Chapter 1 defined a dynamical system as a type of mathematical system, \( S = (X, G, U, \phi) \), where \( X \) is a normed linear space, \( G \) is a group, \( U \) is a linear space of input functions defined over the same field as \( X \) and \( \phi : G \times X \times U \to X \) is a continuous map such that for any \( u \in U \) and \( p \in X \), \( \phi(s + t; p, u) = \phi(t, \phi(t; p, u[0,s]), u[0,s+t]) \). In this abstract setting \( X \) is called the state space or phase space and \( G \) represents time. If the input function is fixed to a specific \( u \in U \), then the dynamical system is unforced or homogeneous. If the time group, \( G \) is fixed to the real line, \( \mathbb{R} \), then the system is continuous-time. We will further confine our attention to a state space \( X \subset \mathbb{R}^n \) which subset of Euclidean space.

This chapter confines its attention to homogeneous continuous-time dynamical systems that evolve over a subset, \( X \), of Euclidean space, \( \mathbb{R}^n \). We can therefore simplify our description of the dynamical system to the continuous map \( \phi : \mathbb{R} \times X \to X \) that we call a one parameter group of transition maps of \( X \) back onto itself that satisfies for any \( p \in X \) and \( s, t \in \mathbb{R} \) the following two relations; 1) \( \phi(s + t; p) = \phi(t, \phi(s; p)) \) and 2) \( \phi(0, p) = p \). We refer to \( \phi \) as the dynamical system (since both the phase space, time group, and input are fixed). It will be convenient to introduce two partial maps of \( \phi \). The first partial map, \( \Phi_t : X \to X \), is called the system’s flow and takes values \( \Phi_t(p) = \phi(t; p) \) for any \( p \in X \) and \( t \in \mathbb{R} \). The second partial map \( \xi_p : \mathbb{R} \to X \) is called the system’s trajectory and takes values \( \xi_p(t) = \phi(t; p) \) for any \( p \in X \) and \( t \in \mathbb{R} \).

An ordinary differential equation (ODE) is an equation consisting of the time-derivatives of a variable. ODEs played no role in the preceding description of a dynamical system, but they are nonetheless, very useful in providing concrete representations of the system. An ODE, essentially, provides a local law of state evolution that tells one how a given state, \( p \in X \), changes over an infinitesimal interval of time. The dynamical system, \( \phi \), on the other hand provide a global description of how the state varies both in time and space. In general, the local law embodied by an ODE is easier to construct than the global characterization. This is particularly true for mechanical systems where one can build the system so it adheres to Newton’s laws. But it is usually the case that the system’s global behavior is of greatest interest to us. This chapter investigates how the local ODE representation can be used to reconstruct the past and predict the future behavior of the dynamical system, \( \phi \).

1. Vector Fields for Dynamical Systems

This section links the dynamical system \( \phi \) to a differential equation representation. That linkage will be through a mathematical object known as a vector field.
Let $\phi$ be a homogeneous real-time dynamical system (one parameter group of transformations) that evolves over a set $X$ in Euclidean $n$-space. The trajectories of $\phi$ are unique in the following sense.

**Theorem 24.** For the dynamical system, $\phi$, and for any $p \in X$, the trajectory $\xi_p$ is unique.

**Proof:** Assume that this is not the case, then there are two states $q_1$ and $q_2$ that are not equal to each other and a single state $p$ such that $\phi(t, p) = q_1 \neq q_2 = \phi(t, p)$. By the group property of $\phi$ we know that $\phi(-t, q_1) = p = \phi(-t, q_2)$. If we apply $\phi$ to both sides then

$$\phi(t, \phi(-t, q_1)) = q_1 = \phi(t, \phi(-t, q_2)) = q_2$$

which contradicts the assumption that $q_1 \neq q_2$. ◻

It will be convenient for any $p \in X$ to define a set

$$\Omega_p = \{y \in X : y = \phi(t; p) \text{ for any } t \in \mathbb{R}\}$$

that we call the orbit of $p$. We can introduce a binary relation $\sim$ such that for any $p, q \in X$, that $p \sim q$ if and only if $q \in \Omega_p$. It can be shown that $\sim$ is an equivalence relation and so $\sim$ partitions $X$ into equivalence classes. The equivalence class containing a state $p \in X$ is simply the orbit $\Omega_p$ for that state.

An *equilibrium* or *fixed point* $p^* \in X$ of the dynamical system $\phi$ is a state which is itself an orbit. In other words $\Omega_{p^*} = \{p^*\}$. Fixed points play an important role in how local representations of the flow can be extended to a global characterization. They will also play a critical role in defining what we mean by system stability, which plays a crucial role in the regulation of a dynamical system.

Consider a mapping $f : X \rightarrow Y$ from $X \subset \mathbb{R}^n$ into $Y \subset \mathbb{R}^m$. This mapping is said to be differentiable if each of its component functions $f_i : X \rightarrow \mathbb{R}$ is a continuously differentiable function for $i = 1, 2, \ldots, n$. We say a mapping $f : X \rightarrow Y$ is a *diffeomorphism* if both $f$ and $f^{-1} : Y \rightarrow X$ are differentiable mappings. The dynamical system is said to be smooth (or differentiable) if $\phi$ is a differentiable mapping.

Now consider a smooth dynamical system, $\phi$ and define the *phase velocity* $f : X \rightarrow X$ of the flow $\Phi_t$ at a point $p \in X$ as the vector

$$f(p) \equiv \frac{d}{dt} \bigg|_{t=0} \Phi_t(p)$$

Let $\xi_{x_0}$ be the trajectory of the system from initial state $x_0 \in X$ and let $x_i(t)$ denote the $i$th component of $\xi_{x_0}(t)$ at time instant $t$. Let $x(t) \in X$ denote the vector whose elements are $x_i(t)$ for $i = 1, 2, \ldots, n$. Then the components of the phase velocity, $f(x(t))$, may be written as

$$f_i(x(t)) = \frac{dx_i(t)}{dt} \equiv \dot{x}_i(t)$$

This equation presumes that $x(0) = x_0$, so we can rewrite the system equations in vector form while explicitly indicating the initial condition,

$$\dot{x}(t) = f(x(t)), \quad x(0) = x_0$$
Note that $f$ may be seen as assigning to each vector $x \in X$, another vector $f(x) \in \mathbb{R}^n$ thereby creating a field of vectors in the state space. So the function $f$ is called a vector field. What we’ve just shown is that starting from a smooth dynamical system, $\phi$, with flows $\{\Phi_t\}$, that the system trajectories $\xi_{x_0}$ from an initial $x_0 \in X$ must satisfy the differential equation (33) with the associated initial value, $x_0$. We call equation (33) an initial value problem (IVP). We’ve therefore demonstrated that trajectories of a smooth dynamical system can be represented locally as IVPs, which is summarized in the following theorem.

**Theorem 25.** Given a smooth dynamical system $\phi$ with flows $\{\Phi_t\}$, then the system trajectories $\xi_p : \mathbb{R} \to X$ satisfy the IVP, $\dot{x}(t) = f(x(t))$ with $x(0) = x_0$ in which $f(p) = \frac{d}{dt}|_{t=0} \Phi_t(p)$ for any $p \in X$.

We have just demonstrated that every smooth (differentiable) dynamical system defines a vector field and therefore the trajectories of this system satisfy an IVP. In general, however, we do not start from the dynamical system, $\phi$, and derive a vector field. In many physical systems, the vector fields are determined by the laws of physics governing the physical objects we wish to control. This means that one usually determines the vector field first, based on a mechanistic understanding of how that system functions. The more meaningful question for us is whether or not the vector field constructed on first principles admits a dynamical system, $\phi$. This is not obvious at first glance for it is relatively easy to construct perfectly reasonable IVP’s for which continuously differentiable trajectories may not exist globally. Even if they do exist, these trajectories may fail to be unique thereby contradicting theorem 24.

As an example of a differential equation for which a solution may not exist, let us consider an IVP of the following form,

$$\dot{x}(t) = \begin{cases} 
1 & x(t) \leq 0 \\
-1 & x(t) > 0
\end{cases}, \quad x(0) = 0$$

At the initial time, $x(0)$ is zero and so $\dot{x}(0) = 1$. So an infinitesimal time after 0 we find $x(\varepsilon) > 0$ which means that $\dot{x}(\varepsilon) = -1$. This would immediately force $x$ to go back to zero again, but as soon as it does $\dot{x}$ shifts back to being positive. In other words, this differential equation appears to force the system to chatter back and forth between being slightly positive and zero. Now, one might assert that this is not a reasonable differential equation since its right hand side is not continuous. There are however many real life systems where this type of switching action actually occurs. In particular, the spacecraft example we presented in chapter 1 was one such system and so even if a smooth solution does not exist for this system, it may still be interesting in its own right.

As another example, let us consider the IVP

$$\dot{x}(t) = -x^2(t), \quad x(0) = -1$$

Again this is a perfectly reasonable ODE that may fit some physical process. This is a separable equation, so we can rewrite it as

$$t = -\int_{x_0}^{x} \frac{dx}{x^2} = \frac{1}{x} - \frac{1}{x_0} = \frac{1}{x} + 1$$
which implies that for $t > 0$ that

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

This trajectory only exists over the time interval $[0, 1)$ and so it fails to generate a smooth dynamical system $\phi$, since we define $\phi$ over all time.

The last example we’ll consider is the IVP,

$$\dot{x}(t) = x^{1/3}, \quad x(0) = 0$$

This IVP has two continuously differentiable solutions. The trivial trajectory $x(t) = 0$ satisfies the ODE and the function $x(t) = (\frac{2t}{3})^{3/2}$ also satisfies the ODE. This is problematic for us as well since we already know that smooth dynamical systems, $\phi$, must generate unique trajectories by theorem 24.

The preceding examples demonstrated that the use of first principle modeling of physical processes may give rise to differential equations that do not admit smooth dynamical systems. In the following sections, we want to derive conditions under which the trajectories generated by an ODE system exist and are unique.

The mapping $f : \mathbb{R}^n \to \mathbb{R}^n$ was called a vector field because it assigns to each vector $x$ in $\mathbb{R}^n$ another vector $f(x)$ that is also in the vector space $\mathbb{R}^n$. Recall that the dual space, $(\mathbb{R}^n)^*$, of $\mathbb{R}^n$ is the set of all linear real valued functions defined on $\mathbb{R}^n$. This dual space is also an $n$-dimensional vector space whose elements are called covectors. While vectors in $\mathbb{R}^n$ are denoted in column form, it is customary to represent covectors in $(\mathbb{R}^n)^*$ as "row" vectors.

Now suppose that $\omega_1, \ldots, \omega_n$ are smooth real-valued functions of the real variables $x_1, \ldots, x_n$ and consider the row vector

$$\omega(x) = [\omega_1(x), \omega_2(x), \ldots, \omega_n(x)]$$

We view $\omega$ as a map assigning to each vector $x \in \mathbb{R}^n$ an element $\omega(x)$ of the dual space $(\mathbb{R}^n)^*$.

A covector field of special importance is the so-called differential or gradient of a real-valued function $V : \mathbb{R}^n \to \mathbb{R}$. This covector field is denoted as $dV$ or $\frac{\partial V}{\partial x}$ and is defined as the $1 \times n$ row vector whose $i$th element is the partial derivative of $\lambda$ with respect to $x_i$

$$dV(x) = \left[ \frac{\partial V(x)}{\partial x_i} \right] = \left[ \frac{\partial V(x)}{\partial x_1}, \frac{\partial V(x)}{\partial x_2}, \ldots, \frac{\partial V(x)}{\partial x_n} \right]$$

It is often more customary to use the following notation for the differential

$$dV(x) = \frac{\partial V(x)}{\partial x}$$

We define the derivative of $V$ along $f$ (also known as the directional derivative) as

$$\langle dV(x), f(x) \rangle = \frac{\partial V}{\partial x} f(x) = \sum_{i=1}^{n} \frac{\partial V}{\partial x_i} f_i(x)$$
This derivative is also denoted as $L_f V$ or $D_f V$. Directional derivatives play an important role in expressing relationships certifying the “stability” of a system or its equilibria. We will make use of these stability certificates starting in chapter 5.

We may also apply this operation repeatedly which motivates the following recursive definition for the iterated directional derivative,

$$L_k^f V(x) = \frac{\partial (L_{k-1}^f V(x))}{\partial x} f(x)$$

where $L_0^f V(x) = V(x)$. Such iterated directional derivatives will be useful to us when we consider the geometric theory of nonlinear control discussed in chapter 9.

2. Existence of Solutions to IVPs

As discussed above, it is quite possible to introduce differential equations for dynamical systems with solutions that cannot possibly be generated by a dynamical system $(X, \phi)$. We now turn to establish those conditions under which ODE models for dynamical systems actually do generate orbits of a smooth dynamical system $(X, \phi)$. We first consider the question of the existence of solutions.

Consider the initial value problem

$$\dot{x} = f(t, x), \quad x(t_0) = x_0 \tag{34}$$

where $f \in C(U, \mathbb{R}^n)$ and $U$ is an open subset of $\mathbb{R}^{n+1}$ with $(t_0, x_0) \in U$. We first note that integrating both sides of this equation with respect to $t$ shows that the IVP is equivalent to the following integral equation

$$x(t) = x_0 + \int_0^t f(s, x(s))ds \tag{35}$$

Let $x : \mathbb{R} \to X$ be a $C^1$ solution to the IVP in the sense that it satisfies the above integral equation and note that

$$x_h = x_0 + \dot{x}(0)h + o(h) = x_0 + f(0, x_0)h + o(h)$$

where $\frac{o(h)}{h} \to 0$ as $h \to 0$. We think of $x_h$ as a first order approximation to the solution at time instant $h$. This suggests that an approximate solution to the IVP for all time instants $h, 2h, \ldots, mh, \ldots$ might be obtained through the following recursive procedure

$$x_h(t_{m+1}) = x_h(t_m) + f(t_m, x_h(t_m))h, \quad t_m = mh$$

This procedure is known as Euler’s method.

The question we want to address is whether $x_h(mh)$ asymptotically approaches a function $x(t)$ as $h \downarrow 0$ and if so does that function satisfy the integral equation (35) which means $x$ is a $C^1$ solution of the IVP. To establish conditions on $f$ that assure the existence of this “limiting” solution, we will introduce an infinite set
The set \( \{x_m\}_{m \in \mathbb{N}} \) of functions \( x_m : \mathbb{R} \rightarrow \mathbb{R}^n \) that are equicontinuous. This means that for all \( \epsilon > 0 \) there is a \( \delta > 0 \) (independent of \( m \)) such that

\[
|t - s| \leq \delta \Rightarrow |x_m(t) - x_m(s)| \leq \epsilon,
\]

for all \( m \in \mathbb{N} \). The key thing is that the \( \delta \) is independent of \( m \), so equicontinuity may be seen as an extension of uniform continuity for single functions to a collection of functions, \( \{x_m\} \). The following theorem provides the main result that we will need to establish whether or not the sequence of functions generated by Euler’s method indeed converge to the true solution of the system. We will not prove this theorem as its proof may be found in elementary texts on mathematical analysis such as [Rud64], but we do state it formally below.

**Theorem 26.** (Arzelà-Ascoli) Consider the sequence of functions, \( \{x_m(t)\}_{m=1}^\infty \), is in \( C(I, \mathbb{R}^n) \) where \( I \) is a compact interval. Assume \( \{x_m(t)\} \) is equicontinuous. If the sequence \( \{x_m\} \) is also bounded, then it has a uniformly convergent subsequence.

With the help of the Arzelà-Ascoli’s theorem it then becomes possible to establish Peano’s theorem which simply requires \( f \) to be continuous for the IVP to have a local \( C^1 \) solution. The key idea in proving Peano’s theorem is that since \( f \) is continuous it must be bounded by a constant on any compact interval and this allows us to bound \( x_h(t) \) in a uniform manner that is independent of \( h \). In other words \( x_h(t) \) forms an equicontinuous family of functions that by the Arzelà-Ascoli theorem has a subsequence that converges uniformly to a function that is indeed the solution to the IVP. The theorem is stated below, followed by its proof.

**Theorem 27.** (Peano) Suppose \( f \) is continuous on \( V = [t_0, t_0 + T] \times \overline{N_\delta(x_0)} \) and denote the maximum of \( |f| \) on \( V \) as \( M \). Then there exists at least one solution of the IVP for \( t \in [t_0, t_0 + T_0] \) which remains in \( \overline{N_\delta(x_0)} \) where \( T_0 = \min\{T, \delta / M\} \). The analogous assertion holds for the interval \( [t_0 - T_0, t_0] \).

**Proof:** Pick \( \delta, T > 0 \) such that \( V = [t_0, t_0 + T] \times \overline{N_\delta(x_0)} \subset U = \mathbb{R} \times \mathbb{R}^n \). Define the set of functions \( \{x_h\} \) by

\[
x_h(t) = x_0 + \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \chi(s)f(t_j, x_h(t_j))ds
\]

where \( \chi(s) = 1 \) for \( s \in [t_0, t] \) and zero elsewhere. Since \( f \) is continuous and \( V \) is compact we know \( f \) attains its maximum on \( V \)

\[
M = \max_{(t,x) \in V} |f(t,x)|
\]

From the definition of \( x_h(t) \), we can therefore conclude that for any \( s, t \in [t_0, t_0 + T_0] \) with \( T_0 = \min\{T, \frac{\delta}{M}\} \) that

\[
|x_h(t) - x_h(s)| \leq M|t - s|
\]

Which means that \( \{x_h\} \) is equicontinuous and bounded and so we can invoke the the Arzelà-Ascoli theorem to assert that there exists a uniformly convergent subsequence \( x_{h_i}(t) \rightarrow x(t) \) as \( h_i \rightarrow 0 \). It remains to show that this limit \( x(t) \) solves the IVP. In particular, we use the integral form of the IVP to establish this result.
Let $\Delta(h)$ be a function where

$$|f(t, y) - f(t, x)| \leq \Delta(h) \text{ for } |y - x| \leq Mh \text{ and } |s - t| \leq h$$

Since $f$ is uniformly continuous on $V$ we can always find a sequence of times $\{h_i\}$ such that $h_i \to 0$ and $\Delta(h_i) \to 0$. To estimate the difference between the right and left side of our integral equation for a given $x_h(t)$, we choose an $m$ with $t \leq t_m$ and using the integral expression for $x_h(t)$ we see that

$$|x_h(t) - x_0 - \int_{t_0}^{t} f(s, x_h(s))ds| \leq \Delta(h) \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \chi(s) |f(t_j, x_h(t_j)) - f(s, x_h(s))| ds$$

$$\leq \Delta(h) \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \chi(s) ds = |t - t_0| \Delta(h).$$

If we look at

$$x(t) = \lim_{h \to 0} x_h(t)$$

$$= x_0 + \lim_{h \to 0} \int_{t_0}^{t} f(s, x_h(s))ds$$

$$= x_0 + \int_{0}^{t} f(s, x(s))ds$$

the last equality occurs because uniform convergence implies we can interchange the limit with the integral.

3. Uniqueness of Local Solutions

In an earlier section we examined the IVP, $\dot{x}(t) = x^{1/3}(t)$ with initial condition $x(0) = 0$. The right hand side of this IVP is clearly continuous and a solution exists globally. The problem we saw with this system, however, was that the solution was not unique. We will need to introduce a stronger notion of continuity to ensure that the solutions of the IVP are unique.

Consider a function $f : D \to \mathbb{R}^n$ where $D \subset \mathbb{R}^n$. This function is said to be locally Lipschitz at $x$ on $D$ if for each point $x \in D$ there exists a neighborhood, $D_0$, of $x$ for which there exists a real constant $L \geq 0$ such that

$$|f(x) - f(y)| \leq L|x - y|$$

for all $y \in D_0$. This essentially means that around $(x, f(x))$, the graph of $f$ can be enclosed in a conic sector defined by linear inequalities whose slopes are the Lipschitz constant $L$. If the Lipschitz property holds uniformly over $D$ then we simply say that $f$ is Lipschitz on $D$. The following theorem provides a way to estimate the Lipschitz constant, $L$. This theorem requires the domain $D \subset \mathbb{R}^n$ to be convex. A set, $D$, is said to be convex if for any $x, y \in D$ we can establish that $sx + (1 - s)y \in D$ for all $s \in [0, 1]$. 
THEOREM 28. Let \( f : [a, b] \times D \to \mathbb{R}^n \) be continuous on \( D \). Suppose the Jacobian matrix \( \frac{\partial f}{\partial x} \) exists and is continuous on \([a, b] \times D\). Let \( W \subset D \) be convex and let \( L \geq 0 \) such that

\[
\left\| \frac{\partial f}{\partial x} \right\| \leq L
\]

over \([a, b] \times W\). Then

\[
|f(t, x) - f(t, y)| \leq L|x - y|
\]

for all \( t \in [a, b], x \in W, \) and \( y \in W \).

Proof: Consider a line segment drawn between any two points in \( W \). We know this segment lies within \( W \) because \( W \) is convex. We may therefore represent any point on this segment as

\[
\gamma(s) = (1 - s)x + sy
\]

where \( s \in [0, 1] \) and \( x, y \in W \). Now let

\[
g(s) = z^T f(t, \gamma(s))
\]

where \( |z| = 1 \) and

\[
z^T [f(t, y) - f(t, x)] = |f(t, y) - f(t, x)|
\]

The function \( g(s) \) is real valued and \( C^1 \) so by the mean value theorem \([\text{TM55}]\), there exists an \( s_1 \in [0, 1] \) such that

\[
g(1) - g(0) = \frac{dg}{dx}(s_1) \Rightarrow z^T [f(t, y) - f(t, x)] = z^T \left[ \frac{\partial f}{\partial x} \right] (y - x)
\]

By the choice of \( z \) we also know that

\[
z^T [f(t, y) - f(t, x)] = |f(t, y) - f(t, x)| = |z^T \left[ \frac{\partial f}{\partial x} \right] (y - x)|
\]

\[
\leq |z| \left| \left[ \frac{\partial f}{\partial x} \right] (y - x) \right| \leq L|y - x|
\]

which completes the proof. \( \diamond \)

The Lipschitz condition is “stronger” than continuity in the sense that every Lipschitz function is continuous, though the converse may not be true. If we return to our earlier example where \( f(x) = \frac{x}{3} \) it is apparent that this function is continuous. But it is not locally Lipschitz at zero since \( \frac{df}{dx} = \frac{1}{3}x^{-2/3} \) which goes to infinity as \( x \) goes to zero. From the above theorem \( \frac{df}{dx} \) is a lower bound on the Lipschitz constant \( L \) and so we can conclude \( f \) is not Lipschitz at zero. What we will now establish is if \( f \) is Lipschitz, then it has a unique local solution.

To prove this result, we will make use of a something known as the contraction mapping principle. In particular, consider a map \( G : X \to X \) on a normed linear space \( X \). We say \( G \) is a contraction mapping if and only if for any \( x, y \in X \), there exists \( 0 \leq \gamma < 1 \) such that

\[
\|G[x] - G[y]\| \leq \gamma \|x - y\|
\]
Essentially, if $G$ is a contraction mapping it takes a pair of vector $x, y$ in $X$ and maps it onto another pair $G[x], G[y]$ of vectors in $X$ such that the distance between the vectors “contracts”. If we can also assert that $X$ is a Banach space, then we can prove there exists a unique element $x^* \in X$ that is a fixed point of the mapping $G$. In other words, there is a unique $x^* \in X$ such that $x^* = G[x^*]$. This is known as the contraction mapping principle, which is stated and proven below.

**Theorem 29. (Contraction Mapping Principle)** Let $X$ be a Banach space, let $S \subseteq X$ and let $G : S \to S$ be a contraction mapping, then there exists a unique element $x^* \in X$ such that $x^* = G[x^*]$.

**Proof:** Select an arbitrary $x_1 \in S$ and define a sequence $\{x_k\}$ by the recursive equation

$$x_{k+1} = G[x_k]$$

Note that

$$
\|x_{k+1} - x_k\| = \|G[x_k] - G[x_{k-1}]\| \\
\leq \gamma \|x_k - x_{k-1}\| \\
\leq \gamma^2 \|x_{k-1} - x_{k-2}\| \\
\leq \cdots \\
\leq \gamma^{k-1} \|x_2 - x_1\|
$$

It therefore follows that

$$
\|x_{k+r} - x_k\| \leq \|x_{k+r} - x_{k+r-1}\| + \|x_{k+r-1} - x_{k+r-2}\| + \cdots + \|x_{k+1} - x_k\| \\
\leq \left(\gamma^{k+r-2} + \gamma^{k+r-3} + \cdots + \gamma^{k-1}\right) \|x_2 - x_1\| \\
= \frac{\gamma^{k-1}}{1 - \gamma} \|x_2 - x_1\|
$$

As $k \to \infty$, one can clearly see that $\|x_{k+r} - x_k\| \to 0$, which implies that $\{x_k\}$ is Cauchy. Since $X$ is a Banach space, $\{x_k\}$ must therefore be convergent to a point $x^* \in X$.

The limit point, $x^*$, must also be a fixed point of $G$. This can be seen by noting that

$$
\|x^* - G[x^*]\| \leq \|x^* - x_k\| + \|x_k + G[x^*]\| \\
\leq \|x^* - x_k\| + \gamma \|x_{k-1} - x^*\|
$$

If $k$ is large enough, then the right hand side above can be made arbitrarily small so that $\|x^* - G[x^*]\| = 0$. Since $\|\cdot\|$ is a norm this implies $x^* - G[x^*] = 0$ which establishes that $x^*$ is a fixed point.

To show that the limit point is unique, let us assume that this is not that case. So there are two limit points $x^*$ and $y^*$ and since both are fixed points we know that

$$
\|x^* - y^*\| = \|G[x^*] - G[y^*]\| \leq \gamma \|x^* - y^*\|
$$
where the inequality follows from the fact that \( G \) is a contraction mapping and where \( 0 \leq \gamma < 1 \). We can rearrange the above inequality to see that
\[
(1 - \gamma)\|x^* - y^*\| \leq 0
\]
But because \( 0 \leq \gamma < 1 \), we know \( 1 - \gamma > 0 \) and so this means the above inequality holds if and only if \( \|x^* - y^*\| = 0 \). Again since \( \| \cdot \| \) is a norm for \( X \), this implies \( x^* = y^* = 0 \) which establishes the uniqueness of the fixed point.

To prove that the IVP has a unique solution, we again turn to the integral form of the equation.
\[
x(t) = x_0 + \int_0^t f(s, x(s))\,ds
\]
What we do now, however, is we think of the right hand side of this inequality as an operator mapping a function \( x : \mathbb{R} \to X \) onto another function. We will show that this operator is a contraction mapping when \( f \) is Lipschitz and so we can apply the contraction mapping principle to infer that it has a unique point whose fixed point, \( x^* \), is the function satisfying the above integral equation. The formal statement of this theorem and its proof are given below.

**Theorem 30.** (Local Uniqueness) Let \( f \) be a continuous map from a connected open set \( D \subset \mathbb{R}^n \) into \( \mathbb{R} \). If \( f \) is Lipschitz on \( D \), then there exists a nonzero \( T > 0 \) such that the IVP has at most one \( C^1 \) solution on the interval \([0, T)\).

**Proof:** Let \( X \) be the space of all continuous functions on \([0, T]\) with \( L_\infty \) norm \( \|x\|_{L_\infty} = \max_{t} |x(t)| \).
Consider a subset \( S \subset X \) such that
\[
S = \{ x \in X : \|x - x_0\|_{L_\infty} \leq r \}
\]
where \( r \) is positive and real. Now consider an operator, \( G : X \to X \), acting on elements of \( X \) that satisfy
\[
G[x](t) - x_0 = \int_0^t f(s, x(s))\,ds
\]
for all \( t \in [0, T] \). By continuity of \( f \) and the fact that \([0, T]\) is compact (closed and bounded) we know \( f \) attains its maximum on this set and so there exists a real number \( h \) such that
\[
h = \max_{t \in [0, T]} |f(t, x_0)|
\]
Now consider \( G[x](t) - x_0 \) and let us take it Euclidean norm,
\[
\|G[x](t) - x_0\| = \left\| \int_0^t f(s, x(s))\,ds \right\| = \left\| \int_0^t (f(s, x(s)) - f(s, x_0) + f(s, x_0))\,ds \right\|
\leq \int_0^t (|f(s, x(s)) - f(s, x_0)| + |f(s, x_0)|)\,ds
\]
Since \( f \) is Lipschitz and since \( |f| \) is bounded above by \( h \) on \([0, T]\), we can bound the above inequality as
\[
\|G[x](t) - x_0\| \leq (L|x(s) - x_0| + h)\,ds = (t - 0)(Lr + h) \\
\leq T(Lr + h)
\]
Let us choose $T < \frac{r}{Lr + h}$ so that $P$ maps the set $S$ back into itself. So if we restrict the interval of time $[0, T]$ to meet this constraint then we know $P$ takes a function in the small ball and maps it back onto that ball.

We now consider conditions under which this restricted $P$ is also a contraction mapping. Let $x$ and $y$ be elements of $S$ and consider

$$|G[x](t) - G[y](t)| = \left| \int_0^t (f(s, x(s)) - f(s, y(s))) \, ds \right|$$

$$\leq \int_0^t |f(s, x(s)) - f(s, y(s))| \, ds$$

$$\leq \int_0^t L|x(s) - y(s)| \, ds$$

$$\leq \int_0^T L\|x - y\|_{L_\infty} \, ds = LT\|x - y\|_{L_\infty}$$

Since this must hold for all $t$, it also holds for the maximum so that

$$\|G[x] - G[y]\|_{L_\infty} \leq LT\|x - y\|_{L_\infty}$$

So if we select $T < 1/L$, then $G$ is a contraction mapping in the $L_\infty$ signal space.

So we’ve shown that the interval of existence of the solution should satisfy

$$T < \min \left\{ \frac{1}{L}, \frac{r}{Lr + h} \right\}$$

Since $G$ is an automorphism and a contraction mapping, we can use the contraction mapping principle to conclude there is a unique $x^* \in L_\infty$ that is a fixed point for the operator $G$. In particular this means that

$$x^*(t) - x_0 = \int_0^t f(s, x^*(s)) \, ds$$

for all $t \in [0, T]$ which establishes the integral form of the IVP. \(\diamondsuit\)

### 4. Extension to Global Solutions

Theorem 30 only establishes uniqueness for a solution over a finite interval of time. This interval may be very small. One may try to extend the interval of existence by repeated application of theorem 30. This involves taking the initial condition $x_0$, extending the solution over an interval $T_1$, and then taking $x(T_1)$ and applying the local theorem again to generate a longer interval of existence $T_2$. We can continue doing this indefinitely to generate a sequence of intervals $\{T_i\}$. The problem, with this approach however is that this sequence may be convergent to a finite number thereby limiting the interval of existence.

This section investigates conditions under which this “local” solution in a neighborhood of the initial condition can be extended in a global manner. One approach is to strengthen the local Lipschitz condition in theorem 30 to a global Lipschitz condition. Since the Lipschitz condition is global, the estimate we obtain for the interval existence is uniformly bounded in a manner that prevents $\{T_i\}$ from being convergent to a finite value.
Theorem 31. (Global Uniqueness - Lipschitz) Consider an IVP in which \( f(t, x) \) is piecewise continuous in \( t \) and is globally Lipschitz in \( x \) over the time interval \([0, T_0]\), then the IVP has a unique solution over \([0, T_0]\).

**Proof:** The proof follows that of the local version in theorem 30. The thing that keeps our solution local is found in equation (36) in which the interval of existence \( T \) is dependent on the initial condition through our maximum bound on \(|f|\) over the interval \([0, T]\). Since the Lipschitz condition is global, we can now remove this dependence on \( x_0 \) and it becomes possible to make \( r \) arbitrarily large. This allows us to modify the condition on the interval of existence to \( T = 1/L \). If this is smaller than \( T_0 \), then we can subdivide the interval \( T \) into smaller intervals and then use theorem 30 to obtain uniqueness over each of these subintervals.

Theorem 31, however, is very restrictive. If one considers, for example, the vector field, \( f(x) = x^3 \), then it is rather easy to see that this function is not globally Lipschitz and yet one can easily demonstrate that the IVP has a unique global solution for any initial condition. To relax the global Lipschitz condition, we need to impose a compactness condition.

**Theorem 32.** Let \( f(t, x) \) be piecewise continuous in \( t \) and locally Lipschitz in \( x \) for all \( t \geq 0 \) and all \( x \) in a domain \( D \subset \mathbb{R}^n \). Let \( W \) be a compact subset of \( D \) with \( x_0 \in W \) and suppose it is known that every solution of the IVP lies entirely in \( W \), then there is a unique solution defined for all \( t \geq 0 \).

**Proof:** Recall from our earlier discussion, we know that the local theorem only ensures existence up to a finite time interval \( T \). But if this is the case, then the solution must leave any compact subset of \( D \) since the state \( x(T) \) is a limit point of the trajectory and the closed nature of the compact set (Heine-Borel) requires that \( W \) contain its limit points. This would contradict the assumption that the trajectory remains in a compact set and so we have a contradiction to the assertion that \( T \) is finite.

Let us return to our previous example where \( \dot{x} = -x^3 \). In that case \( f(x) \) is locally Lipschitz about the origin, and so we will only be able to guarantee that the solution is local. However, if we start with a positive initial condition, then \( \dot{x} \) is negative and similarly if \( x(t) \) is negative, then \( \dot{x} \) is positive. So if \( x(0) = a \), then the solution cannot leave the closed and bounded (and hence compact) set \( W = \{ x \in \mathbb{R} : |x| \leq |a| \} \). So we can use theorem 32 to infer there exists a unique solution to this IVP.

5. Solution Sensitivity to IVP Data

In many applications the data used to define the IVP may be approximately known. This data includes the initial condition, \( x_0 \), and any parameters characterizing the right hand side of the ODE. It is not only critical that the system’s behavior “exist” and be “unique”, but it is also critical that this behavior is relatively insensitive to small changes in that data. This section examines how sensitive the solutions of an ODE might be to variations in the system’s initial conditions. The main tool we will use to examine this is the *Gronwall-Bellman inequality.*
THEOREM 33. (Gronwall-Bellman) Let $\lambda : [a, b] \to \mathbb{R}$ be continuous and let $\mu : [a, b] \to \mathbb{R}$ be continuous and non-negative. If a continuous function $y : [a, b] \to \mathbb{R}$ satisfies
\begin{equation}
y(t) \leq \lambda(t) + \int_a^t \mu(s)y(s)\,ds
\end{equation}
for $t \in [a, b]$, then on the same interval
\begin{equation}
y(t) \leq \lambda(t) + \int_0^t \lambda(s)\mu(s)e^\int_0^s \mu(\tau)d\tau\,ds
\end{equation}

**Proof:** Let $\phi(t) = \exp \left( -\int_0^t \lambda(s)ds \right)$. Then one can compute
\[
\frac{d\phi(t)}{dt} \int_0^t \lambda(s)y(s)\,ds = \lambda(t)\phi(t) \left( y(t) - \int_0^t \lambda(s)y(s)\,ds \right) \leq \mu(t)\lambda(t)\phi(t)
\]
from the assumption in equation (37). Integrating this inequality with respect to $t$ and dividing by $\phi(t)$ yields,
\[
\int_0^t \lambda(s)y(s)\,ds \leq \int_0^t \mu(s)\lambda(s)\frac{\phi(s)}{\phi(t)}\,ds
\]
Adding $\mu(t)$ to both sides, one can then use equation (37) to obtain the conclusion in equation (38). $\diamond$

As an example of the use of the Gronwall-Bellman inequality, let us recall from the Peano theorem (theorem 27) that there exists a solution to the IVP $\dot{x} = f(t, x)$ with $x(0) = x_0$. Let us assume that $f$ is locally Lipschitz with Lipschitz constant $L$, but that this solution is not unique. In other words we assume there are two solutions, $x_1 : \mathbb{R} \to \mathbb{R}$ and $x_2 : \mathbb{R} \to \mathbb{R}$ that must satisfy the integral equation
\[
x(t) = x_0 + \int_0^t f(s, x(s))\,ds
\]
The Euclidean norm of the difference between these two solutions at time $t \in \mathbb{R}$ is
\[
|x_1(t) - x_2(t)| \leq \left| \int_0^t [f(s, x_1(s)) - f(s, x_2(s))]\,ds \right| \\
\leq \int_0^t |f(s, x_1(s)) - f(s, x_2(s))|\,ds \\
\leq \int_0^t L|x_1(s) - x_2(s)|\,ds
\]
The hypothesis of the Gronwall-Bellman inequality is satisfied if we let $\lambda(t) = 0, \mu(t) = L$, and $y(t) = |x_1(t) - x_2(t)|$. Applying the Gronwall-Bellman theorem (33) then implies that
\[
y(t) = |x_1(t) - x_2(t)| \leq 0
\]
for all $t \in \mathbb{R}$. This is sufficient to ensure that $x_1(t) = x_2(t)$ for all $t$ in the interval of existence and we’ve just reproven the local uniqueness theorem 30 using a non-constructive approach that does not rely on the Contraction Mapping Principle.
The Gronwall-Bellman inequality can be used to bound how close solutions of an additively perturbed IVP are to each other. This result will then be used to study the dependence of these solutions on the IVP’s parameters and initial conditions.

**THEOREM 34.** Let \( f(t, x) \) be piecewise continuous in \( t \) and Lipschitz in \( x \) on \([t_0, T] \times W \) with Lipschitz constant \( L \), where \( W \subset \mathbb{R}^n \) is an open connected set. Let \( y : \mathbb{R} \to \mathbb{R}^n \) and \( z : \mathbb{R} \to \mathbb{R}^n \) be solutions of

\[
\frac{d}{dt} y = f(t, y), \quad y(t_0) = y_0
\]

and

\[
\frac{d}{dt} z = f(t, z) + g(t, z), \quad z(t_0) = z_0
\]

respectively with \( y(t), z(t) \in W \) for all \( t \in [t_0, T] \). Suppose there exists \( \mu > 0 \) such that

\[
|g(t, x)| \leq \mu \text{ for all } (t, x) \in [t_0, T] \times W
\]

and suppose \( |y_0 - z_0| \leq \gamma \). Then for all \( t \in [t_0, T] \),

\[
|y(t) - z(t)| \leq \gamma e^{L(t-t_0)} + \frac{\mu}{L} \left( e^{L(t-t_0)} - 1 \right)
\]

**Proof:** The solutions \( y \) and \( z \) satisfy for all \( t \in [t_0, T] \),

\[
y(t) = y_0 + \int_{t_0}^{t} f(s, y(s)) \, ds
\]

\[
z(t) = z_0 + \int_{t_0}^{t} \left( f(s, z(s)) + g(s, z(s)) \right) \, ds
\]

Subtracting the two equations and taking the Euclidean norm yields,

\[
|y(t) - z(t)| \leq |y_0 - z_0| + \int_{t_0}^{t} \left| f(s, y(s)) - f(s, z(s)) \right| \, ds + \int_{t_0}^{t} |g(s, z(s))| \, ds
\]

\[
\leq \gamma + \mu(t-t_0) + \int_{t_0}^{t} L |y(s) - z(s)| \, ds
\]

Applying the Gronwall-Bellman inequality to the function \( |y(t) - z(t)| \) yields,

\[
|y(t) - z(t)| \leq \gamma + \mu(t-t_0) + \int_{t_0}^{t} L (\gamma + \mu(s-t_0)) e^{L(t-s)} \, ds
\]

Integrating the right hand side by parts yields,

\[
|y(t) - z(t)| \leq \gamma + \mu(t-t_0) - \gamma - \mu(t-t_0) + \gamma e^{L(t-t_0)} + \int_{t_0}^{t} \mu e^{L(t-s)} \, ds
\]

\[
= \gamma e^{L(t-t_0)} + \frac{\mu}{L} \left( e^{L(t-t_0)} - 1 \right)
\]

which completes the proof. \( \diamondsuit \)

We now use the preceding theorem 34 to establish the continuity of solutions with respect to variations in initial states and parameters. In this case, we consider a perturbed version of the IVP whose ODE is

\[
\dot{x}(t) = f(t, x, \lambda)
\]

with \( \lambda \) being a real parameter.
THEOREM 35. Let \( f(t, x, \lambda) \) be continuous in \((t, x, \lambda)\) and locally Lipschitz in \(x\) (uniformly in \(t\) and \(\lambda\)) on \([t_0, T] \times D \times \{|\lambda - \lambda_0| \leq c\}\) where \(D \subset \mathbb{R}^n\) is an open connected set. Let \(y(t; \lambda_0)\) be a solution of \(\dot{x}(t) = f(t, x, \lambda_0)\) with \(y(t_0, \lambda_0) = y_0 \in D\). Suppose \(y(t, \lambda_0)\) is defined and belongs in \(D\) for all \(t \in [t_0, T]\).

Then given \(\epsilon > 0\) there is a \(\delta > 0\) such that if \(|z_0 - y_0| < \delta\) and \(|\lambda - \lambda_0| < \delta\), then there is a unique solution \(z(t, \lambda)\) of \(\dot{x} = f(t, x, \lambda)\) defined on \([t_0, T]\) with \(z(t_0, \lambda) = z_0\), and \(z(t, \lambda)\) satisfies \(|z(t, \lambda) - y(t, \lambda_0)| < \epsilon\) for all \(t \in [t_0, T]\).

Proof: By the continuity of \(y(t, \lambda_0)\) in \(t\) and the compactness of \([t_0, T]\), we know that \(y(t, \lambda_0)\) is uniformly bounded on \([t_0, T]\). Define a "tube" \(U\) around the solution \(y(t, \lambda_0)\) by

\[
U = \{(t, x) \in [t_0, T] \times \mathbb{R}^n : |x - y(t, \lambda_0)| \leq \epsilon\}
\]

Suppose \(U \subset [t_0, T] \times D\). If not, then replace \(\epsilon\) by a smaller \(\epsilon\) that is small enough to force \(U\) to be a subset of \([t_0, T] \times D\). The set \(U\) is compact and so \(f(t, x, \lambda)\) is Lipschitz in \(x\) on \(U\) with a Lipschitz constant of say \(L\). By continuity of \(f\) in \(\lambda\), for any \(\alpha > 0\) there is \(\beta > 0\) with \(\beta < \epsilon\) such that

\[
|f(t, x, \lambda) - f(t, x, \lambda_0)| < \alpha \text{ for all } (t, x) \in U \text{ and all } |\lambda - \lambda_0| < \beta
\]

Take \(\alpha < \epsilon\) and \(|z_0 - y_0| < \alpha\). By the local existence and uniqueness theorem 30 there is a unique solution \(z(t, \lambda)\) on some time interval \([t_0, t_0 + \Delta]\). The solution starts in side \(U\) and as long as it remains in the tube, this solution can be extended. We can prove this assertion by noting that if we choose \(\alpha\) small enough, then the solution remains in \(U\) for all \(t \in [t_0, T]\). So let \(\tau\) be the first time when the solution leaves the tube. On the time interval \([t_0, \tau]\), all the conditions of the perturbation sensitivity theorem 34 are satisfied with \(\gamma = \mu = \alpha\). So we can conclude

\[
|z(t, \lambda) - y(t, \lambda_0)| \leq \alpha e^{L(t-t_0)} + \alpha \left( e^{L(t-t_0)} - 1 \right) < \alpha \left( 1 + \frac{1}{L} \right) e^{L(t-t_0)}
\]

Choosing \(\alpha < \epsilon L e^{-L(T-t_0)} / (1 + L)\) ensure that solution \(z(t, \lambda)\) cannot leave the tube during the interval \([t_0, T]\). Therefore \(z\) is defined on \([t_0, T]\) and satisfies \(|z(t, \lambda) - y(t, \lambda_0)| < \epsilon\). Taking \(\delta = \min\{\alpha, \beta\}\) therefore completes the proof. \(\diamondsuit\)

6. Comparison Principle

The comparison principle is a useful tool for finding bounds on the solutions to initial value problems. This principle establishes conditions when the solution \(\dot{x}(t) = f(t, x)\) is bounded above by another function \(v : \mathbb{R} \to \mathbb{R}\). This principle holds even if \(v\) is not differentiable, but has an upper right-hand derivative, \(D^+ v\) that is often called the Dini derivative.

Let \(v : \mathbb{R} \to \mathbb{R}\) be a function and define its Dini derivative as the function \(D^+[v] : \mathbb{R} \to \mathbb{R}\) that takes values \(t \in \mathbb{R}\) of

\[
D^+[v](t) = \limsup_{h \to 0} \frac{v(t+h) - v(t)}{h}
\]
If \( v \) is locally Lipschitz, then \( D^+[v](t) \) is finite for all \( t \) and if \( v \) is differentiable then its Dini derivative is the usual derivative. Note that the Dini derivative can exist for discontinuous functions. We can now state and prove the comparison principle.

**Theorem 36. (Comparison Principle)** Consider the scalar differential equation

\[
\dot{u}(t) = f(t, u(t))
\]

with initial condition \( u(0) = u_0 \) with \( f \) being continuous in \( t \) and locally Lipschitz in \( u \) for all \( t \geq 0 \). Let \([0,T]\) be the maximum interval of existence of \( u \). Let \( v \) be a continuous function whose Dini derivative satisfies

\[
D^+[v](t) \leq f(t, v(t))
\]

with \( v(0) < u_0 \). Then \( v(t) \leq u(t) \) for all \( t \in [t_0, T) \).

**Proof:** Consider the differential equation

\[
\dot{z} = f(t, z) + \lambda, \quad z(0) = u_0
\]

where \( \lambda \) is a positive constant. On any compact interval \([0, t_1]\), due to the continuity of solutions, we know that for any \( \epsilon > 0 \) there exists \( \delta > 0 \) such that if \( \lambda < \delta \) then this system has a unique solution \( z(t, \lambda) \) defined on \([0, t_1]\) and \( |z(t, \lambda) - u(t)| < \epsilon \). We first claim that \( v(t) \leq z(t, \lambda) \). This can be proven through contradiction, for if this assertion were not true, then there would be times \( a, b \in [0, t_1] \) when \( v(a) = z(a, \lambda) \) and \( v(t) > z(t, \lambda) \) for \( a < t \leq b \). This observation would imply

\[
v(t) - v(a) > z(t, \lambda) - z(a, \lambda)
\]

for all \( t \in (a, b] \), which means that

\[
D^+[v](a) \geq \dot{z}(a, \lambda) = f(a, z(a, \lambda)) + \lambda > f(a, v(a))
\]

This contradicts the assumption that \( D^+[v](t) \leq f(t, v(t)) \) and so \( v(t) \leq z(t, \lambda) \) for all \( t \in [0, t_1] \).

Now assume that \( v(t) \) is not less than or equal to \( u(t) \). This would mean there exists \( a \in (0, t_1] \) such that \( v(a) > u(a) \). Taking \( \epsilon = (v(a) - u(a))/2 \) implies

\[
v(a) - z(a, \lambda) = v(a) - u(a) + u(a) - z(a, \lambda) \geq \epsilon
\]

which contradicts the first claim. \( \Box \)

Let us now examine some examples illustrating how the comparison principle is used. Consider the scalar differential equation

\[
\dot{x}(t) = f(x(t)) = -(1 + x^2)x
\]

with \( x(0) = a \). Let \( v(t) = x^2(t) \) and its time derivative is

\[
\dot{v}(t) = 2x(t)\dot{x}(t) = -2x(1 + x^2)x = -2x^2 - 2x^4 \leq -2x^2
\]
So \( v \) satisfies the differential inequality

\[
\dot{v}(t) \leq -2v(t)
\]

where \( v(0) = a^2 \). Let \( u(t) \) be the solution to the differential equation

\[
\dot{u}(t) = -2u(t)
\]

with initial condition \( u(0) = a^2 \). The solution of this differential equation is \( u(t) = a^2e^{-2t} \). So by the comparison principle we know that \( v \) satisfies

\[
v(t) \leq a^2e^{-2t}
\]

and since \( |x(t)| = \sqrt{v(t)} \), we can conclude that

\[
|x(t)| = \sqrt{v(t)} \leq e^{-t}|a|
\]

for all \( t \geq 0 \).

Let us now consider a \textit{forced} version of the preceding system. In particular we are looking for an upper bound on \( |x(t)| \) when \( x \) satisfies the differential equation

\[
\dot{x} = -(1 + x^2)x + e^t
\]

with \( x(0) = a \). As before we consider a comparison function \( v(t) = x^2(t) \) and note that

\[
\dot{v}(t) = 2x\dot{x} = -2x^2 - 2x^4 + 2xe^t \leq -2v + 2\sqrt{v}e^t
\]

We can try to solve this differential equation for \( v \), but this may be too hard to do.

So let us consider an alternative choice for \( v \). In particular let \( v(t) = |x(t)| \). This function is only differentiable when \( x(t) \neq 0 \). For those \( x \) we can see

\[
\dot{v}(t) = \frac{d}{dt} \sqrt{x^2} = -|x|(1 + x^2) + \frac{x}{|x|} e^t
\]

Since \( 1 + x^2 \geq 1 \), we know that \(-|x|(1 + x^2) \leq -|x|\) and so the bounding inequality becomes

\[
\dot{v} \leq -v(t) + e^t
\]

which is linear and can be solved. Then application of the comparison principle allows us to conclude that

\[
v(t) = |x(t)| \leq e^{-t}|a| + \frac{1}{2}(e^t - e^{-t})
\]

The use of comparison functions will be important in our later study of system stability (starting in chapter 5). In particular, we will consider a specific type of comparison function called a Lyapunov function whose existence \textit{certifies} that the system’s equilibrium possesses the property of Lyapunov stability. In this regard, these comparison functions become what we will later refer to as \textit{certificates}.
While the existence and uniqueness theory for ordinary differential equations with *continuous* right hand sides is well understood. One often encounters systems in which the right hand side of the equation is discontinuous. These arise frequently in control applications for mechanical systems when one must switch between various types of control actions (on-off) in a discontinuous manner. The basic problem we find is that switched control laws give rise to differential equations with discontinuous right hand sides.

Let \( s : \mathbb{R}^n \to \mathbb{R} \) be a function so that

\[
S_0 = \{ x \in \mathbb{R}^n : s(x) = 0 \}
\]

is an \( n - 1 \) dimensional surface in \( \mathbb{R}^n \). We’ll refer to \( S_0 \) as the *switching boundary*. Now define the differential equation

\[
\dot{x} = f^+(x) \text{ for } \{ x : s(x) > 0 \} = S_+ \\
\dot{x} = f^-(x) \text{ for } \{ x : s(x) < 0 \} = S_-
\]

(39)

where \( f^+ \) and \( f^- \) are smooth functions from \( \mathbb{R}^n \) into \( \mathbb{R}^n \). In general, \( f^+ \) and \( f^- \) do not match on \( S_0 \) so that the dynamics are discontinuous at \( S_0 \). In other words there is a "step" change in the vector field as one traverses \( S_0 \) is a transverse manner. Figure 1 shows the possible phase portraits associated with the discontinuity. In the figure on the left hand side, the trajectories both point towards the discontinuity surface \( S_0 \). Intuitively, one would expect that imperfections in the switching would cause the state trajectory to "chatter" or zig-zag across the discontinuity surface, as suggested by the jagged line in the figure. In the case of the middle figure the trajectories of \( f^+ \) point toward \( S_0 \) and those of \( f^- \) point away from it. There appears, therefore, to be no problem with continuing the solution trajectories across \( S_0 \) in this instance. In the right hand figure the trajectories of \( f^+ \) and \( f^- \) both point away from \( S_0 \). It would appear that the initial condition on \( S_0 \) would follow either one of the trajectories, though it is impossible to say which one.

The standard technique used for dealing with this breakdown in assumptions is to regularize the system. This means one adds a small perturbation to the given system so as to make the system well defined and then study the behavior of this well defined system in the limit as the regularizing perturbation goes to zero. One common regularization for the case of step discontinuities in differential equations is to assume that the "switched differential equation" is the limit as \( \Delta \to 0 \) of a *hysteretic* switching mechanism shown in figure 2. The variable \( y \) represents the switching variable: when \( y = +1 \) the dynamics are described by \( f^+ \) and when \( y = -1 \) they are described by \( f^- \). Applying this regularization yields the phase portraits shown in figure 1’s left hand pane (i.e. both \( f^+ \) and \( f^- \) point into the switching surface). The frequency of crossing \( S_0 \) increases (chatters) as \( \Delta \downarrow 0 \). Also it appears in the limit when \( \Delta = 0 \) that the trajectory is confined to the switching surface \( S_0 \). This particular approach to regularizing differential equations with discontinuous right hand sides was due to Fillipov [FA88]. The "sliding" solution is said to solve the differential equation in the sense of Fillipov.
This chapter sought to link the abstract topological definition of a dynamical system to differential equation representations. Much of this discussion will be found in many textbooks on ordinary differential equations. A central theme was that ODEs are local representations and flows are global characterizations of system dynamics. That viewpoint was drawn from [Arn73]. The particular approach I used in discussing uniqueness was drawn from [Kha96].

What we learned in this chapter was that characterizing a system through a parameterized group of transformations, $\phi$, was a global representation of the system’s dynamics that could always be represented locally by an ordinary differential equation. In practice, however, one usually starts from from a differential equation and not all ODE-based models give rise to global representations of $\phi$. This chapter showed that we had to
limit the right hand side of the ODE to Lipschitz functions to ensure the uniqueness of trajectories implied by our global topological description of a dynamical system. We also looked at results characterizing the sensitivity of these ODE solutions to initial conditions and parameters. We examined methods for extending the local solution of an ODE into a global solution and finally we reviewed some alternative solution concepts for ODE’s that have been useful when smoothness is too strong a requirement on the system’s local behavior.

As suggested in the introductory chapter 1, one way of studying and managing the behavior of dynamical systems is through linearization. The next chapter will justify this approach by showing that the flows of a dynamical system are topologically equivalent to the flows of its linearization.