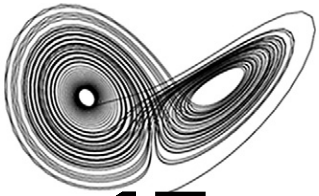
Differential equations, dynamical systems, and an introduction to Chaos - M.W. Hirsh, S. Smale, R. L. Denavey - Elsevier



17 Existence and Uniqueness Revisited

In this chapter we return to the material presented in Chapter 7, this time filling in all of the technical details and proofs that were omitted earlier. As a result, this chapter is more difficult than the preceding ones; it is, however, central to the rigorous study of ordinary differential equations. To comprehend thoroughly many of the proofs in this section, the reader should be familiar with such topics from real analysis as uniform continuity, uniform convergence of functions, and compact sets.

17.1 The Existence and Uniqueness Theorem

Consider the autonomous system of differential equations

$$X' = F(X)$$

where $F\colon \mathbb{R}^n \to \mathbb{R}^n$. In previous chapters, we have usually assumed that F was C^∞; here we will relax this condition and assume that F is only C^1. Recall that this means that F is continuously differentiable. That is, F and its first partial derivatives exist and are continuous functions on $\mathbb{R}^n$. For the first few sections of this chapter, we will deal only with autonomous equations; later we will assume that F depends on t as well as X.

As we know, a solution of this system is a differentiable function $X: J \to \mathbb{R}^n$ defined on some interval $J \subset \mathbb{R}$ such that for all $t \in J$

$$X'(t) = F(X(t)).$$

Geometrically, $X(t)$ is a curve in $\mathbb{R}^n$ whose tangent vector $X'(t)$ equals $F(X(t))$; as in previous chapters, we think of this vector as being based at $X(t)$, so that the map $F: \mathbb{R}^n \to \mathbb{R}^n$ defines a vector field on $\mathbb{R}^n$. An *initial condition* or *initial value* for a solution $X: J \to \mathbb{R}^n$ is a specification of the form $X(t_0) = X_0$ where $t_0 \in J$ and $X_0 \in \mathbb{R}^n$. For simplicity, we usually take $t_0 = 0$.

A nonlinear differential equation may have several solutions that satisfy a given initial condition. For example, consider the first-order nonlinear differential equation

$$x' = 3x^{2/3}.$$

In Chapter 7 we saw that the identically zero function $u_0: \mathbb{R} \to \mathbb{R}$ given by $u_0(t) \equiv 0$ is a solution satisfying the initial condition $u(0) = 0$. But $u_1(t) = t^3$ is also a solution satisfying this initial condition, and, in addition, for any $\tau > 0$, the function given by

$$u_\tau(t) = \begin{cases} 0 & \text{if } t \le \tau \\ (t-\tau)^3 & \text{if } t > \tau \end{cases}$$

is also a solution satisfying the initial condition $u_\tau(0) = 0$.

Besides uniqueness, there is also the question of existence of solutions. When we dealt with linear systems, we were able to compute solutions explicitly. For nonlinear systems, this is often not possible, as we have seen. Moreover, certain initial conditions may not give rise to any solutions. For example, as we saw in Chapter 7, the differential equation

$$x' = \begin{cases} 1 & \text{if } x < 0 \\ -1 & \text{if } x \ge 0 \end{cases}$$

has no solution that satisfies $x(0) = 0$.

Thus it is clear that, to ensure existence and uniqueness of solutions, extra conditions must be imposed on the function F. The assumption that F is continuously differentiable turns out to be sufficient, as we shall see. In the first example above, F is not differentiable at the problematic point $x = 0$, while in the second example, F is not continuous at $x = 0$.

The following is the fundamental local theorem of ordinary differential equations.

The Existence and Uniqueness Theorem. *Consider the initial value problem*

$$X' = F(X), \quad X(0) = X_0$$

where $X_0 \in \mathbb{R}^n$. Suppose that $F\colon \mathbb{R}^n \to \mathbb{R}^n$ is C^1. Then there exists a unique solution of this initial value problem. More precisely, there exists $a > 0$ and a unique solution

$$X\colon (-a, a) \to \mathbb{R}^n$$

of this differential equation satisfying the initial condition

$$X(0) = X_0.$$

■

We will prove this theorem in the next section.

17.2 Proof of Existence and Uniqueness

We need to recall some multivariable calculus. Let $F\colon \mathbb{R}^n \to \mathbb{R}^n$. In coordinates $(x_1, \ldots, x_n)$ on $\mathbb{R}^n$, we write

$$F(X) = (f_1(x_1, \ldots, x_n), \ldots, f_n(x_1, \ldots, x_n)).$$

Let DF_X be the derivative of F at the point $X \in \mathbb{R}^n$. We may view this derivative in two slightly different ways. From one point of view, DF_X is a linear map defined for each point $X \in \mathbb{R}^n$; this linear map assigns to each vector $U \in \mathbb{R}^n$ the vector

$$DF_X(U) = \lim_{h \to 0} \frac{F(X + hU) - F(X)}{h},$$

where $h \in \mathbb{R}$. Equivalently, from the matrix point of view, DF_X is the $n \times n$ Jacobian matrix

$$DF_X = \left(\frac{\partial f_i}{\partial x_j} \right)$$

where each derivative is evaluated at $(x_1, \ldots, x_n)$. Thus the derivative may be viewed as a function that associates different linear maps or matrices to each point in $\mathbb{R}^n$. That is, $DF\colon \mathbb{R}^n \to L(\mathbb{R}^n)$.

As earlier, the function F is said to be continuously differentiable, or C^1, if all of the partial derivatives of the f_j exist and are continuous. We will assume for the remainder of this chapter that F is C^1. For each $X \in \mathbb{R}^n$, we define the norm $|DF_X|$ of the Jacobian matrix DF_X by

$$|DF_X| = \sup_{|U|=1} |DF_X(U)|$$

where $U \in \mathbb{R}^n$. Note that $|DF_X|$ is not necessarily the magnitude of the largest eigenvalue of the Jacobian matrix at X.

Example. Suppose

$$DF_X = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then, indeed, $|DF_X| = 2$, and 2 is the largest eigenvalue of DF_X. However, if

$$DF_X = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

then

$$\begin{aligned} |DF_X| &= \sup_{0\leq\theta\leq 2\pi} \left| \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \right| \\ &= \sup_{0\leq\theta\leq 2\pi} \sqrt{(\cos\theta + \sin\theta)^2 + \sin^2\theta} \\ &= \sup_{0\leq\theta\leq 2\pi} \sqrt{1 + 2\cos\theta\sin\theta + \sin^2\theta} \\ &> 1 \end{aligned}$$

whereas 1 is the largest eigenvalue. ■

We do, however, have

$$|DF_X(V)| \leq |DF_X||V|$$

for any vector $V \in \mathbb{R}^n$. Indeed, if we write $V = (V/|V|)\,|V|$, then we have

$$|DF_X(V)| = \left|DF_X\big(V/|V|\big)\right|\,|V| \leq |DF_X||V|$$

since $V/|V|$ has magnitude 1. Moreover, the fact that $F\colon \mathbb{R}^n \to \mathbb{R}^n$ is C^1 implies that the function $\mathbb{R}^n \to L(\mathbb{R}^n)$, which sends $X \to DF_X$, is a continuous function.

Let $\mathcal{O} \subset \mathbb{R}^n$ be an open set. A function $F\colon \mathcal{O} \to \mathbb{R}^n$ is said to be *Lipschitz* on $\mathcal{O}$ if there exists a constant K such that

$$|F(Y) - F(X)| \leq K|Y - X|$$

for all $X, Y \in \mathcal{O}$. We call K a *Lipschitz constant* for F. More generally, we say that F is *locally Lipschitz* if each point in $\mathcal{O}$ has a neighborhood $\mathcal{O}'$ in $\mathcal{O}$ such that the restriction F to $\mathcal{O}'$ is Lipschitz. The Lipschitz constant of $F|\mathcal{O}'$ may vary with the neighborhoods $\mathcal{O}'$.

Another important notion is that of compactness. We say that a set $\mathcal{C} \subset \mathbb{R}^n$ is *compact* if $\mathcal{C}$ is closed and bounded. An important fact is that, if $f : \mathcal{C} \to \mathbb{R}$ is continuous and $\mathcal{C}$ is compact, then first of all f is bounded on $\mathcal{C}$ and, secondly, f actually attains its maximum on $\mathcal{C}$. See Exercise 13 at the end of this chapter.

Lemma. *Suppose that the function $F\colon \mathcal{O} \to \mathbb{R}^n$ is C^1. Then F is locally Lipschitz.*

Proof: Suppose that $F\colon \mathcal{O} \to \mathbb{R}^n$ is C^1 and let $X_0 \in \mathcal{O}$. Let $\epsilon > 0$ be so small that the closed ball $\mathcal{O}_\epsilon$ of radius ϵ about X_0 is contained in $\mathcal{O}$. Let K be an upper bound for $|DF_X|$ on $\mathcal{O}_\epsilon$; this bound exists because DF_X is continuous and $\mathcal{O}_\epsilon$ is compact. The set $\mathcal{O}_\epsilon$ is *convex*; that is, if $Y, Z \in \mathcal{O}_\epsilon$, then the straight-line segment connecting Y to Z is contained in $\mathcal{O}_\epsilon$. This straight line is given by $Y + sU \in \mathcal{O}_\epsilon$, where $U = Z - Y$ and $0 \leq s \leq 1$. Let $\psi(s) = F(Y + sU)$. Using the chain rule we find

$$\psi'(s) = DF_{Y+sU}(U).$$

Therefore

$$\begin{aligned} F(Z) - F(Y) &= \psi(1) - \psi(0) \\ &= \int_0^1 \psi'(s)\, ds \\ &= \int_0^1 DF_{Y+sU}(U)\, ds. \end{aligned}$$

Thus we have

$$|F(Z) - F(Y)| \leq \int_0^1 K|U|\, ds = K|Z - Y|. \qquad \blacksquare$$

The following remark is implicit in the proof of the lemma: If $\mathcal{O}$ is convex, and if $|DF_X| \leq K$ for all $X \in \mathcal{O}$, then K is a Lipschitz constant for $F \mid \mathcal{O}$.

Suppose that J is an open interval containing zero and $X: J \to \mathcal{O}$ satisfies

$$X'(t) = F(X(t))$$

with $X(0) = X_0$. Integrating, we have

$$X(t) = X_0 + \int_0^t F(X(s))\, ds.$$

This is the integral form of the differential equation $X' = F(X)$. Conversely, if $X: J \to \mathcal{O}$ satisfies this integral equation, then $X(0) = X_0$ and X satisfies $X' = F(X)$, as is seen by differentiation. Thus the integral and differential forms of this equation are equivalent as equations for $X: J \to \mathcal{O}$. To prove the existence of solutions, we will use the integral form of the differential equation.

We now proceed with the proof of existence. Here are our assumptions:

1. $\mathcal{O}_\rho$ is the closed ball of radius $\rho > 0$ centered at X_0.
2. There is a Lipschitz constant K for F on $\mathcal{O}_\rho$.
3. $|F(X)| \leq M$ on $\mathcal{O}_\rho$.
4. Choose $a < \min\{\rho/M,\ 1/K\}$ and let $J = [-a, a]$.

We will first define a sequence of functions $U_0, U_1, \ldots$ from J to $\mathcal{O}_\rho$. Then we will prove that these functions converge uniformly to a function satisfying the differential equation. Later we will show that there are no other such solutions. The lemma that is used to obtain the convergence of the U_k is the following:

Lemma from Analysis. *Suppose $U_k: J \to \mathbb{R}^n$, $k = 0, 1, 2, \ldots$ is a sequence of continuous functions defined on a closed interval J that satisfy: Given $\epsilon > 0$, there is some $N > 0$ such that for every $p, q > N$*

$$\max_{t \in J} |U_p(t) - U_q(t)| < \epsilon.$$

Then there is a continuous function $U: J \to \mathbb{R}^n$ such that

$$\max_{t \in J} |U_k(t) - U(t)| \to 0 \quad \text{as } k \to \infty.$$

Moreover, for any t with $|t| \leq a$,

$$\lim_{k \to \infty} \int_0^t U_k(s)\, ds = \int_0^t U(s)\, ds. \qquad \blacksquare$$

This type of convergence is called *uniform convergence* of the functions U_k. This lemma is proved in elementary analysis books and will not be proved here. See [38].

The sequence of functions U_k is defined recursively using an iteration scheme known as *Picard iteration.* We gave several illustrative examples of this iterative scheme back in Chapter 7. Let

$$U_0(t) \equiv X_0.$$

For $t \in J$ define

$$U_1(t) = X_0 + \int_0^t F(U_0(s))\, ds = X_0 + tF(X_0).$$

Since $|t| \leq a$ and $|F(X_0)| \leq M$, it follows that

$$|U_1(t) - X_0| = |t||F(X_0)| \leq aM \leq \rho$$

so that $U_1(t) \in \mathcal{O}_\rho$ for all $t \in J$. By induction, assume that $U_k(t)$ has been defined and that $|U_k(t) - X_0| \leq \rho$ for all $t \in J$. Then let

$$U_{k+1}(t) = X_0 + \int_0^t F(U_k(s))\, ds.$$

This makes sense since $U_k(s) \in \mathcal{O}_\rho$ so the integrand is defined. We show that $|U_{k+1}(t) - X_0| \leq \rho$ so that $U_{k+1}(t) \in \mathcal{O}_\rho$ for $t \in J$; this will imply that the sequence can be continued to U_{k+2}, U_{k+3}, and so on. This is shown as follows:

$$\begin{aligned} |U_{k+1}(t) - X_0| &\leq \int_0^t |F(U_k(s))|\, ds \\ &\leq \int_0^t M\, ds \\ &\leq Ma \ < \ \rho. \end{aligned}$$

Next, we prove that there is a constant $L \geq 0$ such that, for all $k \geq 0$,

$$\left|U_{k+1}(t) - U_k(t)\right| \leq (aK)^k L.$$

Let L be the maximum of $|U_1(t) - U_0(t)|$ over $-a \leq t \leq a$. By the above, $L \leq aM$. We have

$$|U_2(t) - U_1(t)| = \left| \int_0^t F(U_1(s)) - F(U_0(s))\, ds \right|$$

$$\leq \int_0^t K|U_1(s) - U_0(s)|\, ds$$
$$\leq aKL.$$

Assuming by induction that, for some $k \geq 2$, we have already proved

$$|U_k(t) - U_{k-1}(t)| \leq (aK)^{k-1}L$$

for $|t| \leq a$, we then have

$$\begin{aligned} |U_{k+1}(t) - U_k(t)| &\leq \int_0^t |F(U_k(s)) - F(U_{k-1}(s))|\, ds \\ &\leq K \int_0^t |U_k(s) - U_{k-1}(s)|\, ds \\ &\leq (aK)(aK)^{k-1}L \\ &= (aK)^k L. \end{aligned}$$

Let $\alpha = aK$, so that $\alpha < 1$ by assumption. Given any $\epsilon > 0$, we may choose N large enough so that, for any $r > s > N$ we have

$$\begin{aligned} |U_r(t) - U_s(t)| &\leq \sum_{k=N}^{\infty} \left|U_{k+1}(t) - U_k(t)\right| \\ &\leq \sum_{k=N}^{\infty} \alpha^k L \\ &\leq \epsilon \end{aligned}$$

since the tail of the geometric series may be made as small as we please.

By the lemma from analysis, this shows that the sequence of functions $U_0, U_1, \ldots$ converges uniformly to a continuous function $X\colon J \to \mathbb{R}^n$. From the identity

$$U_{k+1}(t) = X_0 + \int_0^t F(U_k(s))\, ds,$$

we find by taking limits of both sides that

$$X(t) = X_0 + \lim_{k\to\infty} \int_0^t F(U_k(s))\, ds$$

$$= X_0 + \int_0^t \big(\lim_{k\to\infty} F(U_k(s))\big)\, ds$$
$$= X_0 + \int_0^t F(X(s))\, ds.$$

The second equality also follows from the Lemma from analysis. Therefore $X: J \to \mathcal{O}_\rho$ satisfies the integral form of the differential equation and hence is a solution of the equation itself. In particular, it follows that $X: J \to \mathcal{O}_\rho$ is C^1.

This takes care of the existence part of the theorem. Now we turn to the uniqueness part.

Suppose that $X, Y: J \to \mathcal{O}$ are two solutions of the differential equation satisfying $X(0) = Y(0) = X_0$, where, as above, J is the closed interval $[-a, a]$. We will show that $X(t) = Y(t)$ for all $t \in J$. Let

$$Q = \max_{t\in J} |X(t) - Y(t)|.$$

This maximum is attained at some point $t_1 \in J$. Then

$$\begin{aligned} Q = |X(t_1) - Y(t_1)| &= \left| \int_0^{t_1} (X'(s) - Y'(s))\, ds \right| \\ &\le \int_0^{t_1} |F(X(s)) - F(Y(s))|\, ds \\ &\le \int_0^{t_1} K|X(s) - Y(s)|\, ds \\ &\le aKQ. \end{aligned}$$

Since $aK < 1$, this is impossible unless $Q = 0$. Therefore

$$X(t) \equiv Y(t).$$

This completes the proof of the theorem. ■

To summarize this result, we have shown: Given any ball $\mathcal{O}_\rho \subset \mathcal{O}$ of radius ρ about X_0 on which

1. $|F(X)| \le M$;
2. F has Lipschitz constant K; and
3. $0 < a < \min\{\rho/M, 1/K\}$;

there is a unique solution $X: [-a, a] \to \mathcal{O}$ of the differential equation such that $X(0) = X_0$. In particular, this result holds if F is C^1 on $\mathcal{O}$.

Some remarks are in order. First note that two solution curves of $X' = F(X)$ cannot cross if F satisfies the hypotheses of the theorem. This is an immediate consequence of uniqueness but is worth emphasizing geometrically. Suppose $X\colon J \to \mathcal{O}$ and $Y\colon J_1 \to \mathcal{O}$ are two solutions of $X' = F(X)$ for which $X(t_1) = Y(t_2)$. If $t_1 = t_2$ we are done immediately by the theorem. If $t_1 \neq t_2$, then let $Y_1(t) = Y(t_2 - t_1 + t)$. Then Y_1 is also a solution of the system. Since $Y_1(t_1) = Y(t_2) = X(t_1)$, it follows that Y_1 and X agree near t_1 by the uniqueness statement of the theorem, and hence so do $X(t)$ and $Y(t)$.

We emphasize the point that if $Y(t)$ is a solution, then so too is $Y_1(t) = Y(t + t_1)$ for any constant t_1. In particular, if a solution curve $X\colon J \to \mathcal{O}$ of $X' = F(X)$ satisfies $X(t_1) = X(t_1 + w)$ for some t_1 and $w > 0$, then that solution curve must in fact be a periodic solution in the sense that $X(t + w) = X(t)$ for all t.

17.3 Continuous Dependence on Initial Conditions

For the existence and uniqueness theorem to be at all interesting in any physical or even mathematical sense, the result needs to be complemented by the property that the solution $X(t)$ depends continuously on the initial condition $X(0)$. The next theorem gives a precise statement of this property.

Theorem. *Let $\mathcal{O} \subset \mathbb{R}^n$ be open and suppose $F\colon \mathcal{O} \to \mathbb{R}^n$ has Lipschitz constant K. Let $Y(t)$ and $Z(t)$ be solutions of $X' = F(X)$ which remain in $\mathcal{O}$ and are defined on the interval $[t_0, t_1]$. Then, for all $t \in [t_0, t_1]$, we have*

$$|Y(t) - Z(t)| \leq |Y(t_0) - Z(t_0)| \exp(K(t - t_0)). \qquad \blacksquare$$

Note that this result says that, if the solutions $Y(t)$ and $Z(t)$ start out close together, then they remain close together for t near t_0. While these solutions may separate from each other, they do so no faster than exponentially. In particular, we have this corollary:

Corollary. (Continuous Dependence on Initial Conditions) *Let $\phi(t, X)$ be the flow of the system $X' = F(X)$ where F is C^1. Then ϕ is a continuous function of X.* ■

The proof depends on a famous inequality that we prove first.