## 1.4 Stability and Linearization

Suppose we are studying a physical system whose states $x$ are described by points in $\mathbb{R}^n$. Assume that the dynamics of the system is governed by a given evolution equation

$$\frac{dx}{dt} = f(x).$$

Let $x_0$ be a *stationary point* of the dynamics, i.e., $f(x_0) = 0$. Imagine that we perform an experiment on the system at time $t = 0$ and determine that the initial state is indeed $x_0$. Are we justified in predicting that the system will remain at $x_0$ for all future time? The mathematical answer to this question is yes, but unfortunately it is probably not the question we really wished to ask. Experiments in real life seldom yield exact answers to our idealized models, so in most cases we will have to ask whether the system will remain *near* $x_0$ if it started *near* $x_0$. The answer to the revised question is not always yes, but even so, by examining the evolution equation at hand more carefully, one can sometimes make predictions about the future behavior of a system starting near $x_0$. A simple example will illustrate some of the problems involved. Consider the following two differential equations on the real line:

$$x'(t) = -x(t) \tag{1.4.1}$$

and

$$x'(t) = x(t). \tag{1.4.2}$$

The solutions are, respectively,

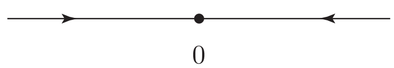
$$x(x_0, t) = x_0 e^{-t} \tag{1.4.3}$$

and

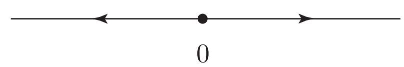
$$x(x_0, t) = x_0 e^{+t}. \tag{1.4.4}$$

Note that 0 is a stationary point of both equations. In the first case, for all $x_0 \in \mathbb{R}$, we have $\lim_{t\to\infty} x(x_0, t) = 0$. The whole real line moves toward the origin, and the prediction that, if $x_0$ is near 0 then $x(x_0, t)$ is near 0, is justified. On the other hand, suppose we are observing a system whose state $x$ is governed by equation (0.1.2). An experiment telling us that at time $t = 0$, $x(0)$ is approximately zero will certainly not permit us to conclude that $x(t)$ stays near the origin for all time, since all points except 0 move rapidly away from 0. Furthermore, our experiment is unlikely to allow us to make an accurate prediction about $x(t)$ because if $x(0) < 0$, $x(t)$ moves rapidly away from the origin toward $-\infty$, but if $x(0) > 0$, $x(t)$ moves toward $+\infty$. Thus, an observer watching such a system would expect sometimes to observe $x(t) \to -\infty$ as $t \to \infty$ and sometimes $x(t) \to +\infty$ as $t \to \infty$. The solution $x(t) = 0$ for all $t$ may be difficult to observe because a slight perturbation of the initial state would destroy this solution. This sort of behavior is frequently observed in nature. It is not due to any nonuniqueness in the solution to the differential equation involved, but to the *instability* of that solution under small perturbations in initial data.

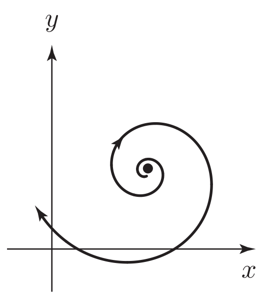
It is convenient to represent the dynamics by sketching representative flow lines (solution curves) in the state space, as in Figure 1.4.1. In this figure we also indicate examples of unstable and stable points on the line, in the plane $(\mathbb{R}^2)$ and in space $(\mathbb{R}^3)$.



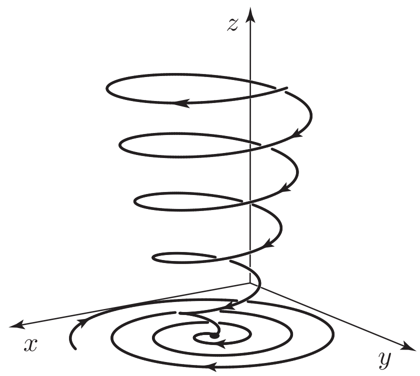

(a) flow lines for a stable point on the line




(b) flow lines for an unstable point on the line




(c) an unstable point for a dynamical system in the plane




(d) a stable point for a dynamical system in space

FIGURE 1.4.1. Stable and unstable equilibria.

Since questions of stability are central in dynamical systems, we will want to define the concept of stability precisely and develop criteria for determining it.

**Definition 1.4.1.** *Consider a $C^1$ dynamical system $\dot{x} = X(x)$ on $\mathbb{R}^n$, and suppose that $x_e$ is a fixed point of $X$; that is, $X(x_e) = 0$.*

1. *We say that the point $x_e$ is* ***stable*** *if for every $b > 0$, there is an $\epsilon > 0$ such that if an initial condition $x_0$ lies in the ball of radius $\epsilon$ around $x(e)$ then it exists for all $t > 0$ and stays in the ball of radius $b$.*

2. *We say that $x_e$ is* ***asymptotically stable*** *if it is stable and in addition, the solutions $x(t)$ with initial conditions in the ball of radius $\epsilon$ converge to $x(e)$ as $t \to +\infty$*

*If $x(e)$ is not stable it is called* ***unstable****. (See Figure 1.4.2.)*

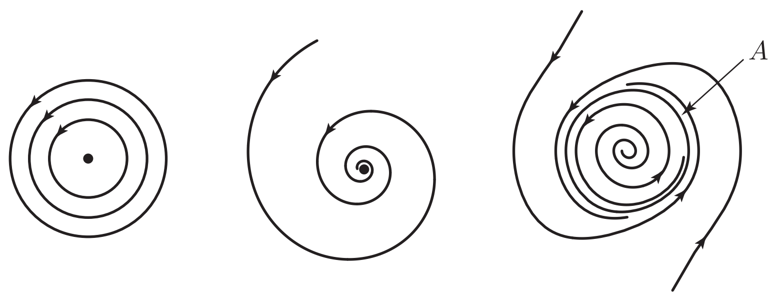



FIGURE 1.4.2. Stability for fixed points and closed orbits.

**Eigenvalue Criteria for Stability.** There is a classical test for stability of fixed points due to Liapunov. To begin, consider the linear case. Letting $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map, the $\dot{x} = Ax$ has flow $F_t(x)$ where $F_t : \mathbb{R}^n \to \mathbb{R}^n$ is the linear map given by

$$F_t = e^{tA}, \tag{1.4.5}$$

where

$$e^{tA} = I + tA + \frac{1}{2}t^2A^2 + \cdots + \frac{1}{n!}t^nA^n + \cdots .$$

Let $\lambda_1, \ldots, \lambda_n$ be the (complex) eigenvalues of $A$.

Recall the following result from the section on *Linear Systems.*

**Proposition 1.4.2.** *For the linear case, the origin is*

**i.** *asymptotically stable if*

$$\text{Re } \lambda_i < 0 \quad \textit{for all} \quad i = i, \ldots, n;$$

**ii.** *unstable if*

$$\text{Re } \lambda_i > 0 \quad \textit{for some } i.$$

When some eigenvalues are on the imaginary axis, further investigation is needed.

Recall that the preceding proposition can be proved using the Jordan normal form; it is especially easy when $A$ is diagonalizable (over the complex numbers) with linearly independent eigenvectors $v_1, \ldots, v_n$, for in that basis $A$ is the diagonal matrix

$$\text{diag}(e^{t\lambda_1}, \ldots, e^{t\lambda_n})$$

and $e^{t\lambda_i} \to 0$ as $t \to \infty$ if $\operatorname{Re}\lambda_i < 0$.

A *local* nonlinear version of this result is as follows.

**Theorem 1.4.3** (Liapunov's Theorem)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be $C^1$ and $x_0 \in \mathbb{R}^n$ be a fixed point of $\dot{x} = f(x)$ (so $f(x_0) = 0$). Let $A = \mathbf{D}f(x_0)$ be the linearization of $f$ (so $A_{ij} = \partial f_i/\partial x_j$ is the Jacobian matrix) and $\lambda_1, \ldots, \lambda_n$ be its eigenvalues. Then $x_0$ is*

**i.** *asymptotically stable if*

$$\operatorname{Re}\lambda_i < 0 \quad \textit{for all} \quad i = 1, \ldots, n;$$

**ii.** *unstable if*

$$\operatorname{Re}\lambda_i > 0 \quad \textit{for some} \quad i.$$

*If the eigenvalues all have real parts zero, then further analysis is necessary.*

We will prove the stability result here. The instability result can be proved by a similar method, but it is a bit more subtle.

**Proof.** Without loss of generality, we can assume that the equilibrium point is at the origin. In the section on linear systems we saw that there is an $\varepsilon > 0$ such that $\|e^{tA}\| \le Me^{-\varepsilon t}$.

From the local existence theory, there is an $r$-ball about 0 for which the time of existence is uniform if the initial condition $x_0$ lies in this ball. Let

$$R(x) = X(x) - Ax.$$

Find $r_2 \le r$ such that $\|x\| \le r_2$ implies $\|R(x)\| \le \alpha\|x\|$, where $\alpha = \varepsilon/2$. This is possible by considering the definition of the derivative at the origin.

Let $D$ be the open $r_2/2$ ball about 0 and let $x_0 \in D$. The strategy is to show that if $x_0 \in D$, then the integral curve starting at $x_0$ remains in $D$ and converges to the origin exponentially as $t \to +\infty$. This will prove the stability result.

To carry out this plan, let $x(t)$ be the integral curve of $X$ starting at $x_0$. Suppose $x(t)$ remains in $D$ for $0 \le t < T$. The equation

$$\dot{x} = Ax(t) + R(x(t))$$

with initial conditions $x(0) = x_0$ has a solution that satisfies the variation of constants formula, namely

$$x(t) = e^{tA}x_0 + \int_0^t e^{(t-s)A}R(x(s))\,ds,$$

and so

$$\|x(t)\| \le Me^{-t\varepsilon}\|x_0\| + \alpha\int_0^t e^{-(t-s)\varepsilon}\|x(s)\|\,ds.$$

Letting $f(t) = e^{t\varepsilon}\|x(t)\|$, the previous inequality becomes

$$f(t) \le M\|x_0\| + \alpha \int_0^1 f(s)\, ds,$$

and so, by Gronwall's inequality, $f(t) \le M\|x_0\|e^{\alpha t}$. Thus

$$\|x(t)\| \le M\|x_0\|e^{(\alpha-\varepsilon)t} = M\|x_0\|e^{-\varepsilon t/2}.$$

Hence $x(t) \in D$, $0 \le t < T$, so as in the local continuation results, $x(t)$ may be indefinitely extended in $t$ and the foregoing estimate holds. ■

**Examples.**

1. Consider the vector field $f$ on $\mathbb{R}^2$ defined by

$$f(x, y) = (y, \mu(1 - x^2)y - x)$$

where $\mu$ is a real parameter. That is, consider the nonlinear system

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= \mu(1 - x^2)y - x \end{aligned}$$

Using Theorem 1.4.3 we see that the linearization at the origin is given by the linear system

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -x + \mu y; \end{aligned}$$

that is,

$$\frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & \mu \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

The eigenvalues of this matrix are

$$\lambda = \frac{1}{2}\left(\mu \pm \sqrt{\mu^2 - 4}\right)$$

and we see that the real parts (examine the cases $|\mu| > 2$ and $|\mu| < 2$ separately) have negative real parts for $\mu < 0$, so it is stable and have at least one positive real part if $\mu > 0$, so the origin is unstable. For $\mu = 0$, the eigenvalues are pure imaginary, so we do not draw a conclusion—but in fact, a direct examination shows that the system is a linear center for $\mu = 0$. A computer simulation confirms these findings.