

Furthermore, as in Section 1.8 of Chapter 1, there is an $n \times n$ invertible matrix C such that

$$B = C^{-1}AC = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}$$

where the eigenvalues $\lambda_1, \ldots, \lambda_k$ of the $k \times k$ matrix P have negative real part and the eigenvalues $\lambda_{k+1}, \ldots, \lambda_n$ of the $(n-k) \times (n-k)$ matrix Q have postive real part. We can choose $\alpha > 0$ sufficiently small that for $j = 1, \ldots, k$,

$$\operatorname{Re}\left(\lambda_j\right) < -\alpha < 0. \tag{5}$$

Letting $y = C^{-1}x$, the system (3) then has the form

$$\dot{\mathbf{y}} = B\mathbf{y} + \mathbf{G}(\mathbf{y}) \tag{6}$$

where $\mathbf{G}(\mathbf{y}) = C^{-1}\mathbf{F}(C\mathbf{y}) \in C^1(\tilde{E})$ where $\tilde{E} = C^{-1}(E)$ and $\mathbf{G}$ satisfies the Lipschitz-type condition (4) above.

It will be shown in the proof that there are $n-k$ differentiable functions $\psi_j(y_1,\ldots,y_k)$ such that the equations

$$y_j = \psi_j(y_1,\ldots,y_k), j = k+1,\ldots,n$$

define a k-dimensional differentiable manifold $\tilde{S}$ in $\mathbf{y}$-space. The differentiable manifold S in $\mathbf{x}$-space is then obtained from $\tilde{S}$ under the linear transformation of coordinates $\mathbf{x} = C\mathbf{y}$.

Proof. Consider the system (6). Let

$$U(t) = \begin{bmatrix} \mathrm{e}^{Pt} & 0 \\ 0 & 0 \end{bmatrix} \text{ and } V(t) = \begin{bmatrix} 0 & 0 \\ 0 & \mathrm{e}^{Qt} \end{bmatrix}.$$

Then $\dot{U} = BU, \dot{V} = BV$ and

$$e^{Bt} = U(t) + V(t)$$

It is not difficult to see that with $\alpha > 0$ chosen as in (5), we can choose $K > 0$ sufficiently large and $\sigma > 0$ sufficiently small that

$$\|U(t)\| \leq Ke^{-(\alpha+\sigma)t} \text{ for all } t \geq 0$$

and

$$\|V(t)\| \leq Ke^{\sigma t} \text{ for all } t \leq 0.$$

Next consider the integral equation

$$\mathbf{u}(t,\mathbf{a}) = U(t)\mathbf{a} + \int_0^t U(t-s)\mathbf{G}(\mathbf{u}(s,\mathbf{a}))ds - \int_t^\infty V(t-s)\mathbf{G}(\mathbf{u}(s,\mathbf{a}))ds. \tag{7}$$

If $\mathbf{u}(t,\mathbf{a})$ is a continuous solution of this integral equation, then it is a solution of the differential equation (6). We now solve this integral equation by the method of successive approximations. Let

$$\mathbf{u}^{(0)}(t,\mathbf{a}) = \mathbf{0}$$

and

$$\begin{aligned} \mathbf{u}^{(j+1)}(t,\mathbf{a}) =& U(t)\mathbf{a} + \int_0^t U(t-s)\mathbf{G}\left(\mathbf{u}^{(j)}(s,\mathbf{a})\right)ds \\ &- \int_t^\infty V(t-s)\mathbf{G}\left(\mathbf{u}^{(j)}(s,\mathbf{a})\right)ds. \end{aligned} \tag{8}$$

Assume that the induction hypothesis

$$\left|\mathbf{u}^{(j)}(t,\mathbf{a}) - \mathbf{u}^{(j-1)}(t,\mathbf{a})\right| \leq \frac{K|\mathbf{a}|e^{-\alpha t}}{2^{j-1}} \tag{9}$$

holds for $j = 1, 2, \ldots, m$ and $t \geq 0$. It clearly holds for $j = 1$ provided $t \geq 0$. Then using the Lipschitz-type condition (4) satisfied by the function $\mathbf{G}$ and the above estimates on $\|U(t)\|$ and $\|V(t)\|$, it follows from the induction hypothesis that for $t \geq 0$

$$\begin{aligned}
\left|\mathbf{u}^{(m+1)}(t,\mathbf{a}) - \mathbf{u}^{(m)}(t,\mathbf{a})\right| \leq & \int_0^t \|U(t-s)\|\varepsilon \left|\mathbf{u}^{(m)}(s,\mathbf{a}) - \mathbf{u}^{(m-1)}(s,\mathbf{a})\right| ds \\
& + \int_t^\infty \|V(t-s)\|\varepsilon \left|\mathbf{u}^{(m)}(s,\mathbf{a}) - \mathbf{u}^{(m-1)}(s,\mathbf{a})\right| ds \\
\leq & \varepsilon \int_0^t K e^{-(\alpha+\sigma)(t-s)} \frac{K|\mathbf{a}|e^{-\alpha s}}{2^{m-1}} ds \\
& + \varepsilon \int_0^\infty K e^{\sigma(t-s)} \frac{K|\mathbf{a}|e^{-\alpha s}}{2^{m-1}} ds \\
\leq & \frac{\varepsilon K^2|\mathbf{a}|e^{-\alpha t}}{\sigma 2^{m-1}} + \frac{\varepsilon K^2|\mathbf{a}|e^{-\alpha t}}{\sigma 2^{m-1}} \\
< & \left(\frac{1}{4} + \frac{1}{4}\right) \frac{K|\mathbf{a}|e^{-\alpha t}}{2^{m-1}} = \frac{K|\mathbf{a}|e^{-\alpha t}}{2^m}
\end{aligned} \tag{10}$$

provided $\varepsilon K/\sigma < 1/4$; i.e., provided we choose $\varepsilon < \dfrac{\sigma}{4K}$. In order that the condition (4) hold for the function $\mathbf{G}$, it suffices to choose $K|\mathbf{a}| < \delta/2$; i.e., we choose $|a| < \dfrac{\delta}{2K}$. It then follows by induction that (9) holds for all $j = 1, 2, 3, \ldots$ and $t \geq 0$. Thus, for $n > m > N$ and $t \geq 0$,

$$\begin{aligned}
\left|\mathbf{u}^{(\mathbf{n})}(t,\mathbf{a}) - \mathbf{u}^{(m)}(t,\mathbf{a})\right| \leq & \sum_{j=N}^{\infty} \left|\mathbf{u}^{(j+1)}(t,\mathbf{a}) - \mathbf{u}^{(j)}(t,\mathbf{a})\right| \\
\leq & K|\mathbf{a}| \sum_{j=N}^{\infty} \frac{1}{2^j} = \frac{K|\mathbf{a}|}{2^{N-1}}.
\end{aligned}$$

This last quantity approaches zero as $N \to \infty$ and therefore $\left\{\mathbf{u}^{(j)}(t,\mathbf{a})\right\}$ is a Cauchy sequence of continuous functions. According to the theorem in Section 2.2 ,

$$\lim_{j\to\infty} \mathbf{u}^{(j)}(t,\mathbf{a}) = \mathbf{u}(t,\mathbf{a})$$

uniformly for all $t \geq 0$ and $|\mathbf{a}| < \delta/2K$. Taking the limit of both sides of (8), it follows from the uniform convergence that the continuous function $\mathbf{u}(t, \mathrm{a})$ satisfies the integral equation (7) and hence the differential equation (6). It follows by induction and the fact that $\mathbf{G} \in C^1(\tilde{E})$ that $\mathbf{u}^{(j)}(t,a)$ is a differentiable function of a for $t \geq 0$ and $|\mathrm{a}| < \delta/2K$. Thus, it follows from the uniform convergence that $\mathbf{u}(t,\mathbf{a})$ is a differentiable function of a for $t \geq 0$ and $|a| < \delta/2K$. The estimate (10) implies that

$$|\mathbf{u}(t, \mathbf{a})| \leq 2K|\mathbf{a}|e^{-\alpha t} \tag{11}$$

for $t \geq 0$ and $|a| < \delta/2K$.

It is clear from the integral equation (7) that the last $n-k$ components of the vector a do not enter the computation and hence they may be taken as zero. Thus, the components $u_j(t, a)$ of the solution $u(t, a)$ satisfy the initial conditions

$$u_j(0, \mathbf{a}) = a_j \text{ for } j = 1, \ldots, k$$

and

$$u_j(0, \mathrm{a}) = -\left(\int_0^\infty V(-s)\mathbf{G}\left(\mathbf{u}\left(s, a_1, \ldots, a_k, 0\right)\right) ds\right)_j \text{ for } j = k+1, \ldots, n.$$

For $j = k+1, \ldots, n$ we define the functions

$$\psi_j\left(a_1, \ldots, a_k\right) = u_j\left(0, a_1, \ldots, a_k, 0, \ldots, 0\right). \tag{12}$$

Then the initial values $y_j = u_j\left(0, a_1, \ldots, a_k, 0, \ldots, 0\right)$ satisfy

$$y_j = \psi_j\left(y_1, \ldots, y_k\right) \text{ for } j = k+1, \ldots, n$$

according to the definition (12). These equations then define a differentiable manifold $\tilde{S}$ for $\sqrt{y_1^2 + \cdots + y_k^2} < \delta/2K$. Furthermore, if $\mathbf{y}(t)$ is a solution of the differential equation (6) with $\mathbf{y}(0) \in \tilde{S}$, i.e., with $\mathbf{y}(0) = \mathbf{u}(0, \mathbf{a})$, then

$$\mathbf{y}(t) = \mathbf{u}(t, \mathbf{a}).$$

It follows from the estimate (11) that if $y(t)$ is a solution of (6) with $\mathbf{y}(0) \in \tilde{S}$, then $\mathbf{y}(t) \to 0$ as $t \to \infty$. It can also be shown that if $\mathbf{y}(t)$ is a solution of (6) with $\mathbf{y}(0) \notin \tilde{S}$ then $\mathbf{y}(t) \nrightarrow 0$ as $t \to \infty$; cf. Coddington and Levinson. It therefore follows from Theorem 2 in Section 2.5 that if $\mathbf{y}(0) \in \tilde{S}$, then $\mathbf{y}(t) \in \tilde{S}$ for all $t \geq 0$. And it can be shown as in [C/L], p. 333 that

$$\frac{\partial \psi_j}{\partial y_i}(0) = 0$$

for $i = 1, \ldots, k$ and $j = k+1, \ldots, n$; i.e., the differentiable manifold $\tilde{S}$ is tangent to the stable subspace $E^s = \{\mathbf{y} \in \mathbf{R}^n \mid y_1 = \cdots y_k = 0\}$ of the linear system $\dot{\mathbf{y}} = B\mathbf{y}$ at 0 .

The existence of the unstable manifold $\tilde{U}$ of (6) is established in exactly the same way by considering the system (6) with $t \to -t$, i.e.,

$$\dot{\mathbf{y}} = -B\mathbf{y} - \mathbf{G}(\mathbf{y}).$$

The stable manifold for this system will then be the unstable manifold $\tilde{U}$ for (6). Note that it is also necessary to replace the vector y by the vector $(y_{k+1}, \ldots, y_n, y_1, \ldots, y_k)$ in order to determine the $n-k$ dimensional manifold U by the above process. This completes the proof of the Stable Manifold Theorem.

Remark 1. The first rigorous results concerning invariant manifolds were due to Hadamard in 1901, Liapunov in 1907 and Perron in 1928. They proved the existence of stable and unstable manifolds of systems of differential equations and of maps. The proof presented in this section is due to Liapunov and Perron. Several recent results generalizing the results of the Stable Manifold Theorem have been given by Hale, Hirsch, Pugh, Shub and Smale to mention a few. We note that if the function f $\in C^r(E)$ and $r \geq 1$, then the stable and unstable differentiable manifolds S and U of (1) are of class C^r. And if f is analytic in E then S and U are analytic manifolds.

Definition 3. Let ϕ_t be the flow of the nonlinear system (1). The global stable and unstable manifolds of (1) at 0 are defined by

$$W^s(0) = \bigcup_{t \leq 0} \phi_t(S)$$

and

$$W^u(0) = \bigcup_{t \geq 0} \phi_t(U)$$

respectively; $W^s(0)$ and $W^u(0)$ are also referred to as the global stable and unstable manifolds of the origin respectively. It can be shown that the global stable and unstable manifolds $W^s(0)$ and $W^u(0)$ are unique and

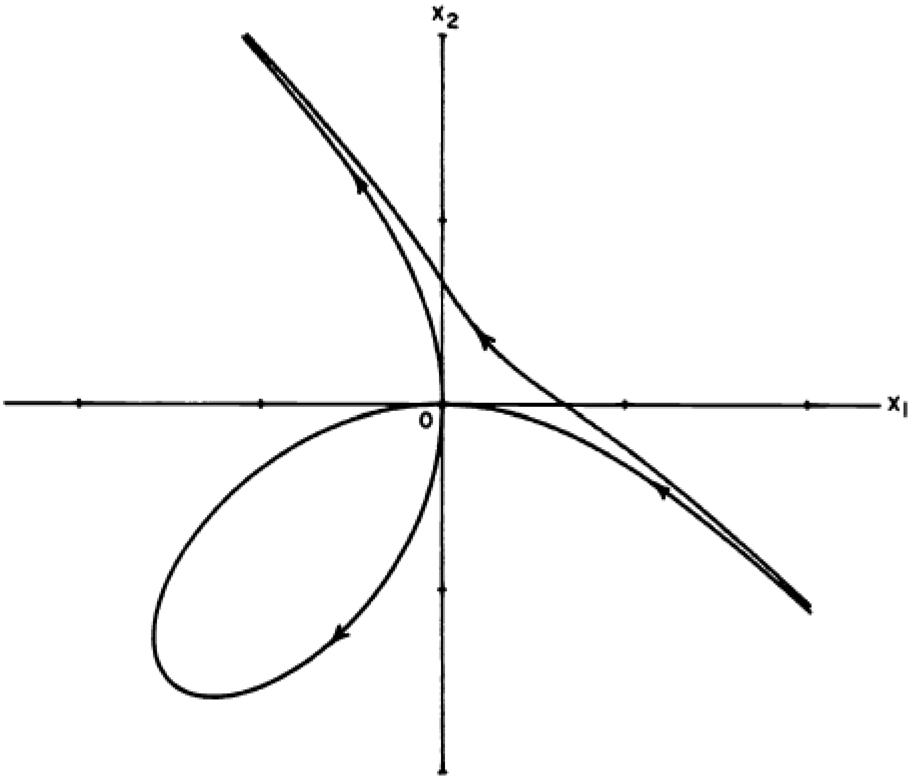

Figure 3

that they are invariant with respect to the flow ϕ_t; furthermore, for all $\mathbf{x} \in W^s(\mathbf{0})$, $\lim_{t\to\infty} \phi_t(\mathbf{x}) = \mathbf{0}$ and for all $\mathbf{x} \in W^u(0)$, $\lim_{t\to-\infty} \phi_t(\mathbf{x}) = \mathbf{0}$.

As in the proof of the Stable Manifold Theorem, it can be shown that in a small neighborhood, N, of a hyperbolic critical point at the origin, the local stable and unstable manifolds, S and U, of (1) at the origin are given by

$$S = \left\{\mathbf{x} \in N \mid \phi_t(\mathbf{x}) \to 0 \text{ as } t \to \infty \text{ and } \phi_t(\mathbf{x}) \in N \text{ for } t \geq 0\right\}$$

and

$$U = \left\{\mathbf{x} \in N \mid \phi_t(\mathbf{x}) \to 0 \text{ as } t \to -\infty \text{ and } \phi_t(\mathbf{x}) \in N \text{ for } t \leq 0\right\}$$

Figure 3 shows some numerically computed solution curves for the system in Example 2. The global stable and unstable manifolds for this example are shown in Figure 4. Note that $W^s(0)$ and $W^u(0)$ intersect in a "homoclinic loop" at the origin. $W^s(0)$ and $W^u(0)$ are more properly called "branched manifolds" in this example.

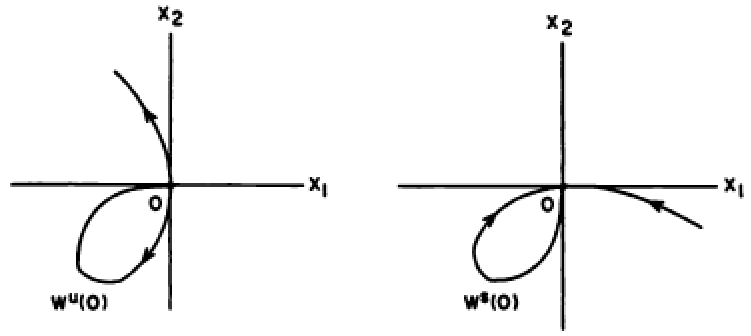

Figure 4

It follows from equation (11) in the proof of the stable manifold theorem that if $\mathbf{x}(t)$ is a solution of the differential equation (6) with $\mathbf{x}(0) \in S$, i.e., if $\mathbf{x}(t) = C\mathbf{y}(t)$ with $\mathbf{y}(0) = \mathbf{u}(0, \mathrm{a}) \in \tilde{S}$, then for any $\varepsilon > 0$ there exists a $\delta > 0$ such that if $|x(0)| < \delta$ then

$$|\mathbf{x}(t)| \leq \varepsilon e^{-\alpha t}$$

for all $t \geq 0$. Just as in the proof of the stable manifold theorem, α is any positive number that satisfies $\mathrm{Re}\left(\lambda_j\right) < -\alpha$ for $j = 1, \ldots, k$ where λ_j, $j = 1, \ldots, k$ are the eigenvalues of $D\mathrm{f}(0)$ with negative real part. This result shows that solutions starting in S, sufficiently near the origin, approach the origin exponentially fast as $t \to \infty$.

Corollary. Under the hypotheses of the Stable Manifold Theorem, if S and U are the stable and unstable manifolds of (1) at the origin and if $\mathrm{Re}\left(\lambda_j\right) < -\alpha < 0 < \beta < \mathrm{Re}\left(\lambda_m\right)$ for $j = 1, \ldots, k$ and $m = k + 1, \ldots, n$, then given $\varepsilon > 0$ there exists a $\delta > 0$ such that if $\mathrm{x}_0 \in N_\delta(0) \cap S$ then

$$\left|\phi_t\left(\mathrm{x}_0\right)\right| \leq \varepsilon \mathrm{e}^{-\alpha t}$$

for all $t \geq 0$ and if $\mathrm{x}_0 \in N_\delta(0) \cap U$ then

$$\left|\phi_t\left(x_0\right)\right| \leq \varepsilon e^{\beta t}$$

for all $t \leq 0$. We add one final result to this section which establishes the existence of an invariant center manifold $W^c(0)$ tangent to E^c at 0. The next theorem follows from the local center manifold theorem, Theorem 2 in Section 2.12, and the stable manifold theorem in this section.

Theorem (The Center Manifold Theorem). Let $\mathbf{f} \in C^r(E)$ where E is an open subset of $\mathbf{R}^n$ containing the origin and $r \geq 1$. Suppose that $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ and that $D\mathbf{f}(\mathbf{0})$ has k eigenvalues with negative real part, j eigenvalues with positive real part, and $m = n - k - j$ eigenvalues with zero real part. Then there exists an m-dimensional center manifold $W^c(0)$ of class C^r tangent to the center subspace E^c of (2) at 0 , there exists a k-dimensional stable manifold $W^s(\mathbf{0})$ of class C^r tangent to the stable subspace E^s of (2) at 0 and there exists a j-dimensional unstable manifold $W^u(0)$ of class C^r tangent to the unstable subspace E^u of (2) at 0 ; furthermore, $W^c(0)$, $W^s(\mathbf{0})$ and $W^u(\mathbf{0})$ are invariant under the flow ϕ_t of (1).

3.7 The Hartman-Grobman Theorem

The Hartman-Grobman Theorem is another very important result in the local qualitative theory of ordinary differential equations. The theorem shows that near a hyperbolic equilibrium point x_0, the nonlinear system (1) has the same qualitative structure as the linear system (2), with $A = D\mathbf{f}(\mathbf{x}_0)$. Throughout this section we shall assume that the equilibrium point x_0 has been translated to the origin.

Definition 1. Two autonomous systems of differential equations such as (1) and (2) are said to be topologically equivalent in a neighborhood of the origin or to have the same qualitative structure near the origin if there is a homeomorphism H mapping an open set U containing the origin onto an open set V containing the origin which maps trajectories of (1) in U onto trajectories of (2) in V and preserves their orientation by time in the sense that if a trajectory is directed from $\mathbf{x}_1$ to $\mathbf{x}_2$ in U, then its image is directed from $H(\mathbf{x}_1)$ to $H(\mathbf{x}_2)$ in V. If the homeomorphism H preserves the parameterization by time, then the systems (1) and (2) are said to be topologically conjugate in a neighborhood of the origin.

Before stating the Hartman-Grobman Theorem, we consider a simple example of two topologically conjugate linear systems.

Theorem (The Hartman-Grobman Theorem). Let E be an open subset of $\mathbf{R}^n$ containing the origin, let $\mathrm{f} \in C^1(E)$, and let ϕ_t be the flow of the nonlinear system (1). Suppose that $\mathrm{f}(\mathbf{0}) = \mathbf{0}$ and that the matrix $A = D\mathrm{f}(\mathbf{0})$ has no eigenvalue with zero real part. Then there exists a homeomorphism H of an open set U containing the origin onto an open set V containing the origin such that for each $\mathbf{x}_0 \in U$, there is an open interval $I_0 \subset \mathbf{R}$ containing zero such that for all $\mathrm{x}_0 \in U$ and $t \in I_0$

$$H \circ \phi_t(\mathbf{x}_0) = e^{At} H(\mathbf{x}_0);$$

i.e., H maps trajectories of (1) near the origin onto trajectories of (2) near the origin and preserves the parameterization by time.

Outline of the Proof. Consider the nonlinear system (1) with $\mathrm{f} \in C^1(E)$, $\mathrm{f}(\mathbf{0}) = \mathbf{0}$ and $A = D\mathrm{f}(\mathbf{0})$.

1. Suppose that the matrix A is written in the form

$$A = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}$$

where the eigenvalues of P have negative real part and the eigenvalues of Q have positive real part.

2. Let ϕ_t be the flow of the nonlinear system (1) and write the solution

$$\mathbf{x}(t, \mathbf{x}_0) = \phi_t(\mathbf{x}_0) = \begin{bmatrix} \mathbf{y}(t, \mathbf{y}_0, \mathbf{z}_0) \\ \mathbf{z}(t, \mathbf{y}_0, \mathbf{z}_0) \end{bmatrix}$$

where

$$\mathbf{x}_0 = \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{z}_0 \end{bmatrix} \in \mathbf{R^n},$$

$\mathbf{y}_0 \in E^s$, the stable subspace of A and $z_0 \in E^u$, the unstable subspace of A.

3. Define the functions

$$\tilde{\mathbf{Y}}(\mathbf{y}_0, z_0) = \mathbf{y}(1, \mathbf{y}_0, \mathbf{z}_0) - e^P \mathbf{y}_0$$

and

$$\tilde{\mathbf{Z}}(\mathbf{y}_0, \mathbf{z}_0) = \mathbf{z}(1, \mathbf{y}_0, \mathbf{z}_0) - e^{Q_{z_0}}.$$

Then $\tilde{\mathbf{Y}}(\mathbf{0}) = \tilde{\mathbf{Z}}(0) = D\tilde{\mathbf{Y}}(\mathbf{0}) = D\tilde{\mathbf{Z}}(\mathbf{0}) = 0$. And since $\mathbf{f} \in C^\mathbf{1}(E)$, $\tilde{\mathbf{Y}}(\mathbf{y}_0, \mathbf{z}_0)$ and $\tilde{\mathbf{Z}}(\mathbf{y}_0, \mathbf{z}_0)$ are continuously differentiable. Thus,

$$\left\|D\tilde{\mathbf{Y}}(\mathbf{y}_0, \mathbf{z}_0)\right\| \le a$$

and

$$\left\|D\tilde{\mathbf{Z}}(\mathbf{y}_0, \mathbf{z}_0)\right\| \le a$$

on the compact set $\left|y_0\right|^2 + |z_0|^2 \le s_0^2$. The constant a can be taken as small as we like by choosing s_0 sufficiently small. We let $\mathbf{Y}(\mathbf{y}_0, \mathbf{z}_0)$ and $\mathbf{Z}(\mathbf{y}_0, \mathbf{z}_0)$ be smooth functions which are equal to $\tilde{\mathbf{Y}}(\mathbf{y}_0, \mathbf{z}_0)$ and $\tilde{\mathbf{Z}}(\mathbf{y}_0, \mathbf{z}_0)$ for $\left|y_0\right|^2 + |z_0|^2 \le (s_0/2)^2$ and zero for $\left|y_0\right|^2 + |z_0|^2 \ge s_0^2$. Then by the mean value theorem

$$\left|\mathbf{Y}(\mathbf{y}_0, \mathbf{z}_0)\right| \le a\sqrt{\left|\mathbf{y}_0\right|^2 + |\mathbf{z}_0|^2} \le a\left(\left|\mathbf{y}_0\right| + |\mathbf{z}_0|\right)$$

and

$$\left|\mathbf{Z}(\mathbf{y}_0, \mathbf{z}_0)\right| \le a\sqrt{\left|\mathbf{y}_0\right|^2 + |z_0|^2} \le a\left(\left|\mathbf{y}_0\right| + |z_0|\right)$$

for all $(\mathbf{y}_0, \mathbf{z}_0) \in \mathbf{R^n}$. We next let $B = e^P$ and $C = e^Q$. Then assuming that we have carried out the normalization in Problem 7 in Section 1.8 of Chapter 1, cf. Hartman [H], p. 233, we have

$$b = \|B\| < 1 \text{ and } c = \left\|C^{-1}\right\| < 1.$$

4. For

$$\mathbf{x} = \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \end{bmatrix} \in \mathbf{R}^n$$

define the transformations

$$L(\mathbf{y},\mathbf{z}) = \begin{bmatrix} B\mathbf{y} \\ C\mathbf{z} \end{bmatrix}$$

and

$$T(\mathbf{y},\mathrm{z}) = \begin{bmatrix} B\mathbf{y} + \mathbf{Y}(\mathbf{y},\mathrm{z}) \\ C\mathrm{z} + \mathbf{Z}(\mathbf{y},\mathrm{z}) \end{bmatrix}$$

i.e., $L(\mathbf{x}) = e^{A}\mathbf{x}$ and locally $T(\mathbf{x}) = \phi_1(\mathbf{x})$.

Lemma. There exists a homeomorphism H of an open set U containing the origin onto an open set V containing the origin such that

$$H \circ T = L \circ H.$$

We establish this lemma using the method of successive approximations. For $\mathbf{x} \in \mathbf{R}^{\mathbf{n}}$, let

$$H(\mathbf{x}) = \begin{bmatrix} \Phi(\mathbf{y},\mathbf{z}) \\ \Psi(\mathbf{y},\mathbf{z}) \end{bmatrix}$$

Then $H \circ T = L \circ H$ is equivalent to the pair of equations

$$\begin{aligned} B\Phi(\mathbf{y},\mathbf{z}) &= \Phi(B\mathbf{y} + \mathbf{Y}(\mathbf{y},\mathbf{z}), C\mathbf{z} + \mathbf{Z}(\mathbf{y},\mathbf{z})) \\ C\Psi(\mathbf{y},\mathbf{z}) &= \Psi(B\mathbf{y} + \mathbf{Y}(\mathbf{y},\mathbf{z}), C\mathbf{z} + \mathbf{Z}(\mathbf{y},\mathbf{z})). \end{aligned} \tag{3}$$

First of all, define the successive approximations for the second equation by

$$\begin{aligned} \Psi_0(\mathbf{y},\mathbf{z}) &= \mathbf{z} \\ \Psi_{k+1}(\mathbf{y},\mathbf{z}) &= C^{-1}\Psi_k(B\mathbf{y} + \mathbf{Y}(\mathbf{y},\mathbf{z}), C\mathbf{z} + \mathbf{Z}(\mathbf{y},\mathbf{z})) \end{aligned} \tag{4}$$

It then follows by an easy induction argument that for $k = 0, 1, 2, \ldots$, the $\Psi_k(\mathbf{y},\mathbf{z})$ are continuous and satisfy $\Psi_k(\mathbf{y},\mathbf{z}) = \mathbf{z}$ for $|\mathbf{y}| + |\mathbf{z}| \geq 2s_0$. We next prove by induction that for $j = 1, 2, \ldots$

$$\left|\Psi_j(\mathbf{y},\mathbf{z}) - \Psi_{j-1}(\mathbf{y},\mathbf{z})\right| \leq Mr^j(|\mathbf{y}| + |\mathbf{z}|)^\delta$$

where $r = \mathrm{c}[2\max(a, b, \mathrm{c})]^\delta$ with $\delta \in (0,1)$ chosen sufficiently small so that $r < 1$ (which is possible since $c < 1$) and $M = ac\,(2s_0)^{1-\delta}/r$. First of all for $j = 1$

$$\begin{aligned} \left|\Psi_1(\mathbf{y},\mathbf{z}) - \Psi_0(\mathbf{y},\mathbf{z})\right| &= C^{-1}\Psi_0(B\mathbf{y} + \mathbf{Y}(\mathbf{y},\mathbf{z}), C\mathbf{z} + \mathbf{Z}(\mathbf{y},\mathbf{z})) - \mathbf{z}\ | \\ &= \left|C^{-1}(C\mathbf{z} + \mathbf{Z}(\mathbf{y},\mathbf{z})) - \mathbf{z}\right| \\ &= \left|C^{-1}\mathbf{Z}(\mathbf{y},\mathbf{z})\right| \leq \left\|C^{-1}\right\| |\mathbf{Z}(\mathbf{y},\mathbf{z})| \\ &\leq ca(|\mathbf{y}| + |\mathbf{z}|) \leq Mr(|\mathbf{y}| + |\mathbf{z}|)^\delta \end{aligned}$$

since $\mathbf{Z}(\mathbf{y},\mathbf{z}) = 0$ for $|\mathbf{y}| + |\mathbf{z}| \geq 2s_0$. And then assuming that the induction hypothesis holds for $j = 1, \ldots, k$ we have

$$\begin{aligned}
\left|\Psi_{k+1}(\mathbf{y},\mathbf{z}) - \Psi_k(\mathbf{y},\mathbf{z})\right| &= |\, C^{-1}\Psi_k(B\mathbf{y} + \mathbf{Y}(\mathbf{y},\mathbf{z}), C\mathbf{z} + \mathbf{Z}(\mathbf{y},\mathbf{z})) \\
&\quad - C^{-1}\Psi_{k-1}(B\mathbf{y} + \mathbf{Y}(\mathbf{y},\mathbf{z}), C\mathbf{z} + \mathbf{Z}(\mathbf{y},\mathbf{z}))\,| \\
&\leq \left\|C^{-1}\right\| \,|\, \Psi_k('') - \Psi_{k-1}('')\,| \\
&\leq \mathrm{c}Mr^k \left[|B\mathbf{y} + \mathbf{Y}(\mathbf{y},\mathbf{z})| + |C\mathbf{z} + \mathbf{Z}(\mathbf{y},\mathbf{z})\right]^\delta \\
&\leq \mathrm{c}Mr^k|b|\mathbf{y}| + 2a(|\mathbf{y}| + |\mathrm{z}|) + \mathrm{c}|\mathbf{z}\,|\Big]^\delta \\
&\leq \mathrm{c}Mr^k[2\max(a,b,\mathrm{c})]^\delta(|\mathbf{y}| + |\mathrm{z}|)^\delta \\
&= Mr^{k+1}(|\mathbf{y}| + |\mathrm{z}|)^\delta
\end{aligned}$$

Thus, just as in the proof of the fundamental theorem in Section 2.2 and the stable manifold theorem in Section 2.7, $\Psi_k(y,z)$ is a Cauchy sequence of continuous functions which converges uniformly as $k \to \infty$ to a continuous function $\Psi(y,z)$. Also, $\Psi(y,z) = z$ for $|y| + |z| \geq 2s_0$. Taking limits in (4) shows that $\Psi(\mathbf{y},\mathbf{z})$ is a solution of the second equation in (3).

The first equation in (3) can be written as

$$B^{-1}\Phi(y,z) = \Phi\left(B^{-1}y + Y_1(y,z), C^{-1}z + Z_1(y,z)\right)$$

where the functions Y_1 and Z_1 are defined by the inverse of T (which exists if the constant a is sufficiently small, i.e., if s_0 is sufficiently small) as follows:

$$T^{-1}(y,z) = \begin{bmatrix} B^{-1}y + Y_1(y,z) \\ C^{-1}z + Z_1(\mathbf{y},z) \end{bmatrix}.$$

Then equation (6) can be solved for $\Phi(y,z)$ by the method of successive approximations exactly as above with $\Phi_0(\mathbf{y},\mathbf{z}) = \mathbf{y}$ since $b = \|B\| < 1$. We therefore obtain the continuous map

$$H(\mathbf{y},\mathrm{z}) = \begin{bmatrix} \Phi(\mathbf{y},\mathrm{z}) \\ \Psi(\mathbf{y},\mathrm{z}) \end{bmatrix}$$

And it follows as on pp. 248-249 in Hartman [H] that H is a homeomorphism of $\mathbf{R}^\mathbf{n}$ onto $\mathbf{R}^n$.

5. We now let H_0 be the homeomorphism defined above and let L^t and T^t be the one-parameter families of transformations defined by

$$L^t\,(\mathrm{x}_0) = \mathrm{e}^{At}\mathrm{x}_0 \quad \text{and} \quad T^t\,(\mathrm{x}_0) = \phi_t\,(\mathrm{x}_0)\,.$$

Define

$$H = \int_0^1 L^{-s} H_0 T^s ds.$$

It then follows using the above lemma that there exists a neighborhood of the origin for which

$$\begin{aligned} L^t H &= \int_0^1 L^{t-s} H_0 T^{s-t} ds T^t \\ &= \int_{-t}^{1-t} L^{-s} H_0 T^s ds T^t \\ &= \left[\int_{-t}^0 L^{-s} H_0 T^s ds + \int_0^{1-t} L^{-s} H_0 T^s ds \right] T^t \\ &= \int_0^1 L^{-s} H_0 T^s ds T^t = H T^t \end{aligned}$$

since by the above lemma $H_0 = L^{-1} H_0 T$ which implies that

$$\begin{aligned} \int_{-t}^0 L^{-s} H_0 T^s ds &= \int_{-t}^0 L^{-s-1} H_0 T^{s+1} ds \\ &= \int_{1-t}^1 L^{-s} H_0 T^s ds \end{aligned}$$

Thus, $H \circ T^t = L^t H$ or equivalently

$$H \circ \phi_t (\mathrm{x}_0) = \mathrm{e}^{At} H (\mathrm{x}_0)$$

and it can be shown as on pp. 250-251 of Hartman [H] that H is a homeomorphism on $\mathbf{R}^n$. This completes the outline of the proof of the HartmanGrobman Theorem.

Theorem (Hartman). Let E be an open subset of $\mathbf{R^n}$ containing the point $\mathbf{x}_0$, let $\mathbf{f} \in C^2(E)$, and let ϕ_t be the flow of the nonlinear system (1). Suppose that $\mathbf{f}(\mathbf{x}_0) = 0$ and that all of the eigenvalues $\lambda_1, \ldots, \lambda_n$ of the matrix $A = D\mathbf{f}(\mathbf{x}_0)$ have negative (or positive) real part. Then there exists a C^1-diffeomorphism H of a neighborhood U of $\mathbf{x}_0$ onto an open set V containing the origin such that for each $\mathbf{x} \in U$ there is an open interval $I(\mathbf{x}) \subset \mathbf{R}$ containing zero such that for all $\mathbf{x} \in U$ and $t \in I(\mathbf{x})$

$$H \circ \phi_t(\mathbf{x}) = e^{At} H(\mathbf{x}).$$

3.8 Stability and Liapunov Functions

In this section we discuss the stability of the equilibrium points of the nonlinear system (1). The stability of any hyperbolic equilibrium point $\mathbf{x}_0$ of (1) is determined by the signs of the real parts of the eigenvalues λ_j of the matrix $Df(\mathrm{x}_0)$. A hyperbolic equilibrium point $\mathbf{x}_0$ is asymptotically stable iff $\mathrm{Re}\left(\lambda_j\right) < 0$ for

$j = 1, \ldots, n$; i.e., iff x_0 is a sink. And a hyperbolic equilibrium point x_0 is unstable iff it is either a source or a saddle. The stability of nonhyperbolic equilibrium points is typically more difficult to determine. A method, due to Liapunov, that is very useful for deciding the stability of nonhyperbolic equilibrium points is presented in this section.

Definition 1. Let ϕ_t denote the flow of the differential equation (1) defined for all $t \in \mathbf{R}$. An equilibrium point $\mathbf{x}_0$ of (1) is stable if for all $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $\mathbf{x} \in N_6(\mathbf{x}_0)$ and $t \geq 0$ we have

$$\phi_t(\mathrm{x}) \in N_c(\mathrm{x}_0) .$$

The equilibrium point $\mathbf{x}_0$ is unstable if it is not stable. And $\mathbf{x}_0$ is asymptotically stable if it is stable and if there exists a $\delta > 0$ such that for all $\mathbf{x} \in N_6(\mathbf{x}_0)$ we have

$$\lim_{t\to\infty} \phi_t(\mathbf{x}) = \mathrm{x}_0.$$

Note that the above limit being satisfied for all $\mathbf{x}$ in some neighborhood of $\mathbf{x}_0$ does not imply that $\mathbf{x}_0$ is stable.

It can be seen from the phase portraits in Section 1.5 of Chapter 1 that a stable node or focus of a linear system in $\mathbf{R}^2$ is an asymptotically stable equilibrium point; an unstable node or focus or a saddle of a linear system in $\mathbf{R}^2$ is an unstable equilibrium point; and a center of a linear system in $\mathbf{R}^2$ is a stable equilibrium point which is not asymptotically stable.

It follows from the Stable Manifold Theorem and the Hartman-Grobman Theorem that any sink of (1) is asymptotically stable and any source or saddle of (1) is unstable. Hence, any hyperbolic equilibrium point of (1) is either asymptotically stable or unstable. The corollary in Section 2.7 provides even more information concerning the local behavior of solutions near a sink:

Theorem 1. If x_0 is a sink of the nonlinear system (1) and $\mathrm{Re}\left(\lambda_j\right) < -\alpha < 0$ for all of the eigenvalues λ_j of the matrix $\mathrm{Df}(\mathrm{x}_0)$, then given $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $\mathbf{x} \in N_\delta(\mathbf{x}_0)$, the flow $\phi_t(\mathbf{x})$ of (1) satisfies

$$\left|\phi_t(\mathbf{x}) - \mathbf{x}_0\right| \leq \varepsilon e^{-\alpha t}$$

for all $t \geq 0$.

Since hyperbolic equilibrium points are either asymptotically stable or unstable, the only time that an equilibrium point x_0 of (1) can be stable but not asymptotically stable is when $D\mathbf{f}(\mathbf{x}_0)$ has a zero eigenvalue or a pair of complex-conjugate, pure-imaginary eigenvalues $\lambda = \pm ib$. It follows from the next theorem, proved in [H/S], that all other eigenvalues λ, of $\mathrm{Df}(\mathrm{x}_0)$ must satisfy $\mathrm{Re}\left(\lambda_j\right) \leq 0$ if x_0 is stable.

Theorem 2. If x_0 is a stable equilibrium point of (1), no eigenvalue of $\mathrm{Df}(\mathrm{x}_0)$ has positive real part.

We see that stable equilibrium points which are not asymptotically stable can only occur at nonhyperbolic equilibrium points. But the question as to whether a nonhyperbolic equilibrium point is stable, asymptotically stable or unstable is a delicate question.

The following method, due to Liapunov (in his 1892 doctoral thesis), is very useful in answering this question.

Definition 2. If $\mathbf{f} \in C^1(E), V \in C^1(E)$ and ϕ_t is the flow of the differential equation (1), then for $\mathbf{x} \in E$ the derivative of the function $V(\mathbf{x})$ along the solution $\phi_t(\mathbf{x})$

$$\dot{V}(\mathbf{x}) = \frac{d}{dt} V\left(\phi_t(\mathbf{x})\right)\Big|_{t=0} = DV(\mathbf{x})\mathbf{f}(\mathbf{x}).$$

The last equality follows from the chain rule. If $\dot{V}(\mathbf{x})$ is negative in E then $V(\mathbf{x})$ decreases along the solution $\phi_t\,(\mathrm{x}_0)$ through $\mathrm{x}_0 \in E$ at $t = 0$. Furthermore, in $\mathbf{R}^2$, if $\dot{V}(\mathbf{x}) \leq 0$ with equality only at $\mathbf{x} = 0$, then for small positive C, the family of curves $V(\mathbf{x}) = C$ constitutes a family of closed curves enclosing the origin and the trajectories of (1) cross these curves from their exterior to their interior with increasing t; i.e., the origin of (1) is asymptotically stable. A function $V : \mathbf{R^n} \to \mathbf{R}$ satisfying the hypotheses of the next theorem is called a Liapunov function.

Theorem 3. Let E be an open subset of $\mathbf{R}^n$ containing $\mathbf{x}_0$. Suppose that $\mathrm{f} \in C^1(E)$ and that $\mathrm{f}\,(\mathrm{x}_0) = \mathbf{0}$. Suppose further that there exists a real valued function $V \in C^1(E)$ satisfying $V\,(\mathrm{x}_0) = 0$ and $V(\mathbf{x}) > 0$ if $\mathbf{x} \neq \mathrm{x}_0$.
Then (a) if $\dot{V}(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in E, \mathbf{x}_0$ is stable;
(b) if $\dot{V}(\mathbf{x}) < 0$ for all $\mathbf{x} \in E \sim \{\mathbf{x}_0\}, \mathbf{x}_0$ is asymptotically stable;
(c) if $\dot{V}(\mathbf{x}) > 0$ for all $\mathbf{x} \in E \sim \{\mathbf{x}_0\}, \mathbf{x}_0$ is unstable.

Proof. Without loss of generality, we shall assume that the equilibrium point $\mathbf{x}_0 = \mathbf{0}$.

(a) Choose $\varepsilon > 0$ sufficiently small that $\overline{N_\varepsilon(0)} \subset E$ and let m_ε be the minimum of the continuous function $V(\mathbf{x})$ on the compact set

$$S_\varepsilon = \{\mathbf{x} \in \mathbf{R^n} | \quad |\mathbf{x}| = \varepsilon\}.$$

Then since $V(\mathbf{x}) > 0$ for $\mathbf{x} \neq 0$, it follows that $m_\varepsilon > 0$. Since $V(\mathbf{x})$ is continuous and $V(0) = 0$, it follows that there exists a $\delta > 0$ such that $|\mathbf{x}| < \delta$ implies that $V(\mathbf{x}) < m_c$. Since $\dot{V}(\mathbf{x}) \leq 0$ for $\mathbf{x} \in E$, it follows that $V(\mathbf{x})$ is decreasing along trajectories of (1). Thus, if ϕ_t is the flow of the differential equation (1), it follows that for all $\mathrm{x}_0 \in N_\delta(0)$ and $t \geq 0$ we have

$$V\left(\phi_t\,(\mathrm{x}_0)\right) \leq V\,(\mathrm{x}_0) < m_\varepsilon.$$

Now suppose that for $|\mathbf{x}_0| < \delta$ there is a $t_1 > 0$ such that $\left|\phi_{t_1}(\mathbf{x}_0)\right| = \varepsilon$; i.e., such that $\phi_{t_1}(\mathrm{x}_0) \in S_\varepsilon$. Then since m_ε is the minimum of $V(\mathbf{x})$ on S_ε, this would imply that

$$V\left(\phi_{t_1}(\mathrm{x}_0)\right) \geq m_\varepsilon$$

which contradicts the above inequality. Thus for $|x_0| < \delta$ and $t \geq 0$ it follows that $\left|\phi_t(\mathrm{x}_0)\right| < \varepsilon$; i.e., 0 is a stable equilibrium point.

(b) Suppose that $\dot{V}(\mathbf{x}) < 0$ for all $\mathbf{x} \in E$. Then $V(\mathbf{x})$ is strictly decreasing along trajectories of (1). Let ϕ_t be the flow of (1) and let $\mathrm{x}_0 \in N_\delta(0)$, the neighborhood defined in part (a). Then, by part (a), if $|x_0| < \delta$, $\phi_t(\mathrm{x}_0) \subset N_\varepsilon(0)$ for all $t \geq 0$. Let $\{t_k\}$ be any sequence with $t_k \to \infty$. Then since $\overline{N_\varepsilon(0)}$ is compact, there is a subsequence of $\left\{\phi_{t_k}(\mathrm{x}_0)\right\}$ that converges to a point in $\overline{N_\varepsilon(\mathbf{0})}$. But for any subsequence $\{t_n\}$ of $\{t_k\}$ such that $\left\{\phi_{t_n}(\mathbf{x}_0)\right\}$ converges, we show below that the limit is zero. It then follows that $\phi_{t_k}(\mathrm{x}_0) \to 0$ for any sequence $t_k \to \infty$ and therefore that $\phi_t(\mathrm{x}_0) \to 0$ as $t \to \infty$; i.e., that 0 is asymptotically stable. It remains to show that if $\phi_{t_n}(x_0) \to y_0$, then $y_0 = 0$. Since $V(x)$ is strictly decreasing along trajectories of (1) and since $V\left(\phi_{t_n}(\mathbf{x}_0)\right) \to V(\mathbf{y}_0)$ by the continuity of V, it follows that

$$V\left(\phi_t(\mathbf{x}_0)\right) > V(\mathbf{y}_0)$$

for all $t > 0$. But if $\mathbf{y}_0 \neq 0$, then for $s > 0$ we have $V\left(\phi_s(\mathbf{y}_0)\right) < V(\mathbf{y}_0)$ and, by continuity, it follows that for all y sufficiently close to y_0 we have $V\left(\phi_s(\mathbf{y})\right) < V(\mathbf{y}_0)$ for $s > 0$. But then for $\mathbf{y} = \phi_{t_n}(\mathrm{x}_0)$ and n sufficiently large, we have

$$V\left(\phi_{s+t_n}(\mathbf{x}_0)\right) < V(\mathbf{y}_0)$$

which contradicts the above inequality. Therefore $\mathbf{y}_0 = \mathbf{0}$ and it follows that 0 is asymptotically stable.

(c) Let M be the maximum of the continuous function $V(\mathbf{x})$ on the compact set $\overline{N_\epsilon(\mathbf{0})}$. Since $\dot{V}(\mathbf{x}) > 0$, $V(\mathbf{x})$ is strictly increasing along trajectories of (1). Thus, if ϕ_t is the flow of (1), then for any $\delta > 0$ and $\mathbf{x}_0 \in N_\delta(\mathbf{0}) \sim \{0\}$ we have

$$V\left(\phi_t(\mathbf{x}_0)\right) > V(\mathrm{x}_0) > 0$$

for all $t > 0$. And since $\dot{V}(\mathbf{x})$ is positive definite, this last statement implies that

$$\inf_{i \geq 0} \dot{V}\left(\phi_t(\mathbf{x}_0)\right) = m > 0$$

Thus,

$$V\left(\phi_t(\mathbf{x}_0)\right) - V(\mathbf{x}_0) \geq mt$$

for all $t \geq 0$. Therefore,

$$V\left(\phi_t\left(\mathbf{x}_0\right)\right) > mt > M$$

for t sufficiently large; i.e., $\phi_t\left(\mathbf{x}_0\right)$ lies outside the closed set $\overline{N_\varepsilon(\mathbf{0})}$. Hence, $\mathbf{0}$ is unstable.

Remark. If $\dot{V}(\mathbf{x}) = 0$ for all $\mathbf{x} \in E$ then the trajectories of (1) lie on the surfaces in $\mathbf{R}^n$ (or curves in $\mathbf{R}^2$) defined by

$$V(\mathbf{x}) = \mathbf{c}.$$

3.9 Saddles, Nodes, Foci and Centers

In Section 1.5 of Chapter 1, a linear system (1), where $\mathbf{x} \in \mathbf{R}^2$ was said to have a saddle, node, focus or center at the origin if its phase portrait was linearly equivalent to one of the phase portraits in Figures 1-4 in Section 1.5 of Chapter 1 respectively; i.e., if there exists a nonsingular linear transformation which reduces the matrix A to one of the canonical matrices B in Cases I-IV of Section 1.5 in Chapter 1 respectively. For example, the linear system (1) of the example in Section 2.8 of this chapter has a saddle at the origin.

In Section 2.6, a nonlinear system (2) was said to have a saddle, a sink or a source at a hyperbolic equilibrium point x_0 if the linear part of f at x_0 had eigenvalues with both positive and negative real parts, only had eigenvalues with negative real parts, or only had eigenvalues with positive real parts, respectively.

In this section, we define the concept of a topological saddle for the nonlinear system (2) with $x \in R^2$ and show that if x_0 is a hyperbolic equilibrium point of (2) then it is a topological saddle if and only if it is a saddle of (2); i.e., a hyperbolic equilibrium point x_0 is a topological saddle for (2) if and only if the origin is a saddle for (1) with $A = D\mathbf{f}\left(\mathbf{x}_0\right)$. We discuss topological saddles for nonhyperbolic equilibrium points of (2) with $\mathbf{x} \in \mathbf{R}^2$ in the next section. We also refine the classification of sinks of the nonlinear system (2) into stable nodes and foci and show that, under slightly stronger hypotheses on the function f, i.e., stronger than $\mathrm{f} \in C^1(E)$, a hyperbolic critical point x_0 is a stable node or focus for the nonlinear system (2) if and only if it is respectively a stable node or focus for the linear system (1) with $A = D\mathrm{f}(\mathrm{x}_0)$. Similarly, a source of (2) is either an unstable node or focus of (2) as defined below. Finally, we define centers and center-foci for the nonlinear system (2) and show that, under the addition of nonlinear terms, a center of the linear system (1) may become either a center, a center-focus, or a stable or unstable focus of (2).

Before defining these various types of equilibrium points for planar systems (2), it is convenient to introduce polar coordinates (r, θ) and to rewrite the system (2) in polar coordinates. In this section we let $\mathbf{x} = (x, y)^T$, $f_1(\mathbf{x}) = P(x, y)$ and $f_2(\mathbf{x}) = Q(x, y)$. The nonlinear system (2) can then be written as

$$\begin{aligned}\dot{x} &= P(x, y)\\ \dot{y} &= Q(x, y).\end{aligned} \tag{3}$$

If we let $r^2 = x^2 + y^2$ and $\theta = \tan^{-1}(y/x)$, then we have

$$r\dot{r} = x\dot{x} + y\dot{y}$$

and

$$r^2\dot{\theta} = x\dot{y} - y\dot{x}.$$

It follows that for $r > 0$, the nonlinear system (3) can be written in terms of polar coordinates as

$$\begin{aligned}\dot{r} &= P(r\cos\theta, r\sin\theta)\cos\theta + Q(r\cos\theta, r\sin\theta)\sin\theta\\ r\dot{\theta} &= Q(r\cos\theta, r\sin\theta)\cos\theta - P(r\cos\theta, r\sin\theta)\sin\theta\end{aligned} \tag{4}$$

or as

$$\frac{dr}{d\theta} = F(r, \theta) \equiv \frac{r[P(r\cos\theta, r\sin\theta)\cos\theta + Q(r\cos\theta, r\sin\theta)\sin\theta]}{Q(r\cos\theta, r\sin\theta)\cos\theta - P(r\cos\theta, r\sin\theta)\sin\theta}. \tag{5}$$

Writing the system of differential equations (3) in polar coordinates will often reveal the nature of the equilibrium point or critical point at the origin.

Definition 1. The origin is called a center for the nonlinear system (2) if there exists a $\delta > 0$ such that every solution curve of (2) in the deleted neighborhood $N_\delta(0) \sim \{0\}$ is a closed curve with 0 in its interior.

Definition 2. The origin is called a center-focus for (2) if there exists a sequence of closed solution curves Γ_n with Γ_{n+1} in the interior of Γ_n such that $\Gamma_n \to 0$ as $n \to \infty$ and such that every trajectory between Γ_n and Γ_{n+1} spirals toward Γ_n or Γ_{n+1} as $t \to \pm\infty$.

Definition 3. The origin is called a stable focus for (2) if there exists a $\delta > 0$ such that for $0 < r_0 < \delta$ and $\theta_0 \in \mathbf{R}$, $r(t, r_0, \theta_0) \to 0$ and $|\theta(t, r_0, \theta_0)| \to \infty$ as $t \to \infty$. It is called an unstable focus if $r(t, r_0, \theta_0) \to 0$ and $|\theta(t, r_0, \theta_0)| \to \infty$ as $t \to -\infty$. Any trajectory of (2) which satisfies $r(t) \to 0$ and $|\theta(t)| \to \infty$ as $t \to \pm\infty$ is said to spiral toward the origin as $t \to \pm\infty$.

Definition 4. The origin is called a stable node for (2) if there exists a $\delta > 0$ such that for $0 < r_0 < \delta$ and $\theta_0 \in \mathbf{R}$, $r(t, r_0, \theta_0) \to 0$ as $t \to \infty$ and $\lim_{t\to\infty} \theta(t, r_0, \theta_0)$ exists; i.e., each trajectory in a deleted neighborhood of the origin approaches the origin along a well-defined tangent line as $t \to \infty$. The origin is called an unstable node if $r(t, r_0, \theta_0) \to 0$ as $t \to -\infty$ and $\lim_{t\to-\infty} \theta(t, r_0, \theta_0)$ exists for all $r_0 \in (0, \delta)$ and $\theta_0 \in \mathbf{R}$. The origin is called a proper node for (2) if it is a node and if every ray through the origin is tangent to some trajectory of (2).

Definition 5. The origin is a (topological) saddle for (2) if there exist two trajectories Γ_1 and Γ_2 which approach 0 as $t \to \infty$ and two trajectories Γ_3 and Γ_4 which approach 0 as $t \to -\infty$ and if there exists a $\delta > 0$ such that all other trajectories which start in the deleted neighborhood of the origin $N_6(\mathbf{0}) \sim \{0\}$ leave $N_6(\mathbf{0})$ as $t \to \pm\infty$. The special trajectories $\Gamma_1, \ldots, \Gamma_4$ are called separatrices.

For a (topological) saddle, the stable manifold at the origin $S = \Gamma_1 \cup \Gamma_2 \cup \{0\}$ and the unstable manifold at the origin $U = \Gamma_3 \cup \Gamma_4 \cup \{0\}$. If the trajectory Γ_i approaches the origin along a ray making an angle θ_i with the x-axis where $\theta_i \in (-\pi, \pi]$ for $i = 1, \ldots, 4$, then $\theta_2 = \theta_1 \pm \pi$ and $\theta_4 = \theta_3 \pm \pi$. This follows by considering the possible directions in which a trajectory of (2), written in polar form (4), can approach the origin; cf. equation (6) below. The following theorems, proved in [A-I], are useful in this regard. The first theorem is due to Bendixson [B].

Theorem 1 (Bendixson). Let E be an open subset of $\mathbf{R}^2$ containing the origin and let f $\in C^1(E)$. If the origin is an isolated critical point of (2), then either every neighborhood of the origin contains a closed solution curve with 0 in its interior or there exists a trajectory approaching 0 as $t \to \pm\infty$.

Theorem 2. Suppose that $P(x, y)$ and $Q(x, y)$ in (3) are analytic functions of x and y in some open subset E of $\mathbf{R}^2$ containing the origin and suppose that the Taylor expansions of P and Q about $(0, 0)$ begin with mth-degree terms $P_m(x, y)$ and $Q_m(x, y)$ with $m \geq 1$. Then any trajectory of (3) which approaches the origin as $t \to \infty$ either spirals toward the origin as $t \to \infty$ or it tends toward the origin in a definite direction $\theta = \theta_0$ as $t \to \infty$. If $xQ_m(x, y) - yP_m(x, y)$ is not identically zero, then all directions of approach, θ_0, satisfy the equation

$$\cos\theta_0 Q_m(\cos\theta_0, \sin\theta_0) - \sin\theta_0 P_m(\cos\theta_0, \sin\theta_0) = 0.$$

Furthermore, if one trajectory of (3) spirals toward the origin as $t \to \infty$ then all trajectories of (3) in a deleted neighborhood of the origin spiral toward 0 as $t \to \infty$.

It follows from this theorem that if P and Q begin with first-degree terms, i.e., if

$$P_1(x, y) = ax + by$$

and

$$Q_1(x, y) = cx + dy$$

with $a, b,$ c and d not all zero, then the only possible directions in which trajectories can approach the origin are given by directions θ which satisfy

$$b \sin^2 \theta + (a - d) \sin \theta \cos \theta - \cos^2 \theta = 0. \tag{6}$$

For $\cos \theta \neq 0$ in this equation, i.e., if $b \neq 0$, this equation is equivalent to

$$b \tan^2 \theta + (a - d) \tan \theta - \mathrm{c} = 0. \tag{6'}$$

This quadratic has at most two solutions $\theta \in (-\pi/2, \pi/2]$ and if $\theta = \theta_1$ is a solution then $\theta = \theta_1 \pm \pi$ are also solutions. Finding the solutions of (6′) is equivalent to finding the directions determined by the eigenvectors of the matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

The next theorem follows immediately from the Stable Manifold Theorem and the Hartman-Grobman Theorem. It establishes that if the origin is a hyperbolic equilibrium point of the nonlinear system (2), then it is a (topological) saddle for (2) if and only if it is a saddle for its linearization at the origin. Furthermore, the directions θ_j along which the separatrices Γ_j approach the origin are solutions of (6).

Theorem 3. Suppose that E is an open subset of $\mathbf{R}^2$ containing the origin and that $\mathrm{f} \in C^1(E)$. If the origin is a hyperbolic equilibrium point of the nonlinear system (2), then the origin is a (topological) saddle for (2) if and only if the origin is a saddle for the linear system (1) with $A = D\mathrm{f}(0)$.

Theorem 4. Let E be an open subset of $\mathbf{R}^2$ containing the origin and let $\mathbf{f} \in C^2(E)$. Suppose that the origin is a hyperbolic critical point of (2). Then the origin is a stable (or unstable) node for the nonlinear system (2) if and only if it is a stable (or unstble) node for the linear system (1) with $A = D\mathbf{f}(\mathbf{0})$. And the origin is a stable (or unstable) focus for the nonlinear system (2) if and only if it is a stable (or unstable) focus for the linear system (1) with $A = D\mathrm{f}(\mathbf{0})$.

Remark. Under the hypotheses of Theorem 4, it follows that the origin is a proper node for the nonlinear system (2) if and only if it is a proper node for the linear system (1) with $A = D\mathrm{f}(0)$. And under the weaker hypothesis that $\mathrm{f} \in C^1(E)$, it still follows that if the origin is a focus for the linear system (1) with $A = D\mathbf{f}(0)$, then it is a focus for the nonlinear system (2).

Theorem 5. Let E be an open subset of $\mathbf{R}^2$ containing the origin and let $\mathbf{f} \in C^1(E)$ with $\mathbf{f}(\mathbf{0}) = \mathbf{0}$. Suppose that the origin is a center for the linear system (1) with $A = D\mathbf{f}(\mathbf{0})$. Then the origin is either a center, a center-focus or a focus for the nonlinear system (2).

Proof. We may assume that the matrix $A = D\mathbf{f}(0)$ has been transformed to its canonical form

$$A = \begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix}$$

with $b \neq 0$. Assume that $b > 0$; otherwise we can apply the linear transformation $t \to -t$. The nonlinear system (3) then has the form

$$\dot{x} = -by + p(x, y)$$

$$\dot{y} = bx + q(x, y)$$

Since f $\in C^1(E)$, it follows that $|p(x, y)/r| \to 0$ and $|q(x, y)/r| \to 0$ as $r \to 0$; i.e., $p = o(r)$ and $q = o(r)$ as $r \to 0$. Thus, in polar coordinates we have $\dot{r} = \mathrm{o}(r)$ and $\dot{\theta} = b + o(1)$ as $r \to 0$. Therefore, there exists a $\delta > 0$ such that $\dot{\theta} \geq b/2 > 0$ for $0 < r \leq \delta$. Thus for $0 < r_0 \leq \delta$ and $\theta_0 \in \mathbf{R}$, $\theta\,(t, r_0, \theta_0) \geq bt/2 + \theta_0 \to \infty$ as $t \to \infty$; and $\theta\,(t, r_0, \theta_0)$ is a monotone increasing function of t. Let $t = h(\theta)$ be the inverse of this monotone function. Define $\tilde{r}(\theta) = r\,(h(\theta), r_0, \theta_0)$ for $0 < r_0 \leq \delta$ and $\theta_0 \in \mathbf{R}$. Then $\tilde{r}(\theta)$ satisfies the differential equation (5) which has the form

$$\frac{d\tilde{r}}{d\theta} = \tilde{F}(\tilde{r}, \theta) = \frac{\cos\theta p(\tilde{r}\cos\theta, \tilde{r}\sin\theta) + \sin\theta q(\tilde{r}\cos\theta, \tilde{r}\sin\theta)}{b + (\cos\theta/\tilde{r})q(\tilde{r}\cos\theta, \tilde{r}\sin\theta) - (\sin\theta/\tilde{r})p(\tilde{r}\cos\theta, \tilde{r}\sin\theta)}$$

Suppose that the origin is not a center or a center-focus for the nonlinear system (3). Then for $\delta > 0$ sufficiently small, there are no closed trajectories of (3) in the deleted neighborhood $N_\delta(0) \sim \{0\}$. Thus for $0 < r_0 < \delta$ and $\theta_0 \in \mathbf{R}$, either $\tilde{r}\,(\theta_0 + 2\pi) < \tilde{r}\,(\theta_0)$ or $\tilde{r}\,(\theta_0 + 2\pi) > \tilde{r}\,(\theta_0)$. Assume that the first case holds. The second case is treated in a similar manner. If $\tilde{r}\,(\theta_0 + 2\pi) < \tilde{r}\,(\theta_0)$ then $\tilde{r}\,(\theta_0 + 2k\pi) < \tilde{r}\,(\theta_0 + 2(k-1)\pi)$ for $k = 1, 2, 3\ldots$ Otherwise we would have two trajectories of (3) through the same point which is impossible. The sequence $\tilde{r}\,(\theta_0 + 2k\pi)$ is monotone decreasing and bounded below by zero; therefore, the following limit exists and is nonnegative:

$$\tilde{r}_1 = \lim_{k\to\infty} \tilde{r}\,(\theta_0 + 2k\pi)$$

If $\tilde{r}_1 = 0$ then $\tilde{r}(\theta) \to 0$ as $\theta \to \infty$; i.e., $r\,(t, r_0, \theta_0) \to 0$ and $\theta\,(t, r_0, \theta_0) \to \infty$ as $t \to \infty$ and the origin is a stable focus of (3). If $\tilde{r}_1 > 0$ then since $|\tilde{F}(r, \theta)| \leq M$ for $0 \leq r \leq \delta$ and $0 \leq \theta \leq 2\pi$, the sequence $\tilde{r}\,(\theta_0 + \theta + 2k\pi)$ is equicontinuous on $[0, 2\pi]$. Therefore, by Ascoli's Lemma, cf. Theorem 7.25 in Rudin [R], there exists a uniformly convergent subsequence of $\tilde{r}\,(\theta_0 + \theta +\ 2k\pi)$ converging to a solution $\tilde{r}_1(\theta)$ which satisfies $\tilde{r}_1(\theta) = \tilde{r}_1(\theta + 2k\pi)$; i.e., $\tilde{r}_1(\theta)$ is a non-zero periodic solution of (5). This contradicts the fact that

there are no closed trajectories of (3) in $N_\delta(\mathbf{0}) \sim \{0\}$ when the origin is not a center or a center focus of (3). Thus if the origin is not a center or a center focus of (3), $\tilde{r}_1 = 0$ and the origin is a focus of (3). This completes the proof of the theorem.

A center-focus cannot occur in an analytic system. This is a consequence of Dulac's Theorem discussed in Section 3.3 of Chapter 3. We therefore have the following corollary of Theorem 5 for analytic systems.

Corollary. Let E be an open subset of $\mathbf{R}^2$ containing the origin and let $\mathbf{f}$ be analytic in E with $\mathbf{f}(\mathbf{0}) = \mathbf{0}$. Suppose that the origin is a center for the linear system (1) with $A = D\mathbf{f}(\mathbf{0})$. Then the origin is either a center or a focus for the nonlinear system (2).

As we noted in the previous section, Liapunov's method is one tool that can be used to distinguish a center from a focus for a nonlinear system. Another approach is to write the system in polar coordinates as in Examples 1-3 above. Yet another approach is to look for symmetries in the differential equations. The easiest symmetries to see are symmetries with respect to the x and y axes.

Definition 6. The system (3) is said to be symmetric with respect to the x-axis if it is invariant under the transformation $(t, y) \to (-t, -y)$; it is said to be symmetric with respect to the y-axis if it is invariant under the transformation $(t, x) \to (-t, -x)$.

Note that the system in Example 1 is symmetric with respect to the x-axis, but not with respect to the y-axis.

Theorem 6. Let E be an open subset of $\mathbf{R}^2$ containing the origin and let $\mathbf{f} \in C^1(E)$ with $\mathbf{f}(\mathbf{0}) = \mathbf{0}$. If the nonlinear system (2) is symmetric with respect to the x-axis or the y-axis, and if the origin is a center for the linear system (1) with $A = D\mathbf{f}(\mathbf{0})$, then the origin is a center for the nonlinear system (2).

The idea of the proof of this theorem is that by Theorem 5, any trajectory of (3) in $N_\delta(0)$ which crosses the positive x-axis will also cross the negative x-axis. If the system (3) is symmetric with respect to the x-axis, then the trajectories of (3) in $N_\delta(0)$ will be symmetric with respect to the x-axis and hence all trajectories of (3) in $N_\delta(0)$ will be closed; i.e., the origin will be a center for (3).

3.10 Nonhyperbolic Critical Points in R^2

In this section we present some results on nonhyperbolic critical points of planar analytic systems. This work originated with Poincaré [P] and was extended by Bendixson [B] and more recently by Andronov et al. [A – I]. We assume that the origin is an isolated critical point of the planar system

$$\begin{aligned} \dot{x} &= P(x, y) \\ \dot{y} &= Q(x, y) \end{aligned} \tag{1}$$

where P and Q are analytic in some neighborhood of the origin. In Sections 2.9 and 2.10 we have already presented some results for the case when the matrix of the linear part $A = D\mathbf{f}(0)$ has pure imaginary eigenvalues, i.e., when the origin is a center for the linearized system. In this section we give some results established in [A – I] for the case when the matrix A has one or two zero eigenvalues, but $A \neq 0$. And these results are extended to higher dimensions in Section 2.12.

First of all, note that if P and Q begin with m th-degree terms P_m and Q_m, then it follows from Theorem 2 in Section 2.10 that if the function

$$g(\theta) = \cos\theta Q_m(\cos\theta, \sin\theta) - \sin\theta P_m(\cos\theta, \sin\theta)$$

is not identically zero, then there are at most $2(m+1)$ directions $\theta = \theta_0$ along which a trajectory of (1) may approach the origin. These directions are given by solutions of the equation $g(\theta) = 0$. Suppose that $g(\theta)$ is not identically zero, then the solution curves of (1) which approach the origin along these tangent lines divide a neighborhood of the origin into a finite number of open regions called sectors. These sectors will be of one of three types as described in the following definitions; cf. [A-I] or [L]. The trajectories which lie on the boundary of a hyperbolic sector are called separatrices. Cf. Definition 1 in Section 3.11.

Definition 1. A sector which is topologically equivalent to the sector shown in Figure 1(a) is called a hyperbolic sector. A sector which is topologically equivalent to the sector shown in Figure 1(b) is called a parabolic sector. And a sector which is topologically equivalent to the sector shown in Figure 1(c) is called an elliptic sector.

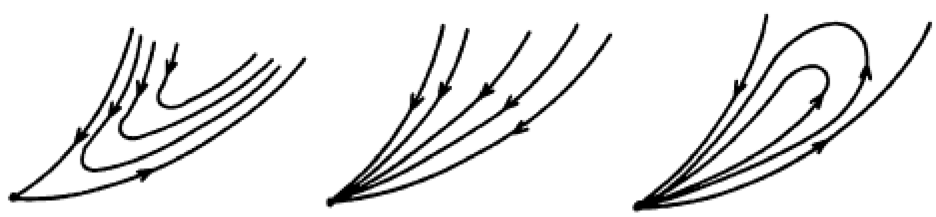

Figure 1. (a) A hyperbolic sector. (b) A parabolic sector. (c) An elliptic sector.

In Definition 1, the homeomorphism establishing the topological equivalence of a sector to one of the sectors in Figure 1 need not preserve the direction of the flow; i.e., each of the sectors in Figure 1 with the arrows reversed are sectors of the same type. For example, a saddle has a deleted neighborhood consisting of four hyperbolic sectors and four separatrices. And a proper node has a deleted neighborhood consisting of one parabolic sector. According to Theorem 2 below, the system

$$\dot{x} = y$$
$$\dot{y} = -x^3 + 4xy$$

has an elliptic sector at the origin; cf. Problem 1 below. The phase portrait for this system is shown in Figure 2. Every trajectory which approaches the origin does so tangent to the x-axis.

A deleted neighborhood of the origin consists of one elliptic sector, one hyperbolic sector, two parabolic sectors, and four separatrices. Cf. Definition 1 and Problem 5 in Section 3.11. This type of critical point is called *a* critical point with an elliptic domain; cf. [A-I].

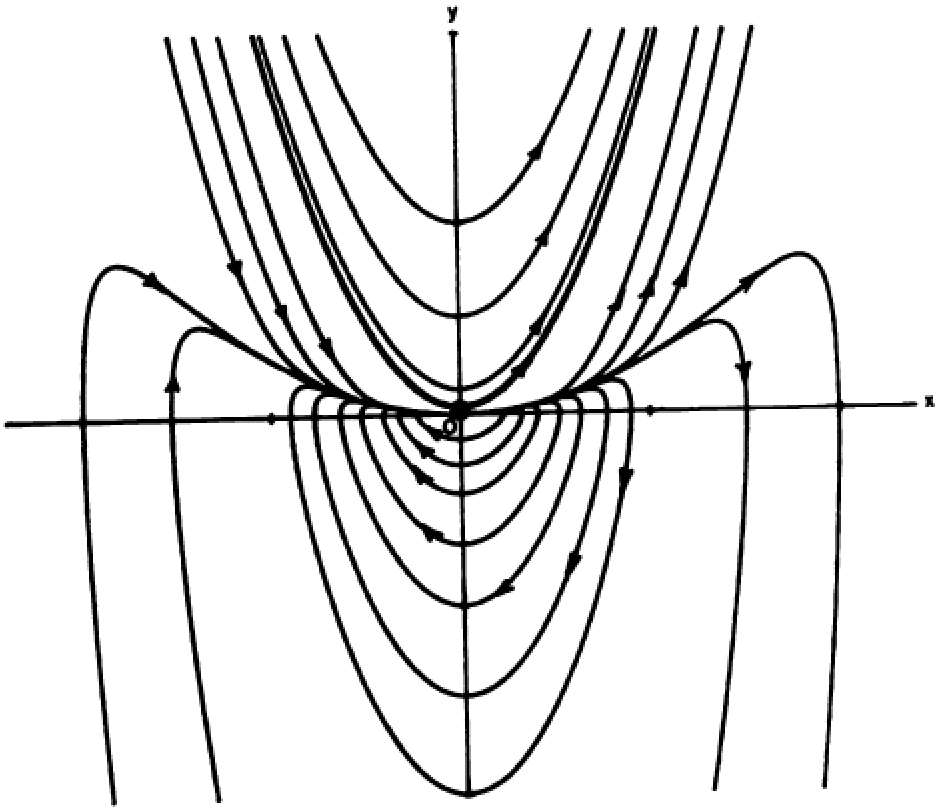

Figure 2. A critical point with an elliptic domain at the origin.

Another type of nonhyperbolic critical point for a planar system is a saddle-node. A saddle-node consists of two hyperbolic sectors and one parabolic sector (as well as three separatrices and the critical point itself). According to Theorem 1 below, the system

$$\dot{x} = x^2$$

$$\dot{y} = y$$

has a saddle-node at the origin. Even without Theorem 1, this system is easy to discuss since it can be solved explicitly for $x(t) = (1/x_0 - t)^{-1}$ and $y(t) = y_0 e^t$. The phase portrait for this system is shown in Figure 3.

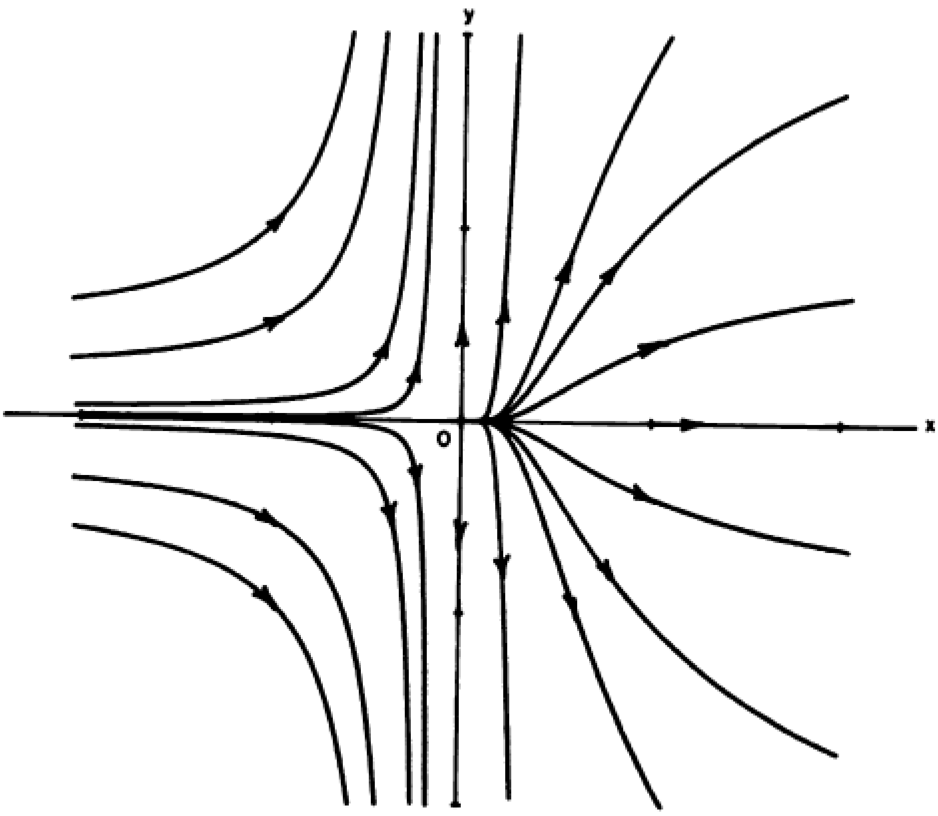

Figure 3. A saddle-node at the origin.

One other type of behavior that can occur at a nonhyperbolic critical point is illustrated by the following example:

$$\dot{x} = y$$
$$\dot{y} = x^2$$

The phase portrait for this system is shown in Figure 4. We see that a deleted neighborhood of the origin consists of two hyperbolic sectors and two separatrices. This type of critical point is called a cusp.

As we shall see, besides the familiar types of critical points for planar analytic systems discussed in Section 2.10, i.e., nodes, foci, (topological) saddles and centers, the only other types of critical points that can occur for (1) when $A \neq 0$ are saddle-nodes, critical points with elliptic domains and cusps.

We first consider the case when the matrix A has one zero eigenvalue, i.e., when $\det A = 0$, but $\operatorname{tr} A \neq 0$. In this case, as in Chapter 1 and as is shown in [A – I] on p. 338 , the system (1) can be put into the form

$$\dot{x} = p_2(x, y) \tag{2}$$
$$\dot{y} = y + q_2(x, y)$$

where p_2 and q_2 are analytic in a neighborhood of the origin and have expansions that begin with second-degree terms in x and y. The following theorem is proved on p. 340 in [A-I]. Cf. Section 2.12.

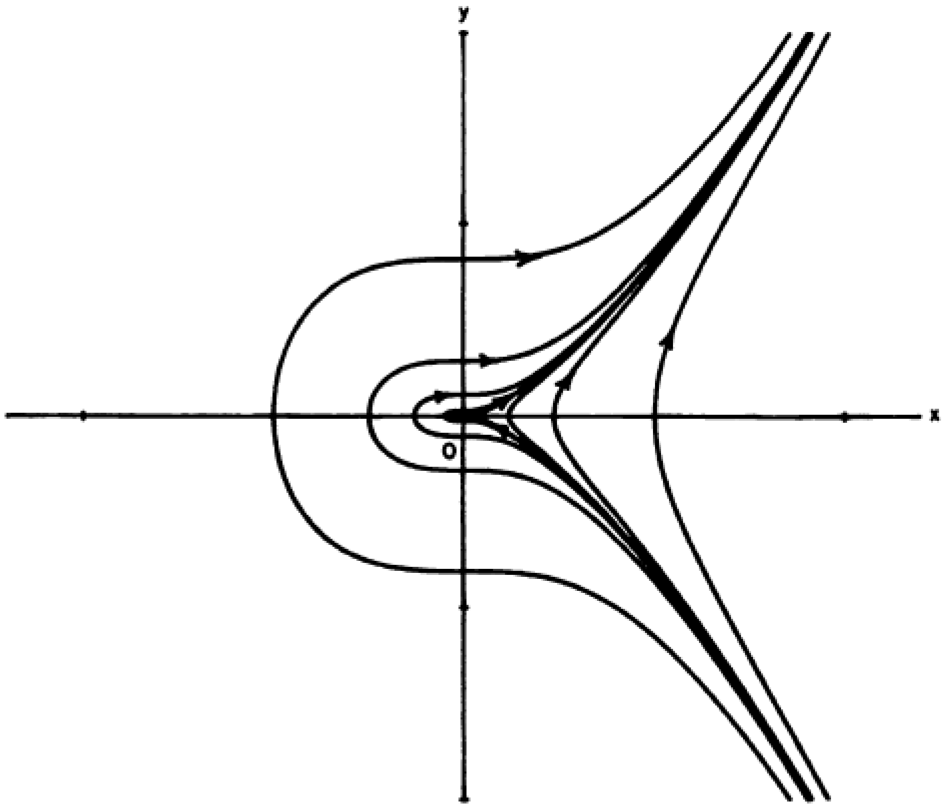

Figure 4. A cusp at the origin.

Theorem 1. Let the origin be an isolated critical point for the analytic system (2). Let $y = \phi(x)$ be the solution of the equation $y + q_2(x, y) = 0$ in a neighborhood of the origin and let the expansion of the function $\psi(x) = p_2(x, \phi(x))$ in a neighborhood of $x = 0$ have the form $\psi(x) = a_m x^m + \cdots$ where $m \geq 2$ and $a_m \neq 0$. Then (1) for m odd and $a_m > 0$, the origin is an unstable node, (2) for m odd and $a_m < 0$, the origin is a (topological) saddle and (3) for m even, the origin is a saddle-node.

Next consider the case when A has two zero eigenvalues, i.e., $\det A = 0$, $\operatorname{tr} A = 0$, but $A \neq 0$. In this case it is shown in [A-I], p. 356, that the system (1) can be put in the "normal" form

$$\begin{aligned}\dot{x} &= y\\ \dot{y} &= a_k x^k[1 + h(x)] + b_n x^n y[1 + g(x)] + y^2 R(x, y)\end{aligned} \tag{3}$$

where $h(x), g(x)$ and $R(x, y)$ are analytic in a neighborhood of the origin, $h(0) = g(0) = 0, k \geq 2, a_k \neq 0$ and $n \geq 1$. Cf. Section 2.13. The next two theorems are proved on pp. 357-362 in [A-I].

Theorem 2. Let $k = 2m + 1$ with $m \geq 1$ in (3) and let $\lambda = b_n^2 + 4(m + 1)a_k$. Then if $a_k > 0$, the origin is a (topological) saddle. If $a_k < 0$, the origin is (1) a focus or a center if $b_n = 0$ and also if $b_n \neq 0$ and $n > m$ or if $n = m$ and $\lambda < 0$, (2) a node if $b_n \neq 0, n$ is an even number and $n < m$ and also if $b_n \neq 0, n$ is an even number, $n = m$ and $\lambda \geq 0$ and (3) a critical point with an elliptic domain if $b_n \neq 0, n$ is an odd number and $n < m$ and also if $b_n \neq 0, n$ is an odd number, $n = m$ and $\lambda \geq 0$.

Theorem 3. Let $k = 2m$ with $m \geq 1$ in (3). Then the origin is (1) a cusp if $b_n = 0$ and also if $b_n \neq 0$ and $n \geq m$ and (2) a saddle-node if $b_n \neq 0$ and $n < m$.

We see that if $D\mathbf{f}(\mathbf{x}_0)$ has one zero eigenvalue, then the critical point $\mathbf{x}_0$ is either a node, a (topological) saddle, or a saddle-node; and if $D\mathbf{f}(\mathrm{x}_0)$ has two zero eigenvalues, then the critical point x_0 is either a focus, a center, a node, a (topological) saddle, a saddle-node, a cusp, or a critical point with an elliptic domain.

Finally, what if the matrix $A = 0$? In this case, the behavior near the origin can be very complex. If P and Q begin with m th-degree terms, then the separatrices may divide a neighborhood of the origin into $2(m + 1)$ sectors of various types. The number of elliptic sectors minus the number of hyperbolic sectors is always an even number and this number is related to the index of the critical point discussed in Section 3.12 of Chapter 3. For example, the homogenous quadratic system

$$\begin{aligned}\dot{x} &= x^2 + xy\\ \dot{y} &= \frac{1}{2}y^2 + xy\end{aligned}$$

has the phase portrait shown in Figure 5. There are two elliptic sectors and two parabolic sectors at the origin. All possible types of phase portraits for homogenous, quadratic systems have been classified by the Russian mathematician L.S. Lyagina [19]. For more information on the topic, cf. the book by Nemytskii and Stepanov [N/S].

Remark. A critical point, $\mathbf{x}_0$, of (1) for which $D\mathbf{f}(\mathbf{x}_0)$ has a zero eigenvalue is often referred to as a multiple critical point. The reason for this is made clear in Section 4.2 of Chapter 4 where it is shown that a multiple critical point of (1) can be made to split into a number of hyperbolic critical points under a suitable perturbation of (1).

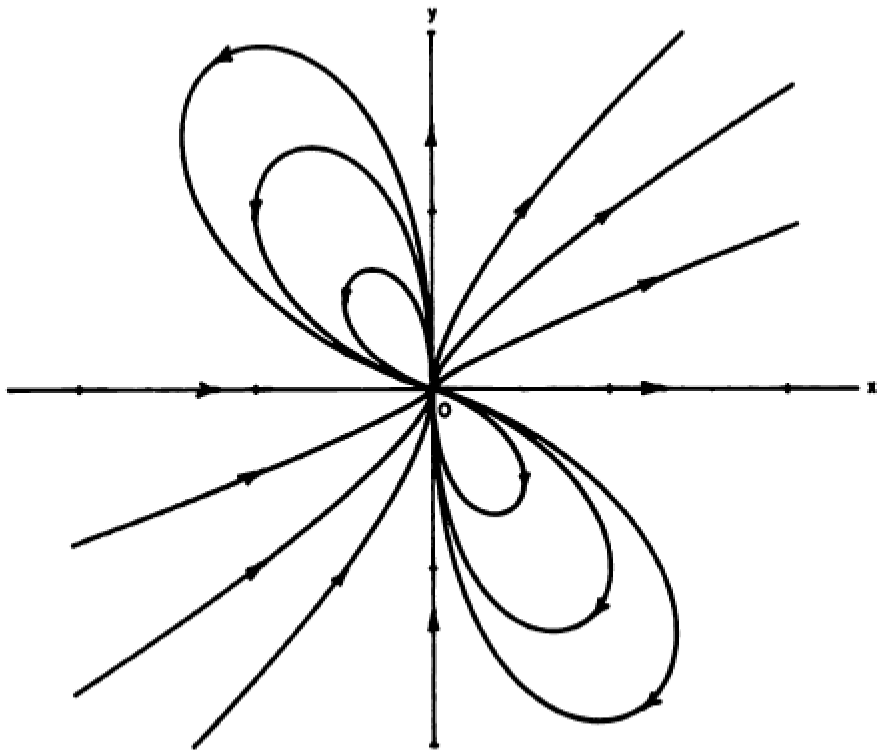

Figure 5. A nonhyperbolic critical point with two elliptic sectors and two parabolic sectors.

BIBLIOGRAPHY

[1] P. Auger, C. Lett et J. C. Poggiale. *Modélisation mathématique en écologie : Cours et exercices corrigés*. Dunod, Paris, 2010.

[2] G. C. Layek. *An introduction to dynamical systems and chaos*. Springer, India, 2015.

[3] J. L. Pac. *Systèmes dynamiques : Cours et exercices corrigés*. Dunod, Paris, 2016.

[4] H. O. Peitgen, H. Jürgens et D. Saupe. *Chaos and Fractals*. Springer-Verlag, New York,Inc , 2004.

[5] L. Perko. *Differential Equations and Dynamical Systems*. Springer, 2006.

[6] S. H. Strogatz, Nonlinear dynamics and chaos : *With Applications to Physics, Biology, Chemistry, and Engineering*. Perseus Books Publishing, L.L.C, 1994.

[7] T. Vialar et A. Goergen. *Complex and Chaotic Nonlinear Dynamics*. Springer-Verlag Berlin, Heidelberg, 2009.

[8] G. Teschl, Ordinary differential equations and dynamical systems, American Mathematical Society, 2011.

[9] M-S. Abdelouahab, Equations différentielles II, Cours L3, Centre universitaire de Mila,2011.

[10] Robinson, C. Dynamical Systems: Stability, Symbolic Dynamics, and Chaos. Boca Raton, FL: CRC Press, 1995

[11] Morris W. Hirsch, Stephen Smale, and Robert L. Devaney. Differential equations, dynamical systems, and an introduction to chaos. Elsevier/Academic Press, Amsterdam, third edition, 2013