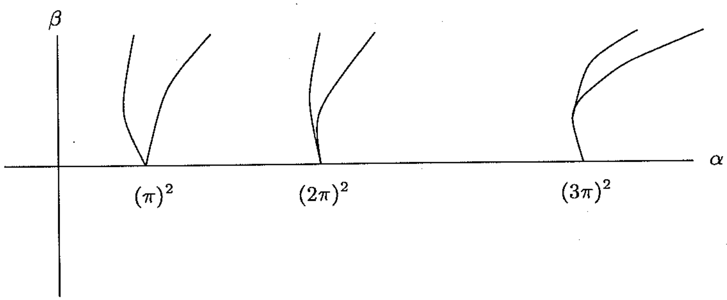

FIGURE 3.6. Regions of instability (Arnold tongues) for the linearized pendulum.

Exercise 3.26. Find an example of a periodic function $s \mapsto a(s)$ with period one such that $\tau_{\beta\beta}(0,0) > 0$. For this choice of the displacement, the inverted pendulum is not stabilized for small $\beta > 0$.

Exercise 3.27. What can you say about stability of the inverted pendulum using Lyapunov's theorem (Theorem 2.68)?

Let us consider the stability of the noninverted pendulum. Note that the linearization of the differential equation (3.75) at $\theta = 0$ is given by

$$w'' + (\alpha - \beta a(s))w = 0,$$

and let $\Psi(t, \alpha, \beta)$ denote the principal fundamental matrix solution of the corresponding homogeneous linear system at $t = 0$. In this case, we have

$$\operatorname{tr} \Psi(1, \alpha, 0) = 2 \cos \sqrt{\alpha}.$$

Because the boundaries of the regions of instability are given by

$$|\operatorname{tr} \Psi(1, \alpha, \beta)| = 2,$$

they intersect the α-axis only if $\sqrt{\alpha}$ is an integer multiple of π. In view of the fact that $\alpha = 1/\Omega^2$, these observations suggest the zero solution is unstable for small amplitude displacements whenever there is an integer n such that the period of the displacement is

$$\frac{1}{\Omega}\Big(\frac{L}{g}\Big)^{1/2} = \frac{n}{2}\Big(2\pi\Big(\frac{L}{g}\Big)^{1/2}\Big);$$

that is, the period of the displacement is a half-integer multiple of the natural frequency of the pendulum. In fact, the instability of the pendulum for a small amplitude periodic displacement with $n = 1$ is demonstrated in every playground by children pumping up swings.

The proof that the instability boundaries do indeed cross the α-axis at the "resonant" points $(\alpha, \beta) = ((n\pi)^2, 0)$, for $n = 1, \ldots, \infty$, is obtained

from an analysis of the Taylor expansion of the function given by $\Psi(1, \alpha, \beta)$ at each resonant point (see Exercise 3.28). Typically, the instability regions are as depicted in Figure 3.6. The instability region with $n = 1$ is "open" at $\beta = 0$ (the tangents to the boundary curves have distinct slopes); the remaining instability regions are "closed." While it is an interesting mathematical problem to determine the general shape of the stability regions ([78], [115]), the model is, perhaps, not physically realistic for large β.

Exercise 3.28. Suppose that $a(s) = \sin(2\pi s)$ and set

$$g(\alpha, \beta) = \operatorname{tr} \Psi(1, \alpha, \beta) - 2.$$

Show that $g_\alpha((n\pi)^2, 0) = 0$ and $g_\beta((n\pi)^2, 0) = 0$. Thus, the implicit function theorem cannot be applied directly to obtain the boundaries of the regions of instability, the boundary curves are singular at the points where they meet the α-axis. By computing appropriate higher order derivatives and analyzing the resulting Taylor expansion of g, show that the regions near the α-axis are indeed as depicted in Figure 3.6. Also, show that the regions become "thinner" as n increases.

3.6 Origins of ODE: Partial Differential Equations

In this section there is an elementary discussion of three "big ideas":

- Certain partial differential equations (PDE) can be viewed as ordinary differential equations with an infinite dimensional phase space.
- Finite dimensional approximations of some PDE are systems of ordinary differential equations.
- Traveling wave fronts in PDE can be described by ordinary differential equations.

While these ideas are very important and therefore have been widely studied, only a few elementary illustrations will be given here. The objective of this section is to introduce these ideas as examples of how ordinary differential equations arise and to suggest some very important areas for further study (see [27], [85], [84], [90], [135], [140], [162], [174], and [190]). We will also discuss the solution of first order PDE as an application of the techniques of ordinary differential equations.

Most of the PDE mentioned in this section can be considered as models of "reaction-diffusion" processes. To see how these models are derived, imagine some substance distributed in a medium. The density of the substance is represented by a function $u : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ so that $(x, t) \mapsto u(x, t)$ gives its density at the site with coordinate x at time t.

If Ω is a region in space with boundary $\partial\Omega$, then the rate of change of the amount of the substance in Ω is given by the flux of the substance through the boundary of Ω plus the amount of the substance generated in Ω; that is,

$$\frac{d}{dt}\int_\Omega u\,d\mathcal{V} = -\int_{\partial\Omega} X\cdot\eta\,dS + \int_\Omega f\,d\mathcal{V}$$

where X is the vector field representing the motion of the substance; $d\mathcal{V}$ is the volume element; dS is the surface element; the vector field η is the outer unit normal field on the boundary of Ω; and f, a function of density, position and time, represents the amount of the substance generated in Ω. The minus sign on the flux term is required because we are measuring the rate of change of the amount of substance *in* Ω. If, for example, the flow is all out of Ω, then $X\cdot\eta \geq 0$ and the minus sign is required because the rate of change of the amount of substance in Ω must be negative.

If Stokes' theorem is applied to rewrite the flux term and the time derivative is interchanged with the integral of the density, then

$$\int_\Omega u_t\,d\mathcal{V} = -\int_\Omega \operatorname{div} X\,d\mathcal{V} + \int_\Omega f\,d\mathcal{V}.$$

Moreover, by using the fact that the region Ω is arbitrary in the integral identity, it is easy to prove the fundamental balance law

$$u_t = -\operatorname{div} X + f. \tag{3.78}$$

To obtain a useful dynamical equation for u from equation (3.78), we need a constitutive relation between the density u of the substance and the flow field X. It is not at all clear how to derive this relationship from the fundamental laws of physics. Thus, we have an excellent example of an important problem where physical intuition must be used to propose a constitutive law whose validity can only be tested by comparing the results of experiments with the predictions of the corresponding model. Problems of this type lie at the heart of applied mathematics and physics.

For equation (3.78), the classic constitutive relation—called Darcy's, Fick's, or Fourier's law depending on the physical context—is

$$X = -K\operatorname{grad} u + \mu V \tag{3.79}$$

where $K \geq 0$ and μ are functions of density, position, and time; and V denotes the flow field for the medium in which our substance is moving. The minus sign on the gradient term represents the assumption that the substance diffuses from higher to lower concentrations.

By inserting the relation (3.79) into the balance law (3.78), we obtain the dynamical equation

$$u_t = \operatorname{div}(K\operatorname{grad} u) - \operatorname{div}(\mu V) + f.$$

Also, if we assume that the diffusion coefficient K is equal to k^2 for some constant k, the function μ is given by $\mu(u,x,t) = \gamma u$ where γ is a constant, and V is an incompressible flow field ($\operatorname{div} V = 0$); then we obtain the most often used reaction-diffusion-convection model equation

$$u_t + \gamma\operatorname{grad} u\cdot V = k^2\Delta u + f. \tag{3.80}$$

In this equation, the gradient term is called the *convection term,* the Laplacian term is called the *diffusion term,* and f is the *source term.* Let us also note that if the diffusion coefficient is zero, the convection coefficient is given by $\gamma = 1$, the source function vanishes, and V is not necessarily incompressible, then equation (3.80) reduces to the law of conservation of mass, also called the *continuity equation,* given by

$$u_t + \operatorname{div}(uV) = 0. \tag{3.81}$$

Because equation (3.80) is derived from general physical principles, this PDE can be used to model many different phenomena. As a result, there is a vast scientific literature devoted to its study. We will not be able to probe very deeply, but we will use equation (3.80) to illustrate a few aspects of the analysis of these models where ordinary differential equations arise naturally.

3.6.1 Infinite Dimensional ODE

A simple special case of the reaction-diffusion-convection model (3.80) is the linear diffusion equation (the heat equation) in one spatial dimension, namely, the PDE

$$u_t = k^2 u_{xx} \tag{3.82}$$

where k^2 is the *diffusivity constant.* This differential equation can be used to model heat flow in an insulated bar. In fact, let us suppose that the bar is idealized to be the interval $[0,\ell]$ on the x-axis so that $u(x,t)$ represents the temperature of the bar at the point with coordinate x at time t. Moreover, because the bar has finite length, let us model the heat flow at the ends of the bar where we will consider just two possibilities: The bar is insulated at both ends such that we have the Neumann boundary conditions

$$u_x(0,t) = 0, \qquad u_x(\ell,t) = 0;$$

or, heat is allowed to flow through the ends of the bar, but the temperature at the ends is held constant at zero (in some appropriate units) such that we have the Dirichlet boundary conditions

$$u(0,t) = 0, \qquad u(\ell,t) = 0.$$

If one set of boundary conditions is imposed and an initial temperature distribution, say $x \mapsto u_0(x)$, is given on the bar, then we would expect that there is a unique scalar function $(x,t) \mapsto u(x,t)$, defined on the set $[0,\ell] \times [0,\infty)$ that satisfies the PDE, the initial condition $u(x,0) = u_0(x)$, and the boundary conditions. Of course, if such a solution exists, then for each $t > 0$, it predicts the corresponding temperature distribution $x \mapsto u(x,t)$ on the bar. In addition, if there is a solution of the boundary value problem corresponding to each initial temperature distribution, then we have a situation that is just like the phase flow of an ordinary differential equation. Indeed, let us consider a linear space $\mathcal{E}$ of temperature distributions on the rod and let us suppose that if a function $v : [0,\ell] \to \mathbb{R}$ is in $\mathcal{E}$, then there is a solution $(x,t) \mapsto u(x,t)$ of the boundary value problem with v as the initial temperature distribution such that $x \mapsto u(x,t)$ is a function in $\mathcal{E}$ whenever $t > 0$. In particular, all the functions in $\mathcal{E}$ must satisfy the boundary conditions. If this is the case, then we have defined a function $(0,\infty) \times \mathcal{E} \to \mathcal{E}$ given by $(t,v) \mapsto \varphi_t(v)$ such that $\varphi_0(v)(x) = v(x)$ and $(x,t) \mapsto \varphi_t(v)(x)$ is the solution of the boundary value problem with initial temperature distribution v. In other words, we have defined a dynamical system with "flow" φ_t whose phase space is the function space $\mathcal{E}$ of possible temperature distributions on the bar. For example, for the Dirichlet problem, we might take $\mathcal{E}$ to be the subset of $C^2[0,\ell]$ consisting of those functions that vanish at the ends of the interval $[0,\ell]$.

Taking our idea a step farther, let us define the linear transformation A on $\mathcal{E}$ by

$$Au = k^2 u_{xx}.$$

Then, the PDE (3.82) can be rewritten as

$$\dot{u} = Au, \tag{3.83}$$

an ordinary differential equation on the infinite dimensional space $\mathcal{E}$. Also, to remind ourselves of the boundary conditions, let us write $A = A_{\mathcal{N}}$ if Neumann boundary conditions are imposed and $A = A_{\mathcal{D}}$ for Dirichlet boundary conditions.

The linear homogeneous differential equation (3.83) is so simple that its solutions can be given explicitly. However, we will see how the general solution of the PDE can be found by treating it as an ordinary differential equation.

Let us begin by determining the rest points of the system (3.83). In fact, a rest point is a function $v : [0,\ell] \to \mathbb{R}$ that satisfies the boundary conditions and the second order ordinary differential equation $v_{xx} = 0$. Clearly, the only possible choices are affine functions of the form $v = cx + d$ where c and d are real numbers. There are two cases: For $A_{\mathcal{N}}$ we must have $c = 0$, but d is a free variable. Thus, there is a line in the function space $\mathcal{E}$ corresponding to the constant functions in $\mathcal{E}$ that consists entirely of rest points. For the Dirichlet case, both c and d must vanish and there is a unique rest point at the origin of the phase space.

Having found the rest points for the differential equation (3.83), let us discuss their stability. By analogy with the finite dimensional case, let us recall that we have discussed two methods that can be used to determine the stability of rest points: linearization and Lyapunov's direct method. In particular, for the finite dimensional case, the method of linearization is valid as long as the rest point is hyperbolic, and, in this case, the eigenvalues of the system matrix for the linearized system at the rest point determine its stability type.

Working formally, let us apply the method of linearization at the rest points of the system (3.83). Since this differential equation is already linear, we might expect the stability of these rest points to be determined from an analysis of the position in the complex plane of the eigenvalues of the system operator A. By definition, if λ is an eigenvalue of the operator $A_{\mathcal{D}}$ or $A_{\mathcal{N}}$, then there must be a nonzero function v on the interval $[0,\ell]$ that satisfies the boundary conditions and the ordinary differential equation

$$k^2 v_{xx} = \lambda v.$$

If v is an eigenfunction with eigenvalue λ, then we have that

$$\int_0^\ell k^2 v_{xx} v \, dx = \int_0^\ell \lambda v^2 \, dx. \tag{3.84}$$

Let us suppose that v is square integrable, that is,

$$\int_0^\ell v^2 \, dx < \infty$$

and also smooth enough so that integration by parts is valid. Then, equation (3.84) is equivalent to the equation

$$v_x v \Big|_0^\ell - \int_0^\ell v_x^2 \, dx = \frac{\lambda}{k^2} \int_0^\ell v^2 \, dx.$$

Thus, if either Dirichlet or Neumann boundary conditions are imposed, then the boundary term from the integration by parts vanishes, and therefore the eigenvalue λ must be a nonpositive real number.

For $A_{\mathcal{D}}$, if $\lambda = 0$, then there is no nonzero eigenfunction. If $\lambda < 0$, then the eigenvalue equation has the general solution

$$v(x) = c_1 \cos \alpha x + c_2 \sin \alpha x$$

where $\alpha := (-\lambda)^{1/2}/k$ and c_1 and c_2 are constants; and, in order to satisfy the Dirichlet boundary conditions, we must have

$$\begin{pmatrix} 1 & 0 \\ \cos \alpha\ell & \sin \alpha\ell \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

for some nonzero vector of constants. In fact, the determinant of the matrix must vanish, and we therefore have to impose the condition that $\alpha\ell$ is an integer multiple of π; or equivalently,

$$\lambda = -\left(\frac{n\pi k}{\ell}\right)^2$$

with a corresponding eigenfunction given by

$$x \mapsto \sin\frac{n\pi}{\ell}x$$

for each integer $n = 1, 2, \ldots, \infty$. By a similar calculation for $A_{\mathcal{N}}$, we have that $\lambda = 0$ is an eigenvalue with a corresponding eigenfunction $v \equiv 1$, and again the same real numbers

$$\lambda = -\left(\frac{n\pi k}{\ell}\right)^2$$

are eigenvalues, but this time with corresponding eigenfunctions

$$x \mapsto \cos\frac{n\pi}{\ell}x.$$

The nature of the real parts of the eigenvalues computed in the last paragraph and the principle of linearized stability suggest that the origin is an asymptotically stable rest point for the Dirichlet problem. On the other hand, the rest points of the Neumann problem seem to be of a different type: each of these rest points would appear to have a one-dimensional center manifold and an infinite dimensional stable manifold. All of these statements are true. But to prove them, certain modifications of the corresponding finite dimensional results are required. For example, the principle of linearized stability is valid for rest points of infinite dimensional ODE under the assumption that all points in the *spectrum* of the operator given by the linearized vector field at the rest point (in our case the operator A) have negative real parts that are bounded away from the imaginary axis in the complex plane (see, for example, [162, p. 114]). More precisely, the required hypothesis is that there is some number $\alpha > 0$ such that the real part of every point in the spectrum of the operator is less than $-\alpha$.

Recall that a complex number λ is in the spectrum of the linear operator A if the operator $A - \lambda I$ does not have a *bounded* inverse. Of course, if $v \neq 0$ is an eigenfunction with eigenvalue λ, then the operator $A - \lambda I$ is not injective and indeed λ is in the spectrum. In a finite dimensional space, if an operator is injective, then it is invertible. Hence, the only complex numbers in the spectrum of a finite dimensional linear operator are eigenvalues. However, in an infinite dimensional space, there can be points in the spectrum that are not eigenvalues (see [60]). For example, let us define the space $L^2(0, \ell)$ to be all (real) functions $v : [0, \ell] \to \mathbb{R}$ such that

$$\int_0^\ell v^2(x)\,dx < \infty \tag{3.85}$$

where we consider two such functions v and w to be equal if

$$\int_0^\ell (v(x) - w(x))^2\,dx = 0,$$

and consider the operator $B : L^2 \to L^2$ given by $(Bf)(x) \mapsto xf(x)$. This operator has no eigenvalues, yet the entire interval $[0, \ell]$ is in its spectrum. (Why?)

The operators $A_{\mathcal{D}}$ and $A_{\mathcal{N}}$, considered as operators defined in $L^2(0, \ell)$, have spectra that consist entirely of eigenvalues (pure point spectrum). However, to prove this claim we must first deal with the fact that these operators are not defined on all of L^2. After all, a square integrable function does not have to be differentiable. Instead, we can view our operators to be defined on the subset of L^2 consisting of those functions that have two derivatives both contained in L^2. Then, the claim about the spectra of $A_{\mathcal{D}}$ and $A_{\mathcal{N}}$ can be proved in two steps. First, if a complex number λ is not an eigenvalue, then for all $w \in L^2$ there is some function v that satisfies the boundary conditions and the differential equation

$$k^2 v_{xx} - \lambda v = w.$$

In other words, there is an operator $B : L^2 \to L^2$ given by $Bw = v$ such that $(A - \lambda I)Bw = w$. The fact that B is bounded is proved from the explicit construction of B as an integral operator. Also, it can be proved that $B(A - \lambda I)v = v$ for all v in the domain of A (see Exercise 3.29). Using these facts and the theorem on linearized stability mentioned above, it follows that the origin is an asymptotically stable rest point for the Dirichlet problem.

Exercise 3.29. Show that the spectrum of the operator in $L^2(0, \ell)$ given by $Av = v_{xx}$ with either Dirichlet or Neumann boundary conditions consists only of eigenvalues. Prove the same result for the operator $Av = av_{xx} + bv_x + cv$ where a, b, and c are real numbers.

In view of our results for finite dimensional linear systems, we expect that if we have a linear evolution equation $\dot{v} = Av$, even in an infinite dimensional phase space, and if $Aw = \lambda w$, then $e^{t\lambda}w$ is a solution. This is indeed the case for the PDE (3.82). Moreover, for linear evolution equations, we can use the principle of superposition to deduce that every linear combination of solutions of this type is again a solution. If we work formally, that is, without proving convergence, and if we use the eigenvalues and eigenvectors computed above, then the "general solution" of the Dirichlet problem is given by

$$u(x, t) = \sum_{n=1}^{\infty} e^{-(\frac{n\pi k}{\ell})^2 t} a_n \sin\frac{n\pi}{\ell}x,$$

and the general solution of the Neumann problem is given by

$$u(x,t) = \sum_{n=0}^{\infty} e^{-(\frac{n\pi k}{\ell})^2 t} b_n \cos \frac{n\pi}{\ell} x$$

where a_n and b_n are real numbers.

If the initial condition $u(x, 0) = u_0(x)$ is given, then, for instance, for the Dirichlet problem we must have that

$$u_0(x) = \sum_{n=1}^{\infty} a_n \sin \frac{n\pi}{\ell} x.$$

In other words, the initial function u_0 must be represented by a Fourier sine series. What does this mean? The requirement is that the Fourier sine series converges to u_0 in the space L^2 endowed with its natural norm,

$$\|v\| := \Big(\int_0^\ell v^2(x)\, dx \Big)^{1/2}.$$

In fact, the inner product space L^2 is a Hilbert space; that is, with respect to this norm, every Cauchy sequence in L^2 converges (see [156]). The precise requirement for u_0 to be represented by a Fourier sine series is that there are real numbers a_n and corresponding L^2 partial sums

$$\sum_{n=1}^{N} a_n \sin \frac{n\pi}{\ell} x$$

such that

$$\lim_{N\to\infty} \|u_0 - u_N\| = 0.$$

If the initial function u_0 is continuous, then for our special case the corresponding solution obtained by Fourier series also converges pointwise to a C^2 function that satisfies the PDE in the classical sense. We will show in a moment that this solution is unique, and therefore the special solutions of the PDE obtained from the eigenvalues and corresponding eigenfunctions do indeed form a fundamental set of solutions for our boundary value problems.

There are several ways to prove that solutions of the diffusion equation with a given initial condition are unique. We will use the "energy method"; an alternative uniqueness proof is based on the maximum principle (see Exercise 3.30). To show the uniqueness result, let us note that if two solutions of either the Dirichlet or Neumann boundary value problem satisfy the same initial condition, then the difference u of these two solutions is

a solution of the same boundary value problem but with initial value the zero function. Using an integration by parts, we also have the equality

$$\frac{d}{dt}\int_0^\ell \frac{1}{2} u^2\, dx = \int_0^\ell u_t u\, dx = k^2 \int_0^\ell u_{xx} u\, dx = -k^2 \int_0^\ell u_x^2\, dx.$$

It follows that the function

$$t \mapsto \int_0^\ell \frac{1}{2} u^2(x,t)\, dx$$

is not increasing, and therefore it is bounded above by its value at $t = 0$, namely,

$$\int_0^\ell \frac{1}{2} u^2(x,0)\, dx = 0.$$

The conclusion is that $u(x,t) \equiv 0$, as required.

Exercise 3.30. Prove the maximum principle: If $u_t(x,t) = k^2 u_{xx}(x,t)$ is a C^2 function on the open rectangle $(0,\ell) \times (0,T)$ and a continuous function on the closure of this rectangle, then the maximum of the function u is assumed either on the line $(0,\ell) \times \{0\}$ or on one of the lines

$$\{0\} \times [0,T], \qquad \{\ell\} \times [0,T].$$

Also, use the maximum principle to prove the uniqueness of solutions of the boundary value problem with initial condition for the diffusion equation. Hint: Use calculus (see [171, p. 41]).

Exercise 3.31. Solve the PDE (3.82) by the method of separation of variables; that is, assume that there is a solution of the form $u(x,t) = p(x)q(t)$, substitute this expression into the PDE, impose the boundary conditions, and determine the general form of the functions p and q.

Using the explicit form of the Fourier series representations of the general solutions of the heat equation with Dirichlet or Neumann boundary conditions, we can see that these solutions are very much like the solutions of a homogeneous linear ordinary differential equation: They are expressed as superpositions of fundamental solutions and they obviously satisfy the flow property $\varphi_s(\varphi_t(v)) = \varphi_{s+t}(v)$ as long as s and t are not negative (the series solutions do not necessarily converge for $t < 0$). Because of this restriction on the time variable, the solutions of our evolution equation are said to be *semi-flows* or *semi-groups*.

In the case of Dirichlet boundary conditions, if we look at the series solution, then we can see immediately that the origin is in fact globally asymptotically stable. For the Neumann problem there is a one-dimensional invariant manifold of rest points, and all other solutions are attracted exponentially fast to this manifold. Physically, if the temperature is held fixed at zero at the ends of the bar, then the temperature at each point of the bar approaches zero at an exponential rate, whereas if the bar is insulated at its ends, then the temperature at each point approaches the average value of the initial temperature distribution.

Our discussion of the scalar diffusion equation, PDE (3.82), has served to illustrate the fact that a (parabolic) PDE can be viewed as an ordinary differential equation on an infinite dimensional space. Moreover, as we have seen, if we choose to study a PDE from this viewpoint, then our experience with ordinary differential equations can be used to advantage as a faithful guide to its analysis.

Exercise 3.32. Verify the semi-flow property $\varphi_s(\varphi_t(v)) = \varphi_{s+t}(v)$ for the solutions of the scalar heat equation with Dirichlet or Neumann boundary conditions. Generalize this result to the equation $u_t = u_{xx} + f(u)$ under the assumption that every initial value problem for this equation has a local solution. Hint: How is the flow property proved for finite dimensional autonomous equations?

Let us now consider the nonlinear PDE

$$u_t = k^2 u_{xx} + f(u, x, t), \qquad 0 < x < \ell, \quad t > 0 \tag{3.86}$$

where f is a smooth function that represents a heat source in our heat conduction model.

To illustrate the analysis of rest points for a nonlinear PDE, let us assume that the source term f for the PDE (3.86) depends only on its first variable, and let us impose, as usual, either Dirichlet or Neumann boundary conditions. In this situation, the rest points are given by those solutions of the ordinary differential equation

$$k^2 u_{xx} + f(u) = 0 \tag{3.87}$$

that also satisfy the Dirichlet or Neumann boundary conditions.

The boundary value problem (3.87) is an interesting problem in ordinary differential equations. Let us note first that if we view the independent variable as "time," then the second order differential equation (3.87) is just Newton's equation for a particle of mass k^2 moving in a potential force field with force $-f(u)$. In addition, the corresponding first order system in the phase plane is the Hamiltonian system

$$\dot{u} = v, \qquad \dot{v} = -f(u)$$

whose total energy is given by

$$H(u, v) := \frac{k^2}{2} v^2 + F(u)$$

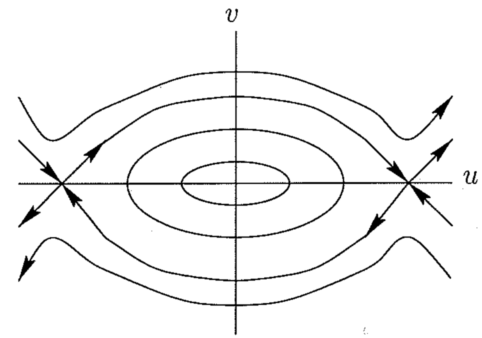

FIGURE 3.7. Phase portrait of the system $\dot{u} = v$ $\dot{v} = -u + u^3$.

where F, the potential energy, can be taken to be

$$F(u) := \int_0^u f(w)\, dw,$$

and, as we know, the phase plane orbits all lie on curves of constant energy. We will use these facts below.

A rest point of the PDE (3.86) with our special form of f and Dirichlet boundary conditions corresponds to a trajectory in the phase plane that starts on the v-axis and returns to the v-axis again exactly at time $x = \ell$. On the other hand, a rest point for the PDE with Neumann boundary conditions corresponds to a trajectory in the phase plane that starts on the u-axis and returns to the u-axis at time $x = \ell$.

Though the nonlinear boundary value problems that have just been described are very difficult in general, they can be "solved" in some important special cases. As an example, let us consider the following Dirichlet boundary value problem

$$u_t = u_{xx} + u - u^3, \qquad u(0, t) = 0, \quad u(\ell, t) = 0 \tag{3.88}$$

(see Exercise 3.35 for Neumann boundary conditions). Note first that the constant functions with values 0 or ± 1 are all solutions of the differential equation $u_{xx} + u - u^3 = 0$. However, only the zero solution satisfies the Dirichlet boundary conditions. Thus, there is exactly one constant rest point. Let us determine if there are any nonconstant rest points.

The phase plane system corresponding to the steady state equation for the PDE (3.88) is given by

$$\dot{u} = v, \qquad \dot{v} = -u + u^3.$$

It has saddle points at $(\pm 1, 0)$ and a center at $(0, 0)$. Moreover, the period annulus surrounding the origin is bounded by a pair of heteroclinic orbits that lie on the curve

$$\frac{1}{2}v^2 + \frac{1}{2}u^2 - \frac{1}{4}u^4 = \frac{1}{4}$$

(see Figure 3.7). Using this fact, it is easy to see that the interval $(0, 1/\sqrt{2})$ on the v-axis is a Poincaré section for the annulus of periodic orbits. Also, a glance at the phase portrait of the system shows that only the solutions that lie on these periodic orbits are candidates for nonconstant steady states for the PDE; they are the only periodic orbits in the phase plane that meet the v-axis at more than one point. Also, let us notice that the phase portrait is symmetric with respect to each of the coordinate axes. In view of this symmetry, if we define the period function

$$T : (0, \frac{1}{\sqrt{2}}) \to \mathbb{R} \tag{3.89}$$

so that $T(a)$ is the minimum period of the periodic solution starting at $u(0) = 0$, $v(0) = a$, then

$$u(\frac{1}{2}T(a)) = 0, \qquad v(\frac{1}{2}T(a)) = -a.$$

Hence, solutions of our boundary value problem that correspond to rest points for the PDE also correspond to periodic solutions whose half periods are exactly some integer submultiple of ℓ; equivalently, these solutions correspond to those real numbers a such that $0 < a < 1/\sqrt{2}$ and $T(a) = 2\ell/n$ for some positive integer n. In fact, each such a corresponds to exactly two rest points of the PDE; namely, $x \mapsto u(x)$ and $x \mapsto u(\ell - x)$ where $x \mapsto (u(x), v(x))$ is the phase trajectory such that $u(0) = 0$ and $v(0) = a$.

The number and position in the phase plane of all rest point solutions of the PDE can be determined from the following three propositions: (*i*) $T(a) \to 2\pi$ as $a \to 0^+$; (*ii*) $T(a) \to \infty$ as $a \to (1/\sqrt{2})^-$; and (*iii*) $T'(a) > 0$ (see Exercise 3.33). Using these facts, it follows that there is at most a finite number of rest points that correspond to the integers $1, 2, \dots, n$ such that $n < \ell/\pi$.

Exercise 3.33. Prove that the period function T given in display (3.89) has a positive first derivative. One way to do this is to find the explicit time-dependent periodic solutions of the first order system $\dot{u} = v$, $\dot{v} = -u + u^3$ using elliptic functions. For a different method, see [29] and [153].

Exercise 3.34. Find the rest points for the Dirichlet boundary value problem

$$u_t = u_{xx} + au - bu^2, \qquad u(x, 0) = 0, \quad u(x, \ell) = 0$$

(see [35]).

Are the rest points of the PDE (3.88) stable? It turns out that the stability problem for nonconstant rest points, even for our scalar PDE, is too difficult to describe here (see [162, p. 530]). However, we can say something about the stability of the constant rest point at the origin for the PDE (3.88). In fact, let us note that if $\ell < \pi$, then it is the only rest point. Moreover, its stability can be determined by linearization.

Let us first describe the linearization procedure for a PDE. The correct formulation is simple if we view the PDE as an ordinary differential equation on a function space. Indeed, we can just follow the recipe for linearizing an ordinary differential equation of the form $\dot{u} = g(u)$. Let us recall that if z is a rest point and g is a smooth function, then the linearization of the ordinary differential equation at z is

$$\dot{x} = Dg(z)(x - z),$$

or equivalently

$$\dot{w} = Dg(z)w$$

where $w := x - z$. Moreover, if the eigenvalues of $Dg(z)$ all have negative real parts, then the rest point z is asymptotically stable (see Section 2.3).

In order to linearize at a rest point of a PDE, let us suppose that the function $x \mapsto z(x)$ is a rest point for the PDE

$$u_t = g(u)$$

where $g(u) := u_{xx} + f(u)$ and $f : \mathbb{R} \to \mathbb{R}$ is a differentiable function. If the domain of $A_{\mathcal{D}}$ is viewed as the function space $C^2[0, \ell]$, then the function $g : C^2[0, \ell] \to C^0[0, \ell]$ is differentiable. This follows because the function $u \mapsto u_{xx}$ is linear and the function f is smooth. However, we have to be careful. In the definition of g we must view the notation $f(u)$ to mean $f \circ u$ where $u \in C^2[0, \ell]$. The difficulty is that the smoothness of the function $u \mapsto f \circ u$ depends on the topology of the function space to which u belongs (see Example 1.153).

Once we know that g is differentiable, its derivative can be easily computed as a directional derivative; in fact,

$$Dg(z)v = \frac{d}{dt}g(z + tv)\Big|_{t=0} = v_{xx} + Df(z)v.$$

Therefore, by definition, the linearized equation at the rest point z is given by

$$\dot{w} = w_{xx} + Df(z(x))w. \tag{3.90}$$

For a nonconstant rest point, the linearized equation (3.90) depends on the space variable x. The determination of stability in this case is often quite difficult—recall the stability analysis for periodic solutions of finite dimensional ordinary differential equations. For a constant rest point, the linearized equation has the form $\dot{w} = Aw$ where A is the linear operator given by $w \mapsto w_{xx} + Df(z)w$ for z a fixed number. In this case, as mentioned previously, it seems natural to expect the following result: If the spectrum of A is in the open left half plane and bounded away from the imaginary axis, then the rest point is asymptotically stable. In fact, this result, when properly interpreted, is true for the PDE (3.88). But to prove it, we have to specify the function space on which the spectrum is to be computed and recast the arguments used for ordinary differential equations in an infinite dimensional setting. For the PDE (3.88) the idea—derived from our study of ordinary differential equations—of applying the principle of linearized stability is justified, but some functional analysis is required to carry it out (see [162, Chapter 11]). However, our example is perhaps too simple; there are PDE where the linearized stability of a steady state can be easily proved, but the stability of the rest point is an open question. The problem for a general PDE of the form

$$u_t = Au + f(u)$$

is that the linear operator A, the function f, and the linearized operator $A + Df(z)$ must all satisfy additional hypotheses before the ODE arguments for the validity of the principle of linearized stability can be verified in the infinite dimensional case. This fact is an important difference between the theory of ordinary differential equations and the theory of PDE.

Let us put aside the theoretical justification of linearized stability and reconsider the rest point at the origin for the PDE (3.88) where the linearized system is given by

$$w_t = w_{xx} + w, \qquad w(0) = 0, \quad w(\ell) = 0.$$

In this case, the spectrum of the differential operator defined by

$$Aw = w_{xx} + w$$

consists only of eigenvalues (see Exercise 3.29). In fact, using the analysis of the spectrum of the operator $w \to w_{xx}$ given above, the spectrum of A is easily obtained by a translation. In fact, the spectrum is

$$\Big\{1 - \Big(\frac{n\pi}{\ell}\Big)^2 : n = 1, 2, \dots, \infty\Big\}.$$

Because

$$1 - \Big(\frac{n\pi}{\ell}\Big)^2 \le 1 - \Big(\frac{\pi}{\ell}\Big)^2,$$

the spectrum of A lies in the left half of the complex plane and is bounded away from the imaginary axis if and only if $1 < \pi^2/\ell^2$. Hence, using this

fact and assuming the validity of the principle of linearized stability, we have the following proposition: If $\ell < \pi$, then the origin is the only steady state and it is asymptotically stable.

Let us go one step further in our qualitative analysis of the PDE $u_t = u_{xx} + f(u)$ by showing that there are no periodic solutions. In fact, this claim is true independent of the choice of $\ell > 0$ and for an arbitrary smooth source function f. The idea for the proof, following the presentation in [162], is to show that there is a function (essentially a Lyapunov function) that decreases on orbits. In fact, let us define

$$E(u) = -\int_0^\ell \Big(\frac{1}{2}u(x)u_{xx}(x) + F(u(x))\Big)\,dx$$

where F is an antiderivative of f and note that

$$\dot{E} = -\int_0^\ell \Big(\frac{1}{2}u_t u_{xx} + \frac{1}{2}u u_{txx} + f(u)u_t\Big)\,dx.$$

After integration by parts twice for the integral of the second term in the integrand, and after imposing either Dirichlet or Neumann boundary conditions, it follows that

$$\dot{E} = -\int_0^\ell (u_{xx} + f(u))u_t\,dx = -\int_0^\ell (u_{xx} + f(u))^2\,dx.$$

Hence, except for the rest points, the time derivative of E is negative along orbits. In particular, there are no periodic orbits. Can the function E be used to give a proof of the stability of the rest point at the origin?

For the PDE (3.88) with $\ell < \pi$ we have now built up a rather complete picture of the phase portrait. In fact, we know enough to conjecture that there is a unique rest point that is globally asymptotically stable. Is this conjecture true?

Exercise 3.35. Analyze the existence of rest points, the stability types of constant rest points, and the phase portrait for the Neumann boundary value problem

$$u_t = u_{xx} + u - u^3, \qquad u_x(0, t) = 0, \quad u_x(\ell, t) = 0.$$

Note that there are three constant rest points. Use equation (3.90) to determine their stability types.

3.6.2 Galërkin Approximation

Since most differential equations, ODE or PDE, cannot be "solved," it is natural to seek approximate solutions. Of course, numerical methods are used all the time to obtain approximate values of state variables. However, in theory and practice the utility of approximation methods goes far beyond number crunching; for example, approximations are used to gain insight into the qualitative behavior of dynamical systems, to test computer codes, and to obtain existence proofs. Indeed, approximation methods are central elements of applied mathematics.

In this section we will take a brief look at a special case of *Galërkin's method,* one of the classic approximation methods for PDE. However, let us note that Galërkin's method is just one of an array of methods that are based on the idea of finding finite dimensional approximations of infinite dimensional dynamical systems. Many other methods are based on the idea of finding finite dimensional invariant (or approximately invariant) submanifolds in the infinite dimensional phase space. Of course, rest points and periodic orbits are finite dimensional invariant submanifolds. But these are only the simplest examples. In fact, let us note that a rest point or a periodic orbit might have an infinite dimensional stable manifold and a finite dimensional center manifold. In this case, the local dynamical behavior is determined by the dynamics on the center manifold because nearby orbits are attracted to the center manifold. An important generalization of this basic situation is the concept of an inertial manifold. By definition, an *inertial manifold* M is a finite dimensional submanifold in the phase space that has two properties: M is positively invariant, and every solution is attracted to M at an exponential rate (see [174]).

In general, if there is an attracting finite dimensional invariant manifold, then the dynamical system restricted to this invariant set is an ordinary differential equation that models the asymptotic behavior of the full infinite dimensional PDE. In particular, the ω-limit set of every solution lies on this manifold. Thus, the existence of such an invariant manifold provides the theoretical basis for a complete understanding of the infinite dimensional dynamical system using the techniques of ordinary differential equations. However, it is usually very difficult to prove the existence of attracting invariant manifolds. Moreover, even if an invariant manifold does exist, it is often very difficult to obtain the detailed specification of this manifold that would be required to reduce the original infinite dimensional dynamical system to an ordinary differential equation. As an alternative, an approximation method—such as Galërkin's method—that does not require the existence of an invariant manifold can often be employed with great success.

The following philosophical question seems to accompany all theoretical approximation methods for PDE "Is the set of reduced equations—presumably a system of nonlinear ordinary differential equations—easier to analyze than the original PDE?" In general, the answer to this question is clearly "no." However, if the finite dimensional approximation is "low dimensional" or of some special form, then often qualitative analysis is possible, and useful insights into the dynamics of the original system can be obtained. Perhaps the best "answer" to the question is to avoid the implied choice between infinite dimensional and finite dimensional analysis. The best approach to an applied problem is with a mind free of prejudice. Often several different methods, including physical thinking and numerical analysis, are required to obtain consistent and useful predictions from a model.

Let us begin our discussion of the Galërkin approximation method with an elementary, but key idea. Recall that a (real) vector space H is an inner product space if there is a bilinear form (denoted here by angle brackets) such that if $h \in H$, then $\langle h, h\rangle \geq 0$ and $\langle h, h\rangle = 0$ if and only if $h = 0$. It follows immediately that if $v \in H$ and $\langle v, h\rangle = 0$ for all $h \in H$, then $v = 0$. We will use this simple fact as the basis for solving equations in the space H. Indeed, suppose that we wish to find a solution of the (linear) equation

$$Au = b. \tag{3.91}$$

If there is a vector $u_0 \in H$ such that $\langle Au_0 - b, h\rangle = 0$ for all $h \in H$, then u_0 is a solution of the equation.

If we identify a subspace $S \subset H$ and find $u_S \in S$ such that

$$\langle Au_S - b, s\rangle = 0$$

for all $s \in S$, then u_S is called a *Galërkin approximation* of a solution of equation (3.91). Of course, every $h \in H$ is an "approximation" of a solution! The idea is to consider a sequence of subspaces, $S_1 \subset S_2 \subset \cdots$ that "converge" to H, and the corresponding Galërkin approximations $u_n \in S_n$ such that $\langle Au_n - b, s\rangle = 0$ for all $s \in S_n$. In this case, we might expect that the sequence $u_1, u_2, \ldots$ converges to a solution of the equation (3.91).

If H is a finite dimensional inner product space and the subspaces

$$S_1, S_2, S_3, \ldots$$

are strictly nested, then a corresponding sequence of Galërkin approximations is finite. Thus, we do not have to worry about convergence. However, if H is an infinite dimensional Hilbert space, then the approximating subspaces must be chosen with care in order to ensure the convergence of the sequence of Galërkin approximations.

Let us recall that a sequence $B = \{\nu_i\}_{i=1}^{\infty}$ of linearly independent elements in H is called a *Hilbert space basis* if the linear manifold S spanned by B—all finite linear combinations of elements in B—is dense in H; that is, if $h \in H$, then there is a sequence in S that converges to h in the natural norm defined from the inner product.

Suppose that H is a Hilbert space, $B = \{\nu_i\}_{i=1}^{\infty}$ is a Hilbert space basis for H, and $A : H \to H$ is a linear operator. Also, for each positive integer n let S_n denote the linear manifold spanned by the finite set $\{\nu_1, \ldots, \nu_n\}$. The *Galërkin principle* may be stated as follows: For each positive integer n, there is some $u_n \in S_n$ such that $\langle Au_n - b, s\rangle = 0$ for all $s \in S_n$. Moreover, the sequence $\{u_n\}_{n=1}^{\infty}$ converges to a solution of the equation $Au = b$.

The Galërkin principle is not a theorem! In fact, the Galërkin approximations may not exist or the sequence of approximations may not converge. The applicability of the method depends on the equation we propose to solve, the choice of the space H, and the choice of the basis B.

As an illustration of the Galërkin method applied to a PDE, let us consider the steady state equation

$$u_{xx} + f(x) = 0, \qquad 0 < x < \ell, \tag{3.92}$$

with either Dirichlet or Neumann boundary conditions where f is a smooth function. We will formulate a variational (weak) form for this boundary value problem. The basic idea is based on the fact that if u is a solution of the PDE (3.92), then

$$\int_0^\ell (u_{xx} + f)\phi \, dx = 0 \tag{3.93}$$

whenever ϕ is a square integrable function defined on $[0, \ell]$. In the Hilbert space $L^2(0, \ell)$ (see display (3.85)), the inner product of two functions v and w is

$$\langle v, w \rangle := \int_0^\ell v(x)w(x)\, dx.$$

Therefore, if u is a solution of the PDE, then equation (3.93) merely states that the inner product of ϕ with the zero function in L^2 vanishes. Moreover, if we define the operator $Au = u_{xx}$ and the function $b = f$, then $\langle Au - f, \phi \rangle = 0$ whenever ϕ is in the Hilbert space $L^2(0, \ell)$. Turning this analysis around, we can look for a function u such that $\langle Au - f, \phi \rangle = 0$ for all ϕ in L^2. Roughly speaking, in this case u is called a *weak solution* of the PDE. However, if we wish to apply the Galërkin method to the PDE (3.92), then we have to face the fact that although L^2 spaces are natural Hilbert spaces of functions, the elements in L^2 are not necessarily differentiable. In particular, the operator A is not defined on $L^2(0, \ell)$.

In which Hilbert space should we look for a solution? By asking this question, we free ourselves from the search for a *classical* or *strong* solution of the PDE (3.92), that is, a twice continuously differentiable function that satisfies the PDE and the boundary conditions. Instead, we will seek a *weak solution* by constructing a Hilbert space H whose elements are in L^2 such that a Galërkin formulation of our partial differential equation makes sense in H. If our boundary value problem has a classical solution, and we choose the Hilbert space H as well as the Galëkin formulation appropriately, then the L^2 equivalence class of the classical solution will also be in H. Moreover, if are fortunate, then the weak solution of the boundary value problem obtained by applying the Galërkin principle in H will be exactly the equivalence class of the classical solution.

To construct the appropriate Hilbert space of candidate solutions for the equation (3.93), let us first formally apply the *fundamental method for PDE*, namely, integration by parts, to obtain the identity

$$\int_0^\ell (u_{xx} + f)\phi \, dx = u_x \phi \Big|_0^\ell - \int_0^\ell (u_x \phi_x - f\phi)\, dx. \tag{3.94}$$

If the functions ϕ and u are sufficiently smooth so that the integration by parts is valid, then equation (3.93) is equivalent to an equation involving functions and only one of their derivatives with respect to the variable x, namely, the equation

$$\int_0^\ell u_x \phi_x \, dx - u_x \phi \Big|_0^\ell = \int_0^\ell f\phi \, dx. \tag{3.95}$$

In other words, to use equation (3.95) as a Galërkin formulation of our boundary value problem, it suffices to find a Hilbert space H whose elements have only one derivative with respect to x in L^2. Moreover, suppose that such a Hilbert space H exists. If we find a function $u \in H$ such that equation (3.95) holds for all $\phi \in H$ and *u happens to be smooth*, then the integration by parts is valid and we also have a solution of equation (3.93) for all smooth functions ϕ. Using this fact, it is easy to prove that u satisfies the PDE (3.92) pointwise, that is, u is a classical solution (see Exercise (3.36)).

Exercise 3.36. Suppose that u is a C^2 function. If equation (3.93) holds for every $\phi \in C^\infty$, then prove that $u_{xx} + f(x) = 0$.

If Dirichlet boundary conditions are imposed, then the boundary conditions must be build into the Hilbert space H of test functions from which we select ϕ. In other words, we must impose the condition that the test functions satisfy the Dirichlet boundary conditions. The appropriate Hilbert space is denoted $H_0^1(0, \ell)$. To define it, let us first define the Sobolev norm for a smooth function ϕ as follows:

$$\|\phi\|_1 := \Big(\int_0^\ell \phi^2(x)\, dx\Big)^{1/2} + \Big(\int_0^\ell \phi_x^2(x)\, dx\Big)^{1/2}.$$

The subscript on the norm indicates that one derivative of ϕ is in L^2. The definition of the Sobolev norms with n derivatives taken into account is similar. Also, note that the Sobolev norm is just the sum of the L^2 norms of ϕ and its first derivative. The Sobolev space $H_0^1(0, \ell)$ is defined to be the completion, with respect to the Sobolev norm, of the set of all smooth functions that satisfy the Dirichlet boundary conditions and have a finite Sobolev norm; informally, "the space of functions with one derivative in L^2."

Using the Sobolev space $H_0^1(0,\ell)$, we have the following Galërkin or weak formulation of our Dirichlet boundary value problem: Find $u \in H_0^1(0,\ell)$ such that

$$(u,\phi) := \int_0^\ell u_x\phi_x\,dx = \int_0^\ell f\phi\,dx = \langle f,\phi\rangle \tag{3.96}$$

for all $\phi \in H_0^1(0,\ell)$. If u is a weak solution of the Dirichlet boundary value problem, then, using the definition of the Sobolev space, we can be sure that u is the limit of smooth functions that satisfy the boundary conditions. However, u itself is only defined abstractly as an equivalence class, thus it only satisfies the boundary conditions in the generalized sense, that is, u is the limit of a sequence of functions that satisfy the boundary conditions.

For the Neumann boundary value problem, again using equation (3.94), the appropriate space of test functions is $H^1(0,\ell)$, the space defined just like H_0^1 except that no boundary conditions are imposed. This requires a bit of explanation. First, we have the formal statement of the weak formulation of the Neumann problem: Find a function u in $H^1(0,\ell)$ such that, with the same notation as in display (3.96),

$$(u,\phi) = \langle f,\phi\rangle$$

for all $\phi \in H^1(0,\ell)$. We will show the following proposition: *If u is smooth enough so that the integration by parts in display (3.94) is valid and the equivalence class of u in $H^1(0,\ell)$ is a weak solution of the Neumann problem, then u satisfies the Neumann boundary conditions.* In fact, if $\phi \in H_0^1(0,\ell)$, then ϕ is a limit of smooth functions that satisfy the Dirichlet boundary conditions. Thus, if we use integration by parts for a sequence of smooth functions converging to ϕ in $H_0^1(0,\ell)$ and pass to the limit, then we have the identity

$$-\int_0^\ell u_{xx}\phi\,dx = \int_0^\ell f\phi\,dx$$

for all $\phi \in H_0^1(0,\ell)$. In other words, $u_{xx} + f(x)$ is the zero element of $H_0^1(0,\ell)$. By Exercise (3.37), the space $H_0^1(0,\ell)$ is a dense subspace of $H^1(0,\ell)$. Thus, it is easy to see that the identity

$$-\int_0^\ell u_{xx}\phi\,dx = \int_0^\ell f\phi\,dx$$

holds for all $\phi \in H^1(0,\ell)$. Finally, by this identity, the boundary term in the integration by parts formula in display (3.94) must vanish for each $\phi \in H^1(0,\ell)$. This fact clearly implies that u satisfies the Neumann boundary conditions, as required. Hence, our weak formulation is consistent with the classical boundary value problem: If a weak solution of the Neumann boundary value problem happens to be smooth, then it will satisfy the Neumann boundary conditions.

Exercise 3.37. Prove that $H_0^1(0,\ell)$ is a dense subspace of $H^1(0,\ell)$.

Our analysis leads to the natural question "If a weak solution exists, then is it automatically a strong (classical) solution?" The answer is "yes" for the example problems that we have formulated here, but this important "regularity" result is beyond the scope of our discussion. Let us simply remark that the regularity of the weak solution depends on the form of the PDE. It is also natural to ask if our *weak* boundary value problems have solutions. The answer is in the affirmative. In fact, the relevant theory is easy to understand. We will formulate and prove a few of its basic results.

Let us suppose that H is a real Hilbert space, that $(\ ,\ )$ is a bilinear form on H (it maps $H \times H \to \mathbb{R}$), $\langle\ ,\ \rangle$ is the inner product on H, and $\|\ \| := \langle\ ,\ \rangle^{1/2}$ is the natural norm. The bilinear form is called *continuous* if there is a constant $a > 0$ such that

$$|(u,v)| \le a\|u\|\|v\|$$

for all $u, v \in H$. The bilinear form is called *coercive* if there is a constant $b > 0$ such that

$$(u,u) \ge b\|u\|^2$$

for all $u \in H$.

Theorem 3.38 (Lax–Milgram). *Suppose that H is a real Hilbert space and $(\ ,\ )$ is a continuous and coercive bilinear form on H. If F is a bounded linear functional $F : H \to \mathbb{R}$, then there is a unique $u \in H$ such that*

$$(u,\phi) = F(\phi)$$

for every $\phi \in H$. Moreover,

$$\|u\| \le \frac{1}{b}\|F\|.$$

Proof. The main tool of the proof is a standard result in Hilbert space theory, the Riesz representation theorem: If F is a bounded linear functional, then there is a unique $f \in H$ such that $F(\phi) = \langle f,\phi\rangle$ for every $\phi \in H$ (see [156]). In particular, this is true for the functional F in the statement of the theorem.

If $u \in H$, then the function given by $\phi \mapsto (u,\phi)$ is a linear functional on H. To see that this functional is bounded, use the continuity of the bilinear form to obtain the estimate

$$|(u,\phi)| \le a\|u\|\|\phi\|$$

and note that $\|u\| < \infty$. The Riesz theorem now applies to each such functional. Therefore, there is a function $A : H \to H$ such that

$$(u, \phi) = \langle Au, \phi\rangle$$

for all $\phi \in H$. Moreover, using the linearity of the bilinear form, it follows that A is a linear transformation.

It is now clear that the equation in the statement of the theorem has a unique solution if and only if the equation $Au = f$ has a unique solution for each $f \in H$.

By the continuity and the coerciveness of the bilinear form, if $u, v, \phi \in H$, then

$$\langle A(u-v), \phi\rangle = (u-v, w) \le a\|u-v\|\|\phi\|, \tag{3.97}$$

$$\langle A(u-v), u-v\rangle = (u-v, u-v) \ge b\|u-v\|^2. \tag{3.98}$$

Also, by the Schwarz inequality, we have that

$$\sup_{\|\phi\|\le 1} |\langle v, \phi\rangle| \le \|v\|,$$

and, for $\phi := (1/\|v\|)v$, this upper bound is attained. Thus, the norm of the linear functional $\phi \mapsto \langle w, \phi\rangle$ is $\|w\|$. In particular, using the inequality (3.97), we have

$$\|Au - Av\| = \sup_{\|w\|\le 1} \langle A(u-v), \phi\rangle \le a\|u-v\|. \tag{3.99}$$

Define the family of operators $\mathcal{A}^\lambda : H \to H$ by

$$\mathcal{A}^\lambda \phi = \phi - \lambda(A\phi - f), \qquad \lambda > 0,$$

and note that $\mathcal{A}^\lambda u = u$ if and only if $Au = f$. Thus, to solve the equation $Au = f$, it suffices to show that for at least one choice of $\lambda > 0$, the operator $\mathcal{A}^\lambda$ has a unique fixed point.

By an easy computation using the definition of the norm, equation (3.97), the Schwarz inequality, and equation (3.99), we have that

$$\|\mathcal{A}^\lambda u - \mathcal{A}^\lambda v\|^2 = (1 - 2\lambda a + \lambda^2 a^2)\|u-v\|^2.$$

Note that the polynomial in λ vanishes at $\lambda = 0$ and that its derivative at this point is negative. It follows that there is some $\lambda > 0$ such that the corresponding operator is a contraction on the complete metric space H. By the contraction mapping theorem, there is a unique fixed point $u \in H$. Moreover, for this u we have proved that $(u, u) = F(u)$. Therefore,

$$\|u\|\|F\| \ge \langle f, u\rangle \ge b\|u\|^2,$$

and the last statement of the theorem follows. □

The Lax–Milgram theorem is a classic result that gives us a "hunting license" to seek weak solutions for our boundary value problems. One way to construct a solution is to use the Galërkin method described above. In fact, with the previously defined notation, let us consider one of the finite dimensional Hilbert spaces S_n of H, and note that by the Lax–Milgram theorem there is a unique $u_n \in S_n$ such that

$$(u_n, s) = \langle f, s\rangle \tag{3.100}$$

for all $s \in S_n$ with the additional property that

$$\|u_n\| \le \|f\|. \tag{3.101}$$

The Galërkin principle is the statement that the sequence $\{u_n\}_{n=1}^\infty$ converges to the unique solution u of the weak boundary value problem. The approximation u_n can be expressed as a linear combination of the vectors $\nu_1, \ldots, \nu_n$ that, by our choice, form a basis of the subspace S_n. Thus, there are real numbers $c_1, \ldots, c_n$ such that

$$u_n = \sum_{j=1}^n c_j \nu_j.$$

Also, each element $s \in S_n$ is given in coordinates by

$$s = \sum_{i=1}^n s_i \nu_i.$$

Thus, the equation (3.100) is given in coordinates by the system of equations

$$\sum_{j=1}^n c_j(\nu_j, \nu_i) = \langle f, \nu_i\rangle, \qquad i = 1, \ldots n,$$

or, in the equivalent matrix form for the unknown vector $(c_1, \ldots c_n)$, we have the equation

$$S\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} \langle f, \nu_1\rangle \\ \vdots \\ \langle f, \nu_n\rangle \end{pmatrix}$$

where S, called the *stiffness matrix*—the terminology comes from the theory of elasticity—is given by $s_{ij} := (\nu_j, \nu_i)$. Of course, by the Lax–Milgram theorem, S is invertible and the matrix system can be solved to obtain the approximation u_n.

Does the sequence of approximations $\{u_n\}_{n=1}^\infty$ converge? The first observation is that, by the inequality (3.101), the sequence of approximates is bounded. Let u be the weak solution given by the Lax–Milgram theorem. Subtract the equality $(u_n, s) = \langle f, s\rangle$ from the equality $(u, s) = \langle f, s\rangle$ to see that

$$(u - u_n, s) = 0 \tag{3.102}$$

for all $s \in S_n$. Also, using the coerciveness of the bilinear form, if $\phi \in S_n$, then

$$\begin{aligned} b\|u - u_n\|^2 &\le (u - u_n, u - u_n) = (u - u_n, u - u_n + \phi - \phi) \\ &= (u - u_n, \phi - u_n) + (u - u_n, u - \phi). \end{aligned}$$

By equation (3.102) and the fact that both u_n and ϕ are in S_n, we have the inequality

$$b\|u - u_n\|^2 \le (u - u_n, u - \phi) \le a\|u - u_n\| \|u - \phi\|.$$

It follows that

$$\|u - u_n\| \le \frac{a}{b}\|u - \phi\| \tag{3.103}$$

for all $\phi \in S_n$.

Recall that the linear span of the sequence $\{\nu_j\}_{j=1}^{\infty}$ is assumed to be dense in H. Hence, for each $\epsilon > 0$ there is some integer m and constants $c_1, \ldots, c_m$ such that

$$\|u - \sum_{j=1}^{m} c_j \nu_j\| < \epsilon.$$

If we set $n = m$ and $v = \sum_{j=1}^{m} c_j \nu_j$ in the inequality (3.103), then

$$\|u - u_n\| \le \frac{a}{b}\epsilon.$$

In other words, the sequence of Galërkin approximations converges to the weak solution, as required.

In the context of the steady state problem with which we started, namely, the PDE (3.92), the Lax–Milgram theorem applies (see Exercise 3.39). If, for example, we consider Dirichlet boundary conditions, the bilinear form

$$(u, v) = \int_0^{\ell} u_x v_x \, dx$$

in H_0^1, and

$$\nu_j(x) := \sin \frac{j\pi}{\ell} x, \qquad f(x) = \sum_{j=1}^{\infty} f_j \sin \frac{j\pi}{\ell} x,$$

then the Galërkin approximation is easily computed to be

$$u_n(x) = \sum_{i=1}^{n} \left(\frac{L}{i\pi}\right)^2 f_i \sin \frac{i\pi}{\ell} x, \tag{3.104}$$

exactly the partial sum of the usual Fourier series approximation (see Exercise 3.40).

Exercise 3.39. Prove that the bilinear form

$$(u, v) = \int_0^{\ell} u_x v_x \, dx$$

is continuous and coercive on the spaces H_0^1 and H^1.

Exercise 3.40. Find the stiffness matrix for the Galërkin approximation for the PDE (3.92) with Dirichlet boundary conditions using the basis given by

$$\nu_j(x) := \sin \frac{j\pi}{\ell} x, \qquad j = 1, 2, \ldots, \infty$$

for H_0^1, and verify the approximation (3.104). Also, consider the PDE (3.92) with Neumann boundary conditions, and find the Galërkin approximations corresponding to the basis

$$1, \cos \frac{\pi x}{\ell}, \sin \frac{\pi x}{\ell}, \ldots .$$

We have now seen one very simple example where the Galërkin principle can be turned into a theorem. Let us take this as a prototype argument to justify the Galërkin principle. However, our main objective in this section is to see how the Galërkin method leads to problems in ordinary differential equations. For this, let us consider first the PDE

$$u_t = u_{xx} + f(x, t), \qquad 0 < x < \ell, \quad t > 0 \tag{3.105}$$

with either Dirichlet or Neumann boundary conditions, and let us work formally.

The weak form of our boundary value problem is derived from the integration by parts formula

$$\int_0^{\ell} (u_t - u_{xx} - f(x,t))\phi \, dx = \int_0^{\ell} (u_t\phi + u_x\phi_x - f(x,t))\phi \, dx - u_x\phi\Big|_0^{\ell}.$$

Just as before, we can formulate two weak boundary value problems.

The Dirichlet Problem: Find $u(x, t)$, a family of functions in $H_0^1(0, \ell)$ such that

$$\int_0^{\ell} (u_t\phi + u_x\phi_x) \, dx = \int_0^{\ell} f\phi \, dx$$

for all $\phi \in H_0^1(0,\ell)$.

The Neumann Problem: Find $u(x,t)$, a family of functions in $H^1(0,\ell)$ with the same integral condition satisfied for all $\phi \in H^1(0,\ell)$.

To apply the Galërkin method, choose $\nu_1, \nu_2, \ldots$ a linearly independent sequence whose span is dense in the Hilbert space $H_0^1(0,\ell)$ or $H^1(0,\ell)$, and define the finite dimensional spaces S_n as before. The new wrinkle is that we will look for an approximate solution in the subspace S_n of the form

$$u_n(x,t) = \sum_{j=1}^{n} c_j(t)\nu_j(x)$$

where the coefficients are differentiable functions of time. According to the Galërkin principle, let us search for the unknown functions $c_1, \ldots, c_n$ so that we have $(u_n, s) = \langle f, s\rangle$ for all $s \in S_n$. Expressed in coordinates, the requirement is that the unknown functions satisfy the system of n ordinary differential equations

$$\sum_{j=1}^{n} c_j'(t)\int_0^\ell \nu_j\nu_i\,dx + \sum_{j=1}^{n} c_j(t)\int_0^\ell (\nu_j)_x(\nu_i)_x\,dx = \int_0^\ell f\nu_i\,dx$$

indexed by $i = 1, \ldots, n$. In matrix form, we have the linear system of ordinary differential equations

$$MC' + SC = F(t)$$

where M, given by

$$M_{ij} := \int_0^\ell \nu_j\nu_i\,dx$$

is called the *mass matrix*, S, given by

$$S_{ij} := \int_0^\ell (\nu_j)_x(\nu_i)_x\,dx$$

is the stiffness matrix, and $C := (c_1, \ldots, c_n)$. If the initial condition for the PDE (3.105) is $u(x,0) = u_0(x)$, then the usual choice for the initial condition for the approximate system of ordinary differential equations is the element $u_0^n \in S_n$ such that

$$\langle u_0^n, s\rangle = \langle u_0, s\rangle$$

for all $s \in S_n$. This "least squares" approximation always exists. (Why?)

We have, in effect, described some aspects of the theoretical foundations of the finite element method for obtaining numerical approximations of PDE (see [170]). But a discussion of the techniques that make the finite element method a practical computational tool is beyond the scope of this book.

The Galërkin method was originally developed to solve problems in elasticity. This application yields some interesting dynamical problems for the corresponding systems of ordinary differential equations. Let us consider, for instance, the PDE (more precisely the *integro-PDE*),

$$u_{xxxx} + \Big(\alpha - \beta\int_0^1 u_x^2\,dx\Big)u_{xx} + \gamma u_x + \delta u_t + \epsilon u_{tt} = 0$$

that is derived in the theory of aeroelasticity as a model of panel flutter (see for example the book of Raymond L. Bisplinghoff and Holt Ashley [22, p. 428] where the physical interpretation of this equation and its parameters are given explicitly). We note in passing that this reference is full of Galërkin approximations, albeit Galërkin approximations of linearized equations. In fact, Galërkin approximations are commonplace in the theory of aeroelasticity.

At any rate, let us take the boundary conditions

$$u(0,t) = u(1,t) = 0, \qquad u_{xx}(0,t) = u_{xx}(1,t) = 0$$

given for this equation for "simply supported" panel edges. Of course, $u(x,t)$ represents the deflection of the panel. If we take just the first Fourier mode, that is, the Galërkin approximation with trial function

$$u_1(x,t) = c(t)\sin \pi x,$$

then we obtain the equation

$$\epsilon\ddot{c} + \delta\dot{c} + \pi^2(\pi^2 - \alpha)c + \frac{\pi^4}{2}\beta c^3 = 0. \tag{3.106}$$

Let us note that if $\pi^2 - \alpha < 0$, then this Galëkin approximation is a form of Duffing's equation with damping. We have already developed some of the tools needed to analyze this equation. In fact, most solutions are damped oscillations whose ω-limits are one of two possible asymptotically stable rest points (see Exercise 3.41). However, if a periodic external force is added to this system, then very complex dynamics are possible (see [96] and Chapter 6).

Exercise 3.41. Draw representative phase portraits for the family of differential equations (3.106). How does the phase portrait depend on the choice of the parameters?

Exercise 3.42. Consider the basis functions

$$\nu_j(x) := \sin(j\pi x/\ell)$$

for $H_0^1(0,\ell)$. Find the mass matrix and the stiffness matrix for the Galërkin approximations for the weak Dirichlet boundary value problem (3.105) with $f(x,t) := \sin(\pi x/\ell)\cos\omega t$. Solve the corresponding system of linear differential equations for the nth approximation $u_n(x,t)$. What can you say qualitatively about the solutions of the Galërkin approximations? What long term dynamical behavior of the PDE (3.105) is predicted by the Galërkin approximations? Find a steady state solution? Repeat the analysis for $f(x,t) = \cos\omega t$. Do you see a problem with our formal computations? Formulate and solve analogous problems for Neumann boundary conditions.

Exercise 3.43. Consider a two (Fourier) mode Galërkin approximation for the PDE

$$u_t = k^2 u_{xx} + u - u^3 + a\cos\omega t, \qquad 0 < x < \ell, \quad t > 0$$

with either Dirichlet or Neumann boundary conditions. What is the "general character" of the solutions in the phase plane? Start, for example, with the case where there is a time-independent source term ($a = 0$) and consider the stability of the steady state solution of the PDE at $u \equiv 0$. Is the (linearized) stability criterion for the PDE reflected in the stability of the corresponding rest point in the phase plane of the approximating ordinary differential equation? Is the ω-limit set of every solution of the approximation a rest point?

3.6.3 Traveling Waves

The concept of traveling wave solutions will be introduced in this section for the classic model system

$$u_t = k^2 u_{xx} + au(1-u), \qquad x \in \mathbb{R}, \quad t > 0 \tag{3.107}$$

where k and $a > 0$ are constants.

The PDE (3.107), often called *Fisher's equation,* can be used to model many different phenomena. For example, this equation is a model of logistic population growth with diffusion ([67], [132]), and it is also a model of neutron flux in a nuclear reactor (see [140]). For a general description of this and many other models of this type see [132] and [140].

Let us begin with the observation that equation (3.107) can be rescaled to remove the explicit dependence on the system parameters. In fact, with respect to the new time and space variables

$$\tau = kt, \qquad \xi = x\Big(\frac{a}{k}\Big),$$

equation (3.107) can be recast in the form

$$u_\tau = u_{\xi\xi} + u(1-u).$$

Therefore, with no loss of generality, we will consider the original model equation (3.107) for the case $a = 1$ and $k = 1$.

The basic idea is to look for a solution of equation (3.107) in the form of a *traveling wave*, that is,

$$u(x,t) = U(x - ct)$$

where the wave form is given by the function $U : \mathbb{R} \to \mathbb{R}$ and where the wave speed is $|c| \neq 0$. For definiteness and simplicity, let us assume that $c > 0$. If we substitute the *ansatz* into Fisher's equation, we obtain the second order nonlinear ordinary differential equation

$$\ddot{U} + c\dot{U} + U - U^2 = 0$$

that is equivalent to the phase plane system

$$\dot{U} = V, \qquad \dot{V} = -U - cV + U^2. \tag{3.108}$$

All solutions of the system (3.108) correspond to traveling wave solutions of Fisher's equation. However, for applications, the traveling wave solutions must satisfy additional properties. For example in biological applications, u represents a population. Thus, to be physically meaningful, we must have $u \geq 0$.

In the original model equation, if there is no diffusion, then the model reduces to the one-dimensional ordinary differential equation for logistic growth $\dot{u} = u - u^2$ where there is an unstable rest point at $u = 0$, a stable rest point at $u = 1$, and a connecting orbit, that is, an orbit with α-limit set $\{0\}$ and ω-limit set $\{1\}$.

Is there a traveling wave solution u for the PDE (3.107) such that $0 < u(x,t) < 1$, and

$$\lim_{t\to\infty} u(x,t) = 1, \qquad \lim_{t\to-\infty} u(x,t) = 0?$$

In other words, is there an orbit—for the PDE viewed as an infinite dimensional ordinary differential equation—connecting the steady states $u \equiv 0$ and $u \equiv 1$ as in the case of the one-dimensional logistic model? An answer to this question is given by the following proposition.

Proposition 3.44. *There is a traveling wave solution $(x,t) \mapsto u(x,t)$ whose orbit connects the steady states $u \equiv 0$ and $u \equiv 1$ with $0 < u(x,t) < 1$ if and only if $c \geq 2$.*

Proof. Note that the solution $u(x,t) = U(x - ct)$ is a connecting orbit if $0 < U(s) < 1$, and

$$\lim_{s\to\infty} U(s) = 0, \qquad \lim_{s\to-\infty} U(s) = 1.$$

The system matrix of the linearized phase plane equations (3.108) at the origin has eigenvalues

$$\frac{1}{2}(-c \pm \sqrt{c^2 - 4}),$$

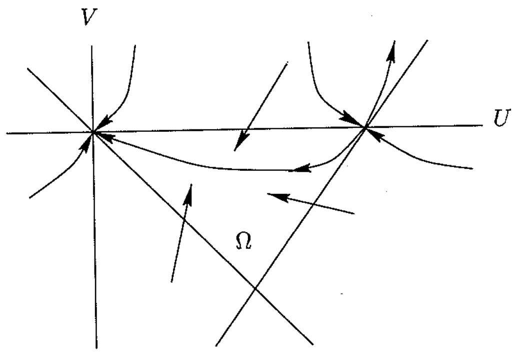

FIGURE 3.8. The invariant region Ω for the system (3.108) in case $c \geq 2$.

and its eigenvalues at the point $(1,0)$ are given by

$$\frac{1}{2}(-c \pm \sqrt{c^2+4}).$$

Therefore, if $c > 0$, then there is a hyperbolic sink at the origin and a hyperbolic saddle at the point $(1,0)$. Moreover, if a connecting orbit exists, then the corresponding phase plane solution $s \mapsto (U(s), V(s))$ must be on the unstable manifold of the saddle and the stable manifold of the sink.

Note that if $c < 2$, then the sink at the origin is of *spiral* type. Hence, even if there is a connecting orbit in this case, the corresponding function U cannot remain positive.

Assume that $c \geq 2$ and consider the lines in the phase plane given by

$$V = \frac{1}{2}\big(-c+\sqrt{c^2-4}\big)U, \qquad V = \frac{1}{2}\big(-c+\sqrt{c^2+4}\big)(U-1). \tag{3.109}$$

They correspond to eigenspaces at the rest points. In particular, the second line is tangent to the unstable manifold of the saddle point at $(U,V) = (1,0)$. The closed triangular region Ω (see Figure 3.8) in the phase plane bounded by the lines (3.109) and the line given by $V = 0$ is positively invariant. In fact, the phase plane vector field points into this region at every point on the boundary of Ω except the rest points. This fact is easily checked by computing the dot product of the vector field with the appropriate normal fields along the lines. In fact, *along the lines (3.109)*, we have

$$\begin{aligned} \dot V - \frac{1}{2}(-c+\sqrt{c^2-4})\dot U &= U^2 \geq 0, \\ \dot V - \frac{1}{2}(-c+\sqrt{c^2+4})\dot U &= (U-1)^2 \geq 0, \end{aligned} \tag{3.110}$$

and $\dot V = -U + U^2$ is negative for $0 < U < 1$.

Suppose (as we will soon see) that the unstable manifold at the saddle intersects the region Ω. Then a solution that starts on this portion of the unstable manifold must remain in the region Ω for all positive time. Thus, the ω-limit set of the corresponding orbit is also in Ω. Because $\dot U \leq 0$ in Ω, there are no periodic orbits in Ω and no rest points in the interior of Ω. By the Poincaré–Bendixson theorem, the ω-limit set must be contained in the boundary of Ω. In fact, this ω-limit set must be the origin.

To complete the proof, we will show that the unstable manifold at the saddle has nonempty intersection with the interior of Ω. To prove this fact, let us first recall that the unstable manifold is tangent to the line given by the second equation in display (3.109). We will show that the tangency is quadratic and that the unstable manifold lies "above" this line. Our proof of this fact is more complicated than is necessary. However, the method used can be generalized.

In the new coordinates given by

$$Z = U - 1, \qquad W = V,$$

the saddle rest point is at the origin for the equivalent first order system

$$\dot Z = W, \qquad \dot W = Z - cW + Z^2.$$

The additional change of coordinates

$$Z = P, \qquad W = Q + \alpha P := Q + \frac{1}{2}(-c+\sqrt{c^2+4})P$$

transforms the system so that the unstable manifold of the saddle point is tangent to the horizontal P-axis. We will show that the unstable manifold is above this axis in some neighborhood of the origin; it then follows from the second formula in display (3.110) that the unstable manifold lies above the P-axis globally.

Note that the unstable manifold is given, locally at least, by the graph of a smooth function $Q = h(P)$ with $h(0) = h'(0) = 0$. Since this manifold is invariant, we must have that $\dot Q = h'(P)\dot P$, and therefore, by an easy computation,

$$P^2 - (c+\alpha)h(P) = h'(P)(h(P) + \alpha P). \tag{3.111}$$

The function h has the form $h(P) = \beta P^2 + O(P^3)$. By substitution of this expression into equation (3.111), we obtain the inequality

$$\beta = (3\alpha + c)^{-1} > 0,$$

as required. □

Much more can be said about the traveling wave solutions that we have just found. A surprising fact is that all orbits of the PDE (3.107) starting with physically realistic initial conditions have as their ω-limit set the traveling wave solution with wave speed $c = 2$. This fact was proved by Andrei N. Kolmogorov, Ivan G. Petrovskii, and Nikolai S. Piskunov [103] (see also [15] and [21]). For a detailed mathematical account of traveling wave solutions see the book of Paul C. Fife [67] and also [132] and [162].

Exercise 3.45. Show that the PDE

$$u_t - u^2 u_x = u_{xx} + u, \qquad x \in \mathbb{R}, \quad t \geq 0$$

has a nonconstant solution that is periodic in both space and time.

3.6.4 First Order PDE

Consider the special case of the model equation (3.80) where there is no diffusion, but the medium moves with velocity field V; that is, consider the differential equation

$$u_t + \gamma \operatorname{grad} u \cdot V = f. \tag{3.112}$$

This is an important example of a first order partial differential equation. Other examples are equations of the form

$$u_t + (f(u))_x = 0,$$

called *conservation laws* (see [162]), and equations of the form

$$S_t + H(S_q, q, t) = 0,$$

called *Hamilton–Jacobi equations* (see [10]). We will show how these PDE can be solved using ordinary differential equations.

Let us consider the case of one space variable; the general case is similar. If $\gamma = 1$, then the equation (3.112) is given by

$$u_t + v(x,t)u_x = f(u,x,t),$$

or, with a redefinition of the names of the functions, it has the more general form

$$f(x,y,u)u_x + g(x,y,u)u_y = h(x,y,u). \tag{3.113}$$

We will "solve" the PDE (3.113) using the following basic idea: If the graph G of a function $z = u(x,y)$ is an invariant manifold for the first order system

$$\dot{x} = f(x,y,z), \quad \dot{y} = g(x,y,z), \quad \dot{z} = h(x,y,z), \tag{3.114}$$

then u is a solution of the PDE (3.113). Indeed, using the results of Section 1.7 and the fact that

$$(x,y) \mapsto (x, y, u(x,y), u_x(x,y), u_y(x,y), -1)$$

is a normal vector field on G, it follows that the manifold G is invariant if and only if the dot product of the vector field associated with the system (3.114) and the normal field is identically zero; that is, if and only if equation (3.113) holds. The orbits of the system (3.114) are called *characteristics* of the PDE (3.113).

Perhaps it is possible to find an invariant manifold for the first order system (3.114) by an indirect method. However, we can also construct the invariant manifold directly from appropriate initial data. To see how this is done, let us suppose that we have a curve in space given by $\gamma : \mathbb{R} \to \mathbb{R}^3$ such that in coordinates

$$\gamma(s) = (\gamma_1(s), \gamma_2(s), \gamma_3(s)).$$

This curve is called *noncharacteristic* at $\gamma(0)$ if

$$f(\gamma(0))\dot{\gamma}_2(0) - g(\gamma(0))\dot{\gamma}_3(0) \neq 0.$$

Let φ_t denote the flow of the system (3.114), and define $\mathcal{H} : \mathbb{R}^2 \to \mathbb{R}^3$ by

$$(s,t) \mapsto \varphi_t(\gamma(s)). \tag{3.115}$$

Also, define $H : \mathbb{R}^2 \to \mathbb{R}^2$ by projection of the image of $\mathcal{H}$ onto its first two components. More precisely, let e_1, e_2, e_3 be the usual basis vectors for $\mathbb{R}^3$ and let the usual inner product be denoted by angle brackets. Then H is given by

$$(s,t) \mapsto (\langle \varphi_t(\gamma(s)), e_1\rangle, \langle \varphi_t(\gamma(s)), e_2\rangle).$$

We will show that H is locally invertible at $\gamma(0)$. For this, compute

$$\begin{aligned} DH(0,0)e_1 &= \frac{d}{d\tau} H(\tau, 0)\Big|_{\tau=0} \\ &= \frac{d}{d\tau}(\gamma_1(s), \gamma_2(s))\Big|_{\tau=0} \\ &= (\dot{\gamma}_1(0), \dot{\gamma}_2(0)), \end{aligned}$$

and similarly

$$DH(0,0)e_2 = (f(\gamma_1(0)), g(\gamma_2(0))).$$

Because the curve γ is noncharacteristic at $\gamma(0)$, the matrix representation of $DH(0,0)$ has nonzero determinant and is therefore invertible. By the inverse function theorem, H is locally invertible at the origin.

Using the local inverse of H, let us note that

$$\mathcal{H}(H^{-1}(x,y)) = (x, y, \mathcal{H}_3(H^{-1}(x,y))).$$

In other words, if $u(x,y) := \mathcal{H}_3(H^{-1}(x,y))$, then the surface given by the range of $\mathcal{H}$ is locally the graph of the function u. This completes the construction of u; it is a local solution of the PDE (3.113).

We have now proved that if we are given initial data on a noncharacteristic curve, then there is a corresponding local solution of the PDE (3.113). Also, we have a method to construct such a solution.

As an example, let us consider the model equation

$$u_\tau + a\sin(\omega\tau)u_x = u - u^2, \qquad 0 \le x \le 1, \quad t \ge 0$$

with initial data $u(x,0) = u_0(x)$ defined on the unit interval. A phenomenological interpretation of this equation is that u is the density of a species with logistic growth in a moving medium that is changing direction with frequency ω and amplitude a. We have used τ to denote the time parameter so that we can write the first order system for the characteristics in the form

$$\dot\tau = 1, \qquad \dot x = a\sin(\omega\tau), \qquad \dot z = z - z^2.$$

To specify the initial data, let us define the noncharacteristic curve given by $s \mapsto (0, s, u_0(s))$. Then, after solving the first order system and using the definition (3.115), we have that

$$\mathcal{H}(s,t) = \Big(t, s + \frac{a}{\mu}(1-\cos\mu t), \frac{e^t u_0(s)}{1 + u_0(s)(e^t - 1)}\Big).$$

Also, because H^{-1} is given explicitly by

$$H^{-1}(\tau, x) = (\tau, x - \frac{a}{\mu}(1 - \cos\mu\tau)),$$

we have the solution

$$u(x,t) = \frac{e^\tau u_0(x - \frac{a}{\mu}(1-\cos\mu\tau))}{1 + (e^\tau - 1)u_0(x - \frac{a}{\mu}(1 - \cos\mu\tau))}. \tag{3.116}$$

What does our model predict? For example, if the initial condition is given by a positive function u_0, then the ω-limit set of the corresponding solution of the PDE is the constant function $u \equiv 1$, the solution corresponding to no drift. However, if the initial population is distributed so that some regions have zero density, then the fate of the initial population is more complicated (see Exercise 3.46).

Exercise 3.46. What long term behavior for the corresponding model equation is predicted by the solution (3.116)? How does your answer depend on the choice of u_0, a, and ω?

Exercise 3.47. Solve the PDE $xu_x + yu_y = 2u$ with u prescribed on the unit circle. Hint: Define the noncharacteristic curve

$$s \mapsto (\cos s, \sin s, h(\cos s, \sin s)).$$

Exercise 3.48. Find solutions of the PDE $xu_x - yu_y = 2u$. How should the data be prescribed?

Exercise 3.49. A function U that is constant along the orbits of an ordinary differential equation is called an *invariant function*, or a *first integral*. In symbols, if we have a differential equation $\dot x = f(x)$ with flow φ_t, then U is invariant provided that $U(\phi_t(x)) = U(x)$ for all x and t for which the flow is defined. Show that U is invariant if and only if $\langle \operatorname{grad} U(x), f(x)\rangle \equiv 0$. Equivalently, the Lie derivative of U in the direction of the vector field given by f vanishes. Consider the differential equation

$$\dot\theta = 1, \qquad \dot\phi = \alpha$$

where $\alpha \in \mathbb{R}$. Also, consider both θ and ϕ as angular variables so that the differential equation can be viewed as an equation on the torus. Give necessary and sufficient conditions on α so that there is a smooth invariant function *defined on the torus.*

Exercise 3.50. A simple example of a conservation law is the (nonviscous) Burgers' equation $u_t + uu_x = 0$. Burgers' equation with viscosity is given by

$$u_t + uu_x = \frac{1}{Re}u_{xx}$$

where Re is called the *Reynold's number*. This is a simple model that incorporates two of the main features in fluid dynamics: convection and diffusion. Solve the nonviscous Burgers' equation with initial data $u(x,0) = (1-x)/2$ for $-1 < x < 1$. Note that the solution cannot be extended for all time. This is a general phenomenon that appears in the study of conservation laws that is related to the existence of *shock waves* (see [162]). Also, consider the viscous Burgers' equation on the same interval with the same initial data and with boundary conditions

$$u(-1,t) = 1, \qquad u(1,t) = 0.$$

How can we find Galërkin approximations? The problem is that with the nonhomogeneous boundary conditions, there is no vector space of functions that satisfy the boundary conditions. To overcome this problem, we can look for a solution of our problem in the form

$$u(x,t) = v(x,t) + \frac{1}{2}(1-x)$$

where v satisfies the equation

$$v_t + (v + \frac{1}{2}(1-x))(v_x - \frac{1}{2}) = v_{xx}$$

and Dirichlet boundary conditions. Determine the Galërkin approximations using trigonometric trial functions. Use a numerical method to solve the resulting differential equations, and thus approximate the solution of the PDE. For a numerical analyst's approach to this problem, consider the Galërkin approximations with respect to the "test function basis" of Chebyshev polynomials given by

$$T_0(x) = 1, \qquad T_1(x) = x, \qquad T_2(x) = 2x^2 - 1$$

and

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

The Chebyshev polynomials are orthogonal (but not orthonormal) with respect to the inner product defined by

$$\langle f, g\rangle := \int_{-1}^{1} f(x)g(x)(1-x^2)^{-1/2}\,dx.$$

Moreover, the Chebyshev polynomials do not satisfy the boundary conditions. However, proceed as follows: Look for a Galërkin approximation in the form

$$u_n(x,t) = \sum_{i=1}^{n} c_i(t)T_{n-1}(x),$$

but only construct the corresponding system of differential equations for

$$c_1, \ldots, c_{n-2}.$$

Then, define the last two coefficients so that the boundary conditions are satisfied (see [69]). Compare numerical results. Finally, note that Burgers' equation can, in principle, be solved explicitly by the Hopf–Cole transformation. In fact, if u is a solution of Burgers' equation and ψ is defined so that $\psi_x = u$, then ψ is defined up to a function that depends only on the time variable. An appropriate choice of the antiderivative satisfies the equation

$$\psi_t + \frac{1}{2}\psi_x^2 = \frac{1}{Re}\psi_{xx}.$$

If ϕ is defined by the equation $\psi = -(2/Re)\phi$, then

$$\phi_t = \frac{1}{Re}\phi_{xx}.$$

Thus, solutions of the heat equation can be used to construct solutions of Burgers' equation. The fact that Burgers' equation can be solved explicitly makes this PDE a very useful candidate for testing numerical codes.

4
Hyperbolic Theory

The chapter is an introduction to the theory of hyperbolic structures in differential equations. The basic idea might be called "the principle of hyperbolic linearization." Namely, if the linearized flow of a differential equation has "no eigenvalues with zero real parts," then the nonlinear flow behaves locally like the linear flow. This idea has far-reaching consequences that are the subject of many important and useful mathematical results. Here we will discuss two fundamental theorems: the center and stable manifold theorem for a rest point and Hartman's theorem.

4.1 Invariant Manifolds

One of the important results in the theory of ordinary differential equations is the stable manifold theorem. This and many closely related results, for example, the center manifold theorem, form the foundation for analyzing the dynamical behavior of a dynamical system in the vicinity of an invariant set. In this section we will consider some of the theory that is used to prove such results, and we will prove the existence of invariant manifolds related to the simplest example of an invariant set, namely, a rest point. However, the ideas that we will discuss can be used to prove much more general theorems. In fact, some of the same ideas can be used to prove the existence and properties of invariant manifolds for infinite dimensional dynamical systems.