CHAPTER 1

Introduction

We will begin by discussing some general properties of initial value problems (IVPs) for ordinary differential equations (ODEs) as well as the basic underlying mathematical theory.

1.1. First-order systems of ODEs

> Does the Flap of a Butterfly's Wings in Brazil Set off a Tornado in Texas?
>
> Edward Lorenz, 1972

We consider an autonomous system of first-order ODEs of the form

$$x_t = f(x) \tag{1.1}$$

where $x(t) \in \mathbb{R}^d$ is a vector of dependent variables, $f : \mathbb{R}^d \to \mathbb{R}^d$ is a vector field, and x_t is the time-derivative, which we also write as dx/dt or $\dot{x}$. In component form, $x = (x_1, \dots, x_d)$,

$$f(x) = (f_1(x_1 \dots, x_d), \dots, f_n(x_1, \dots x_d)),$$

and the system is

$$\begin{aligned} x_{1t} &= f_1(x_1 \dots, x_d), \\ x_{2t} &= f_2(x_1 \dots, x_d), \\ &\dots, \\ x_{dt} &= f_d(x_1 \dots, x_d). \end{aligned}$$

We may regard (1.1) as describing the evolution in continuous time t of a dynamical system with finite-dimensional state $x(t)$ of dimension d.

Autonomous ODEs arise as models of systems whose laws do not change in time. They are invariant under translations in time: if $x(t)$ is a solution, then so is $x(t + t_0)$ for any constant t_0.

Example 1.1. The Lorenz system for $(x, y, z) \in \mathbb{R}^3$ is

$$\begin{aligned} x_t &= \sigma(y - x), \\ y_t &= rx - y - xz, \\ z_t &= xy - \beta z. \end{aligned} \tag{1.2}$$

The system depends on three positive parameters σ, r, β; a commonly studied case is $\sigma = 10$, $r = 28$, and $\beta = 4/3$. Lorenz (1963) obtained (1.2) as a truncated model of thermal convection in a fluid layer, where σ has the interpretation of a Prandtl number(the ratio of kinematic viscosity and thermal diffusivity), r corresponds to a Rayleigh number, which is a dimensionless parameter proportional to the temperature difference across the fluid layer and the gravitational acceleration acting on the fluid, and β is a ratio of the height and width of the fluid layer.

Lorenz discovered that solutions of (1.2) may behave chaotically, showing that even low-dimensional nonlinear dynamical systems can behave in complex ways. Solutions of chaotic systems are sensitive to small changes in the initial conditions, and Lorenz used this model to discuss the unpredictability of weather (the "butterfly effect").

If $\bar{x} \in \mathbb{R}^d$ is a zero of f, meaning that

$$f(\bar{x}) = 0, \tag{1.3}$$

then (1.1) has the constant solution $x(t) = \bar{x}$. We call $\bar{x}$ an *equilibrium solution*, or *steady state solution*, or *fixed point* of (1.1). An equilibrium may be stable or unstable, depending on whether small perturbations of the equilibrium decay — or, at least, remain bounded — or grow. (See Definition 1.14 below for a precise definition.) The determination of the stability of equilibria will be an important topic in the following.

Other types of ODEs can be put in the form (1.1). This rewriting does not simplify their analysis, and may obscure the specific structure of the ODEs, but it shows that (1.1) is rather a general form.

EXAMPLE 1.2. A non-autonomous system for $x(t) \in \mathbb{R}^d$ has the form

$$x_t = f(x, t) \tag{1.4}$$

where $f : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$. A nonautonomous ODE describes systems governed by laws that vary in time *e.g.* due to external influences. Equation (1.4) can be written as an autonomous ('suspended') system for $y = (x, s) \in \mathbb{R}^{n+1}$ with $s = t$ as

$$x_t = f(x, s), \qquad s_t = 1.$$

Note that this increases the order of the system by one, and even if the original system has an equilibrium solution $x(t) = \bar{x}$ such that $f(\bar{x}, t) = 0$, the suspended system has no equilibrium solutions for y.

Higher-order ODEs can be written as first order systems by the introduction of derivatives as new dependent variables.

EXAMPLE 1.3. A second-order system for $x(t) \in \mathbb{R}^d$ of the form

$$x_{tt} = f(x, x_t) \tag{1.5}$$

can be written as a first-order system for $z = (x, y) \in \mathbb{R}^{2d}$ with $y = x_t$ as

$$x_t = y, \qquad y_t = f(x, y).$$

Note that this doubles the dimension of the system.

EXAMPLE 1.4. In Newtonian mechanics, the position $x(t) \in \mathbb{R}^d$ of a particle of mass m moving in d space dimensions in a spatially-dependent force-field $F(x)$, such as a planet in motion around the sun, satisfies

$$m x_{tt} = F(x).$$

If $p = mx_t$ is the momentum of the particle, then (x, p) satisfies the first-order system

$$x_t = \frac{1}{m}p, \qquad p_t = F(x). \tag{1.6}$$

A conservative force-field is derived from a potential $V : \mathbb{R}^d \to \mathbb{R}$,

$$F = -\frac{\partial V}{\partial x}, \qquad (F_1, \dots, F_d) = \left(\frac{\partial V}{\partial x_1}, \dots, \frac{\partial V}{\partial x_d}\right).$$

We use $\partial/\partial x$, $\partial/\partial p$ to denote the derivatives, or gradients with respect to x, p respectively. In that case, (1.6) becomes the Hamiltonian system

$$x_t = \frac{\partial H}{\partial p}, \qquad p_t = -\frac{\partial H}{\partial x} \tag{1.7}$$

where the Hamiltonian

$$H(x, p) = \frac{1}{2m}p^2 + V(x)$$

is the total energy (kinetic + potential) of the particle. The Hamiltonian is a conserved quantity of (1.7), since by the chain rule

$$\begin{aligned}\frac{d}{dt}H\left(x(t), p(t)\right) &= \frac{\partial H}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial H}{\partial p} \cdot \frac{dx}{dt}\\ &= -\frac{\partial H}{\partial x} \cdot \frac{\partial H}{\partial p} + \frac{\partial H}{\partial p} \cdot \frac{\partial H}{\partial x}\\ &= 0.\end{aligned}$$

Thus, solutions (x, p) of (1.7) lie on the level surfaces $H(x, p) = \text{constant}$.

1.2. Existence and uniqueness theorem for IVPs

An initial value problem (IVP) for (1.1) consists of solving the ODE subject to an initial condition (IC) for x:

$$\begin{aligned} x_t &= f(x),\\ x(0) &= x_0. \end{aligned} \tag{1.8}$$

Here, $x_0 \in \mathbb{R}^d$ is a given constant vector. For an autonomous system, there is no loss of generality in imposing the initial condition at $t = 0$, rather than some other time $t = t_0$.

For a first-order system, we impose initial data for x. For a second-order system, such as (1.5), we impose initial data for x and x_t, and analogously for higher-order systems. The ODE in (1.8) determines $x_t(0)$ from x_0, and we can obtain all higher order derivatives $x^{(n)}(0)$ by differentiating the equation with respect to t and evaluating the result at $t = 0$. Thus, it reasonable to expect that (1.8) determines a unique solution, and this is indeed true provided that $f(x)$ satisfies a mild smoothness condition, called Lipschitz continuity, which is nearly always met in applications. Before stating the existence-uniqueness theorem, we explain what Lipschitz continuity means.

We denote by

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

the Euclidean norm of a vector $x \in \mathbb{R}^d$.

Definition 1.5. A function $f : \mathbb{R}^d \to \mathbb{R}^d$ is *locally Lipschitz continuous* on $\mathbb{R}^d$, or Lipschitz continuous for short, if for every $R > 0$ there exists a constant $M > 0$ such that

$$|f(x) - f(y)| \leq M|x - y| \qquad \text{for all } x, y \in \mathbb{R}^d \text{ such that } |x|, |y| \leq R.$$

We refer to M as a Lipschitz constant for f.

A sufficient condition for $f = (f_1, \dots, f_d)$ to be a locally Lipschitz continuous function of $x = (x_1, \dots, x_d)$ is that f is continuous differentiable (C^1), meaning that all its partial derivatives

$$\frac{\partial f_i}{\partial x_j}, \qquad 1 \leq i, j \leq d$$

exist and are continuous functions.

To show this, note that from the fundamental theorem of calculus

$$\begin{aligned} f(x) - f(y) &= \int_0^1 \frac{d}{ds} f\left(y + s(x-y)\right)\, ds \\ &= \int_0^1 Df\left(y + s(x-y)\right)(x-y)\, ds. \end{aligned}$$

Here Df is the derivative of f, whose matrix is the Jacobian matrix of f with components $\partial f_i / \partial x_j$. Hence

$$\begin{aligned} |f(x) - f(y)| &\leq \int_0^1 \left|Df\left(y + s(x-y)\right)(x-y)\right|\, ds \\ &\leq \left(\int_0^1 \left\|Df\left(y + s(x-y)\right)\right\|\, ds\right) |x - y| \\ &\leq M|x - y| \end{aligned}$$

where $\|Df\|$ denotes the Euclidean matrix norm of Df and

$$M = \max_{0 \leq s \leq 1} \left\|Df\left(y + s(x-y)\right)\right\|.$$

For scalar-valued functions, this result also follows from the mean value theorem.

Example 1.6. The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is locally Lipschitz continuous on $\mathbb{R}$, since it is continuously differentiable. The function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = |x|$ is Lipschitz continuous, although it is not differentiable at $x = 0$. The function $h : \mathbb{R} \to \mathbb{R}$ defined by $h(x) = |x|^{1/2}$ is not Lipschitz continuous at $x = 0$, although it is continuous.

The following result, due to Picard and Lindelöf, is the fundamental local existence and uniqueness theorem for IVPs for ODEs. It is a local existence theorem because it only asserts the existence of a solution for sufficiently small times, not necessarily for all times.

Theorem 1.7 (Existence-uniqueness). *If $f : \mathbb{R}^d \to \mathbb{R}^d$ is locally Lipschitz continuous, then there exists a unique solution $x : I \to \mathbb{R}^d$ of (1.8) defined on some time-interval $I \subset \mathbb{R}$ containing $t = 0$.*

In practice, to apply this theorem to (1.8), we usually just have to check that the right-hand side $f(x)$ is a continuously differentiable function of the dependent variables x.

We will not prove Theorem 1.7 here, but we explain the main idea of the proof. Since it is impossible, in general, to find an explicit solution of a nonlinear IVP such as (1.8), we have to construct the solution by some kind of approximation procedure. Using the method of Picard iteration, we rewrite (1.8) as an equivalent integral equation

$$x(t) = x_0 + \int_0^t f\left(x(s)\right)\, ds. \tag{1.9}$$

This integral equation formulation includes both the initial condition and the ODE. We then define a sequence $x_n(t)$ of functions by iteration, starting from the constant initial data x_0:

$$x_{n+1}(t) = x_0 + \int_0^t f\left(x_n(s)\right)\, ds, \qquad n = 1, 2, 3, \dots. \tag{1.10}$$

Using the Lipschitz continuity of f, one can show that this sequence converges uniformly on a sufficiently small time interval I to a unique function $x(t)$. Taking the limit of (1.10) as $n \to \infty$, we find that $x(t)$ satisfies (1.9), so it is the solution of (1.8).

Two simple scalar examples illustrate Theorem 1.7. The first example shows that solutions of nonlinear IVPs need not exist for all times.

Example 1.8. Consider the IVP

$$x_t = x^2, \qquad x(0) = x_0.$$

For $x_0 \neq 0$, we find by separating variables that the solution is

$$x(t) = -\left(\frac{1}{t - 1/x_0}\right). \tag{1.11}$$

If $x_0 > 0$, the solution exists only for $-\infty < t < t_0$ where $t_0 = 1/x_0$, and $x(t) \to -\infty$ as $t \to t_0$. Note that the larger the initial data x_0 the smaller the ‘blow-up’ time t_0. If $x_0 < 0$, then $t_0 < 0$ and the solution exists for $t_0 < t < \infty$. Only if $x_0 = 0$ does the solution $x(t) = 0$ exists for all times $t \in \mathbb{R}$.

One might consider using (1.11) past the time t_0, but continuing a solution through infinity does not make much sense in evolution problems. In applications, the appearance of a singularity typically signifies that the assumptions of the mathematical model have broken down in some way.

The second example shows that solutions of (1.8) need not be unique if f is not Lipschitz continuous.

Example 1.9. Consider the IVP

$$x_t = |x|^{1/2}, \qquad x(0) = 0. \tag{1.12}$$

The right-hand side of the ODE, $f(x) = |x|^{1/2}$, is not differentiable or Lipschitz continuous at the initial data $x = 0$. One solution is $x(t) = 0$, but this is not the only solution. Separating variables in the ODE, we get the solution

$$x(t) = \frac{1}{4}(t - t_0)^2.$$

Thus, for any $t_0 \geq 0$, the function

$$x(t) = \begin{cases} 0 & \text{if } t \leq t_0 \\ (1/4)(t - t_0)^2 & \text{if } t > t_0 \end{cases}$$

is also a solution of the IVP (1.12). The parabolic solution can 'take off' spontanteously with zero derivative from the zero solution at any nonnegative time t_0. In applications, a lack of uniqueness typically means that something is missing from the mathematical model.

If $f(x)$ is only assumed to be a continuous function of x, then solutions of (1.8) always exist (this is the Peano existence theorem) although they may fail to be unique, as shown by Example 1.9. In future, we will assume that f is a smooth function; typically, f will be C^∞, meaning that is has continuous derivatives of all orders. In that case, the issue of non-uniqueness does not arise.

Even for arbitrarily smooth functions f, the solution of the nonlinear IVP (1.8) may fail to exist for all times if $f(x)$ grows faster than a linear function of x, as in Example 1.8. According to the following theorem, the only way in which global existence can fail is if the solution 'escapes' to infinity. We refer to this phenomenon informally as 'blow-up.'

THEOREM 1.10 (Extension). *If $f : \mathbb{R}^d \to \mathbb{R}^d$ is locally Lipschitz continuous, then the solution $x : I \to \mathbb{R}^d$ of (1.8) exists on a maximal time-interval*

$$I = (T_-, T_+) \subset \mathbb{R}$$

where $-\infty \leq T_- < 0$ and $0 < T_+ \leq \infty$. If $T_+ < \infty$, then $|x(t)| \to \infty$ as $t \uparrow T_+$, and if $T_- > -\infty$, then $|x(t)| \to \infty$ as $t \downarrow T_-$,

This theorem implies that we can continue a solution of the ODE so long as it remains bounded.

EXAMPLE 1.11. Consider the function defined for $t \neq 0$ by

$$x(t) = \sin\left(\frac{1}{t}\right).$$

This function cannot be extended to a differentiable, or even continuous, function at $t = 0$ even though it is bounded. This kind of behavior cannot happen for solutions of ODEs with continuous right-hand sides, because the ODE implies that the derivative x_t remains bounded if the solution x remains bounded. On the other hand, an ODE may have a solution like $x(t) = 1/t$, since the derivative x_t only becomes large when x itself becomes large.

EXAMPLE 1.12. Theorem 1.7 implies that the Lorenz system (1.1) with arbitrary initial conditions

$$x(0) = x_0, \qquad y(0) = y_0, \qquad z(0) = z_0$$

has a unique solution defined on some time interval containing 0, since the right hand side is a smooth (in fact, quadratic) function of (x, y, z). The theorem does not imply, however, that the solution exists for all t.

Nevertheless, we claim that when the parameters (σ, r, β) are positive the solution exists for all $t \geq 0$. From Theorem 1.10, this conclusion follows if we can show that the solution remains bounded, and to do this we introduce a suitable Lyapunov function. A convenient choice is

$$V(x, y, z) = rx^2 + \sigma y^2 + \sigma(z - 2r)^2.$$

Using the chain rule, we find that if (x, y, z) satisfies (1.2), then

$$\begin{aligned}\frac{d}{dt}V(x,y,z) &= 2rxx_t + 2\sigma yy_t + 2\sigma(z-2r)z_t \\ &= 2r\sigma x(y-x) + 2\sigma y(rx - y - xz) + 2\sigma(z-2r)(xy-\beta z) \\ &= -2\sigma\left[rx^2 + y^2 + \beta(z-r)^2\right] + 2\beta\sigma r^2.\end{aligned}$$

Hence, if $W(x,y,z) > \beta r^2$, where

$$W(x,y,z) = rx^2 + y^2 + \beta(z-r)^2,$$

then $V(x,y,z)$ is decreasing in time. This means that if C is sufficiently large that the ellipsoid $V(x,y,z) < C$ contains the ellipsoid $W(x,y,z) \leq \beta r^2$, then solutions cannot escape from the region $V(x,y,z) < C$ forward in time, since they move 'inwards' across the boundary $V(x,y,z) = C$. Therefore, the solution remains bounded and exists for all $t \geq 0$.

Note that this argument does not preclude the possibility that solutions of (1.2) blow up backwards in time. The Lorenz system models a forced, dissipative system and its dynamics are not time-reversible. (This contrasts with the dynamics of conservative, Hamiltonian systems, which are time-reversible.)

1.3. Linear systems of ODEs

An IVP for a (homogeneous, autonomous, first-order) linear system of ODEs for $x(t) \in \mathbb{R}^d$ has the form

$$\begin{aligned} x_t &= Ax, \\ x(0) &= x_0 \end{aligned} \tag{1.13}$$

where A is a $d \times d$ matrix and $x_0 \in \mathbb{R}^d$. This system corresponds to (1.8) with $f(x) = Ax$. Linear systems are much simpler to study than nonlinear systems, and perhaps the first question to ask of any equation is whether it is linear or nonlinear.

The linear IVP (1.13) has a unique global solution, which is given explicitly by

$$x(t) = e^{tA}x_0, \qquad -\infty < t < \infty$$

where

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

is the matrix exponential.

If A is nonsingular, then (1.13) has a unique equilibrium solution $x = 0$. This equilibrium is stable if all eigenvalues of A have negative real parts and unstable if some eigenvalue of A has positive real part. If A is singular, then there is a ν-dimensional subspace of equilibria where ν is the nullity of A.

Linear systems are important in their own right, but they also arise as approximations of nonlinear systems. Suppose that $\bar{x}$ is an equilibrium solution of (1.1), satisfying (1.3). Then writing

$$x(t) = \bar{x} + y(t)$$

and Taylor expanding $f(x)$ about $\bar{x}$, we get

$$f(\bar{x} + y) = Ay + \dots$$

where A is the derivative of f evaluated at $\bar{x}$, with matrix (a_{ij}):

$$A = Df(\bar{x}), \qquad a_{ij} = \frac{\partial f_i}{\partial x_j}(\bar{x}).$$

The linearized approximation of (1.1) at the equilibrium $\bar{x}$ is then

$$y_t = Ay.$$

An important question is if this linearized system provides a good local approximation of the nonlinear system for solutions that are near equilibrium. This is the case under the following condition.

DEFINITION 1.13. An equilibrium $\bar{x}$ of (1.1) is hyperbolic if $Df(\bar{x})$ has no eigenvalues with zero real part.

Thus, for a hyperbolic equilibrium, all solutions of the linearized system grow or decay exponentially in time. According to the Hartman-Grobman theorem, if $\bar{x}$ is hyperbolic, then the flows of the linearized and nonlinear system are (topologically) equivalent near the equilibrium. In particular, the stability of the nonlinear equilibrium is the same as the stability of the equilibrium of the linearized system. One has to be careful, however, in drawing conclusions about the behavior of the nonlinear system from the linearized system if $Df(\bar{x})$ has eigenvalues with zero real part. In that case the nonlinear terms may cause the growth or decay of perturbations from equilibrium, and the behavior of solutions of the nonlinear system near the equilibrium may differ qualitatively from that of the linearized system.

Non-hyperbolic equilibria are not typical for specific systems, since one does not expect the eigenvalues of a given matrix to have a real part that is exactly equal to zero. Nevertheless, non-hyperbolic equilibria arise in an essential way in bifurcation theory when an eigenvalue of a system that depends on some parameter has real part that passes through zero.

1.4. Phase space

> it may happen that small differences in the initial conditions produce very great ones in the final phenomena
>
> Henri Poincaré, 1908

Very few nonlinear systems of ODEs are explicitly solvable. Therefore, rather than looking for individual analytical solutions, we try to understand the qualitative behavior of their solutions. This global, geometrical approach was introduced by Poincaré (1880).

We may represent solutions of (1.8) by solution curves, trajectories, or orbits, $x(t)$ in phase, or state, space $\mathbb{R}^d$. These trajectories are integral curves of the vector field f, meaning that they are tangent to f at every point. The existence-uniqueness theorem implies if the vector field f is smooth, then a unique trajectory passes through each point of phase space and that trajectories cannot cross. We may visualize f as the steady velocity field of a fluid that occupies phase space and the trajectories as particle paths of the fluid.

Let $x(t; x_0)$ denote the solution of (1.8), defined on its maximal time-interval of existence $T_-(x_0) < t < T_+(x_0)$. The existence-uniqueness theorem implies that we can define a flow map, or solution map, $\Phi_t : \mathbb{R}^d \to \mathbb{R}^d$ by

$$\Phi_t(x_0) = x(t; x_0), \qquad T_-(x_0) < t < T_+(x_0).$$

That is, Φ_t maps the initial data x_0 to the solution at time t. Note that $\Phi_t(x_0)$ is not defined for all $t \in \mathbb{R}$, $x_0 \in \mathbb{R}^d$ unless all solutions exist globally. In the fluid analogy, Φ_t may be interpreted as the map that takes a particle from its initial location at time 0 to its location at time t.

The flow map Φ_t of an autonomous system has the group property that

$$\Phi_t \circ \Phi_s = \Phi_{t+s}$$

where $\circ$ denotes the composition of maps *i.e.* solving the ODE for time $t + s$ is equivalent to solving it for time s then for time t. We remark that the solution map of a non-autonomous IVP,

$$x_t = f(x, t), \qquad x(t_0) = x_0$$

with solution $x(t; x_0, t_0)$, is defined by

$$\Phi_{t,t_0}(x_0) = x(t; x_0, t_0).$$

The map depends on both the initial and final time, not just their difference, and satisfies

$$\Phi_{t,s} \circ \Phi_{s,r} = \Phi_{t,r}.$$

If $\bar{x}$ is an equilibrium solution of (1.8), with $f(\bar{x}) = 0$, then

$$\Phi_t(\bar{x}) = \bar{x},$$

which explains why equilibria are referred to as fixed points (of the flow map). We may state a precise definition of stability in terms of the flow map. There are many different, and not entirely equivalent definitions, of stability; we give only the simplest and most commonly used ones.

Definition 1.14. An equilibrium $\bar{x}$ of (1.8) is Lyapunov stable (or stable, for short) if for every $\epsilon > 0$ there exists $\delta > 0$ such that if $|x - \bar{x}| < \delta$ then

$$|\Phi_t(x) - \bar{x}| \qquad \text{for all } t \geq 0.$$

The equilibrium is asymptotically stable if it is Lyapunov stable and there exists $\eta > 0$ such that if $|x - \bar{x}| < \eta$ then

$$\Phi_t(x) \to \bar{x} \qquad \text{as } t \to \infty.$$

Thus, stability means that solutions which start sufficiently close to the equilibrium remain arbitrarily close for all $t \geq 0$, while asymptotic stability means that in addition nearby solutions approach the equilibrium as $t \to \infty$. Lyapunov stability does not imply asymptotic stability since, for example, nearby solutions might oscillate about an equilibrium without decaying toward it. Also, it is not sufficient for asymptotic stability that all nearby solutions approach the equilibrium, because they could make large excursions before approaching the equilibrium, which would violate Lyapunov stability.

The next result implies that the solution of an IVP depends continuously on the initial data, and that the flow map of a smooth vector field is smooth. Here, 'smooth' means, for example, C^1 or C^∞.

Theorem 1.15 (Continuous dependence on initial data). *If the vector field f in (1.8) is locally Lipschitz continuous, then the corresponding flow map $\Phi_t : \mathbb{R}^d \to \mathbb{R}^d$ is locally Lipschitz continuous. Moreover, the existence times T_+ (respectively, T_-) are lower (respectively, upper) semi-continuous function of x_0. If the vector field f in (1.8) is smooth, then the corresponding flow map $\Phi_t : \mathbb{R}^d \to \mathbb{R}^d$ is smooth.*

Here, the lower-semicontinuity of T_+ means that

$$T_+(x_0) \leq \liminf_{x \to x_0} T_+(x),$$

so that solutions with initial data near x_0 exist for essentially as long, or perhaps longer, than the solution with initial data x_0.

Theorem 1.15 means that solutions remain close over a finite time-interval if their initial data are sufficiently close. After long enough times, however, two solutions may diverge by an arbitrarily large amount however close their initial data.

Example 1.16. Consider the scalar, linear ODE $x_t = x$. The solutions $x(t)$, $y(t)$ with initial data $x(0) = x_0$, $y(0) = y_0$ are given by

$$x(t) = x_0 e^t, \qquad y(t) = y_0 e^t.$$

Suppose that $[0, T]$ is any given time interval, where $T > 0$. If $|x_0 - y_0| \leq \epsilon e^{-T}$, then the solutions satisfy $|x(t) - y(t)| \leq \epsilon$ for all $0 \leq t \leq T$, so the solutions remain close on $[0, T]$, but $|x(t) - y(t)| \to \infty$ as $t \to \infty$ whenever $x_0 \neq y_0$.

Not only do the trajectories depend continuously on the initial data, but if f is Lipschitz continuous they can diverge at most exponentially quickly in time. If M is the Lipschitz constant of f and $x(t)$, $y(t)$ are two solutions of (1.8), then

$$\frac{d}{dt}|x - y| \leq M|x - y|.$$

It follows from Gronwall's inequality that if $x(0) = x_0$, $y(0) = y_0$, then

$$|x(t) - y(t)| \leq |x_0 - y_0| e^{Mt}.$$

The local exponential divergence (or contraction) of trajectories may be different in different directions, and is measured by the Lyapunov exponents of the system. The largest such exponent is called *the* Lyapunov exponent of the system. Chaotic behavior occurs in systems with a positive Lyapunov exponent and trajectories that remain bounded; it is associated with the local exponential divergence of trajectories (essentially a linear phenomenon) followed by a global folding (typically as a result of nonlinearity).

One way to organize the study of dynamical systems is by the dimension of their phase space (following the Trolls of Discworld: one, two, three, many, and lots). In one or two dimensions, the non-intersection of trajectories strongly restricts their possible behavior: in one dimension, solutions can only increase or decrease monotonically to an equilibrium or to infinity; in two dimensions, oscillatory behavior can occur. In three or more dimensions complex behavior, including chaos, is possible.

For the most part, we will consider dynamical systems with low-dimensional phase spaces (say of dimension $d \leq 3$). The analysis of high-dimensional dynamical systems is usually very difficult, and may require (more or less well-founded) probabilistic assumptions, or continuum approximations, or some other type of approach.

Example 1.17. Consider a gas composed of N classical particles of mass m moving in three space dimensions with an interaction potential $V : \mathbb{R}^3 \to \mathbb{R}$. We denote the positions of the particles by $x = (x_1, x_2, \dots, x_N)$ and the momenta by $p = (p_1, p_2, \dots, p_N)$, where $x_i, p_i \in \mathbb{R}^3$. The Hamiltonian for this system is

$$H(x,p) = \frac{1}{2m}\sum_{i=1}^{N} p_i^2 + \frac{1}{2}\sum_{1\le i\neq j\le N} V\left(x_i - x_j\right),$$

and Hamilton's equations are

$$\frac{dx_i}{dt} = \frac{1}{m}p_i, \qquad \frac{dp_i}{dt} = -\sum_{j\neq i}\frac{\partial V}{\partial x}\left(x_i - x_j\right).$$

The phase space of this system has dimension $6N$. For a mole of gas, we have $N = N_A$ where $N_A \approx 6.02 \times 10^{23}$ is Avogadro's number, and this dimension is extremely large.

In kinetic theory, one considers equations for probability distributions of the particle locations and velocities, such as the Boltzmann equation. One can also approximate some solutions by partial differential fluid equations, such as the Navier-Stokes equations, for suitable averages.

We will mostly consider systems whose phase space is $\mathbb{R}^d$. More generally, the phase space of a dynamical system may be a manifold. We will not give the precise definition of a manifold here; roughly speaking, a d-dimensional manifold is a space that 'looks' locally like $\mathbb{R}^d$, with a d-dimensional local coordinate system about each point, but which may have a different global, topological structure. The d-dimensional sphere $\mathbb{S}^d$ is a typical example. Phase spaces that are manifolds arise, for example, if some of the state variables represent angles.

Example 1.18. The motion of an undamped pendulum of length ℓ in a gravitational field with acceleration g satisfies the pendulum equation

$$\theta_{tt} + \frac{g}{\ell}\sin\theta = 0$$

where $\theta \in \mathbb{T}$ is the angle of the pendulum to the vertical, measured in radians. Here, $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$ denotes the circle; angles that differ by an integer multiple of 2π are equivalent. Writing the pendulum equation as a first-order system for (θ, v) where $v = \theta_t \in \mathbb{R}$ is the angular velocity, we get

$$\theta_t = v, \qquad v_t = -\frac{g}{\ell}\sin\theta$$

The phase space of this system is the cylinder $\mathbb{T} \times \mathbb{R}$. This phase space may be 'unrolled' into $\mathbb{R}^2$ with points on the θ-axis identified modulo 2π, but it is often conceptually clearer to keep the actual cylindrical structure and θ-periodicity in mind.

Example 1.19. The phase space of a rotating rigid body, such as a tumbling satellite, may be identified with the group $SO(3)$ of rotations about its center of mass from some fixed reference configuration. The three Euler angles of a rotation give one possible local coordinate system on the phase space.

Solutions of an ODE with a smooth vector field on a compact phase space without boundaries, such as $\mathbb{S}^d$, exist globally in time since they cannot escape to infinity (or hit a boundary).

1.5. Bifurcation theory

Most applications lead to equations which depend on parameters that characterize properties of the system being modeled. We write an IVP for a first-order system of ODEs for $x(t) \in \mathbb{R}^d$ depending on a vector of parameters $\mu \in \mathbb{R}^m$ as

$$\begin{aligned} x_t &= f(x; \mu), \\ x(0) &= x_0 \end{aligned} \tag{1.14}$$

where $f : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d$.

In applications, it is important to determine a minimal set of dimensionless parameters on which the problem depends and to know what parameter regimes are relevant *e.g.* if some dimensionless parameters are very large or small.

Example 1.20. The Lorentz system (1.2) for $(x, y, z) \in \mathbb{R}^3$ depends on three parameters $(\sigma, r, \beta) \in \mathbb{R}^3$, which we assume to be positive. We typically think of fixing (σ, β) and increasing r, which in the original convection problem corresponds to fixing the fluid properties and the dimensions of the fluid layer and increasing the temperature difference across it.

If the vector field in (1.14) depends smoothly ($e.g.$ C^1 or C^∞) on the parameter μ, then so does the flow map. Explicitly, if $x(t; x_0; \mu)$ denotes the solution of (1.14), then we define the flow map Φ_t by

$$\Phi_t(x_0; \mu) = x(t; x_0; \mu).$$

Theorem 1.21 (Continuous dependence on parameters). *If the vector field $f : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d$ in (1.14) is smooth, then the corresponding flow map $\Phi_t : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d$ is smooth.*

Bifurcation theory is concerned with changes in the qualitative behavior of the solutions of (1.14) as the parameter μ is varied. It may be difficult to carry out a full bifurcation analysis of a nonlinear dynamical system, especially when it depends on many parameters.

The simplest type of bifurcation is the bifurcation of equlbria. The equilibrium solutions of (1.14) satisfy

$$f(\bar{x}; \mu) = 0,$$

so an analysis of equilibrium bifurcations corresponds to understanding how the solutions $\bar{x}(\mu) \in \mathbb{R}^d$ of this $d \times d$ system of nonlinear, algebraic equations depend upon the parameter μ. We refer to a smooth solution $\bar{x} : I \to \mathbb{R}^d$ in a maximal domain I as a solution branch or a branch of equilibria.

There is a closely related dynamical aspect concerning how the stability of the equilibria change as the parameter μ varies. If $\bar{x}(\mu)$ is a branch of equilibrium solutions, then the linearization of the system about $\bar{x}$ is

$$x_t = A(\mu)x, \qquad A(\mu) = D_x f\left(\bar{x}(\mu); \mu\right).$$

Equilibria lose stability if some eigenvalue $\lambda(\mu)$ of A crosses from the left-half of the complex plane into the right-half plane as μ varies. By the implicit function theorem, equilibrium bifurcations are necessarily associated with a real eigenvalue passing through zero, so that A is singular at the bifurcation point.

Equilibrium bifurcations are not the only kind, and the dynamic behavior of a system may change without a change in the equilibrium solutions For example, time-periodic solutions may appear or disappear in a Hopf bifurcation, which occurs where a complex-conjugate pair of complex eigenvalues of A crosses into the right-half plane, or there may be global changes in the geometry of the trajectories in phase space, as in a homoclinic bifurcation.

1.6. Discrete dynamical systems

> Not only in research, but also in the everyday world of politics and economics, we would all be better off if more people realised that simple nonlinear systems do not necessarily possess simple dynamical properties
>
> Robert May, 1976

A (first-order, autonomous) discrete dynamical system for $x_n \in \mathbb{R}^d$ has the form

$$x_{n+1} = f(x_n) \tag{1.15}$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$ and $n \in \mathbb{Z}$ is a discrete time variable.

The orbits, or trajectories of (1.15) consist of a sequence of points $\{x_n\}$ that is obtained by iterating the map f. (They are not curves like the orbits of a continuous dynamical system.) If $f^n = f \circ f \circ \cdots \circ f$ denotes the n-fold composition of f, then

$$x_n = f^n(x_0).$$

If f is invertible, these orbits exists forward and backward in time ($n \in \mathbb{Z}$), while if f is not invertible, they exist in general only forward in time ($n \in \mathbb{N}$). An equilibrium solution $\bar{x}$ of (1.15) is a fixed point of f that satisfies

$$f(\bar{x}) = \bar{x},$$

and in that case $x_n = \bar{x}$ for all n.

A linear discrete dynamical system has the form

$$x_{n+1} = Bx_n, \tag{1.16}$$

where B is a linear transformation on $\mathbb{R}^d$. The solution is

$$x_n = B^n x_0.$$

The linear system (1.16) has the unique fixed point $\bar{x} = 0$ if $I - B$ is a nonsingular linear map. This fixed point is asymptotically stable if all eigenvalues $\lambda \in \mathbb{C}$ of B lie in the unit disc, meaning that $|\lambda| < 1$. It is unstable if B has some eigenvalue with $|\lambda| > 1$ in the exterior of the unit disc.

The linearization of (1.15) about a fixed point $\bar{x}$ is

$$x_{n+1} = Bx_n, \qquad B = Df(\bar{x}).$$

Analogously to the case of continuous systems, we can determine the stability of the fixed point from the stability of the linearized system under a suitable hyperbolicity assumption.

DEFINITION 1.22. A fixed point $\bar{x}$ of (1.15) is hyperbolic if $Df(\bar{x})$ has no eigenvalues with absolute value equal to one.

If $\bar{x}$ is a hyperbolic fixed point of (1.15), then it is asymptotically stable if all eigenvalues of $Df(\bar{x})$ lie inside the unit disc, and unstable if some eigenvalue lies outside the unit disc.

The behavior of even one-dimensional discrete dynamical systems may be complicated. The biologist May (1976) drew attention to the fact that the logistic map,

$$x_{n+1} = \mu x_n \left(1 - x_n\right),$$

leads to a discrete dynamical system with remarkably intricate behavior, even though the corresponding continuous logistic ODE

$$x_t = \mu x(1 - x)$$

is simple to analyze completely. Another well-known illustration of the complexity of discrete dynamical systems is the fractal structure of Julia sets for complex dynamical systems (with two real dimensions) obtained by iterating rational functions $f : \mathbb{C} \to \mathbb{C}$.

Discrete dynamical systems may arise directly as models *e.g.* in population ecology, x_n might represent the population of the nth generation of species. They also arise from continuous dynamical systems.

EXAMPLE 1.23. If Φ_t is the flow map of a continuous dynamical system with globally defined solutions, then the time-one map Φ_1 defines an invertible discrete dynamical system. The dimension of the discrete system is the same as the dimension of the continuous one.

EXAMPLE 1.24. The time-one map of a linear system of ODEs $x_t = Ax$ is

$$B = e^A.$$

Eigenvalues of A in the left-half of the complex plane, with negative real part, map to eigenvalues of B inside the unit disc, and eigenvalues of A is the right-half-plane maps to eigenvalues of B outside the unit disc. Thus, the stability properties of the fixed point $\bar{x} = 0$ in the continuous and discrete descriptions are consistent.

EXAMPLE 1.25. Consider a non-autonomous system for $x \in \mathbb{R}^d$ that depends periodically on time,

$$x_t = f(x, t), \qquad f(x, t + 1) = f(x, t).$$

We define the corresponding Poincaré map $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ by

$$\Phi : x(0) \mapsto x(1).$$

Then Φ defines an autonomous discrete dynamical system of dimension d, which is one less than the dimension $d + 1$ of the original system when it is written in autonomous form. This reduction in dimension makes the dynamics of the Poincaré map easier to visualize than that the original flow, especially when $d = 2$. Moreover, by continuous dependence, trajectories of the original system remain arbitrarily close over the entire time-interval $0 \le t \le 1$ if their initial conditions are sufficient close, so replacing the full flow map by the Poincaré map does not lead to any essential loss of qualitative information.

Fixed points of the Poincaré map correspond to periodic solutions of the original system, although their minimal period need not be one; for example any solution of the original system with period $1/n$ where $n \in \mathbb{N}$ is a fixed point of the Poincaré map, as is any equilibrium solution with $f(\bar{x}, t) = 0$.

Example 1.26. Consider the forced, damped pendulum with non-dimensionalized equation

$$x_{tt} + \delta x_t + \sin x = \gamma \cos \omega t$$

where γ, δ, and ω are parameters, measuring the strength of the damping, the strength of the forcing, and the (angular) frequency of the forcing, respectively. Or a parametrically forced oscillator (such as a swing)

$$x_{tt} + (1 + \gamma \cos \omega t) \sin x = 0.$$

Here, the Poincaré map $\Phi : \mathbb{T} \times \mathbb{R} \to \mathbb{T} \times \mathbb{R}$ is defined by

$$\Phi : (x(0), x_t(0)) \mapsto (x(T), x_t(T)), \qquad T = \frac{2\pi}{\omega}.$$

Floquet theory is concerned with such time-periodic ODEs, including the stability of their time-periodic solutions, which is equivalent to the stability of fixed points of the Poincaré map.

1.7. References

For introductory discussions of the theory of ODEs see [**1, 9**]. Detailed accounts of the mathematical theory are given in [**2, 8**]. A general introduction to nonlinear dynamical systems, with an emphasis on applications, is in [**10**]. An introduction to bifurcation theory in continuous and discrete dynamical systems is [**6**]. For a rigorous but accessible introduction to chaos in discrete dynamical systems, see [**3**]. A classic book on nonlinear dynamical systems is [**7**].