Nonlinear Evolution Equations[1]

John K. Hunter

October 1996

Contents

Chapter 1

Introduction

1.1 Evolution equations

The evolution of a system depending on a continuous time variable t is described by an equation of the form

$$\dot{u} = f(u). \tag{1.1}$$

Here, the dot denotes a time derivative, $u(t) \in X$ is the state of the system at time t, and f is a given vector field on X. The space X is the state space of the system; a point in X specifies the instantaneous state of the system. We will assume that X is a Banach space. When X is finite dimensional, the evolution equation is a system of ordinary differential equations (ODE's). Partial differential equations (PDE's) can be regarded as evolution equations on an infinite dimensional state space. The solution $u(x,t)$ belongs to a function space in x at each instant of time t. Abusing notation slightly, we write $u(t) = u(\cdot, t)$.

Other types of evolution equations can be written in the form (1.1). For example, the nonautonomous equation

$$\dot{u} = f(t, u)$$

can be written in an autonomous form by the introduction of t as a new dependent variable,

$$\begin{aligned} \dot{u} &= f(t, u), \\ \dot{t} &= 1. \end{aligned}$$

Higher order equations can be written as first order equations by the introduction of time derivatives as new dependent variables. For example, the second order equation

$$\ddot{u} = f(u)$$

can be written as the first order system,

$$\begin{aligned} \dot{u} &= v, \\ \dot{v} &= f(u). \end{aligned}$$

Systems with memory in which f is a functional of the history of u (as, for instance, occurs in equations modelling the flow of visco-elastic fluids),

$$\dot{u}(t) = f[u(s) : s \leq t],$$

can be rewritten abstractly in the form (1.1) by the introduction of the history up to time t as a new dependent variable,

$$U(t) = \{u(s) : s \leq t\}.$$

Of course, these artifacts do not make the problem any easier. Although the form of the equation is simplified, the size of the state space is increased, and the special structure of the evolution equation is obscured. Nevertheless, when considering evolution equations in an abstract sense, we may restrict our attention to equations of the form (1.1).

There are many questions which can be asked about evolution equations: the existence of particular types of solutions, such as equilibrium solutions, travelling waves, self-similar solutions, time-periodic solutions; the dynamic stability of these solutions; the long time asymptotic behavior of solutions; chaotic dynamics; complete integrability; singular perturbation expressions for solutions; evolution of random solutions; convergence of numerical schemes; and so on.

The most basic question concerns the existence and uniqueness of solutions. The initial value problem for (1.1) is

$$\begin{aligned} \dot{u} &= f(u), \\ u(0) &= u_0. \end{aligned}$$

If the evolution equation is to provide a self-consistent mathematical model of some real system then, at a minimum, there should exist a unique solution of this initial value problem. These notes focus on this basic question.

1.2 Blow up

A central issue in the study of nonlinear evolution equations is that solutions may exist locally in time (that is, for short times) but not globally in time. This is caused by a phenomenon called "blow-up". Let us illustrate this phenomenon with three very simple ODE's for $u(t) \in \mathbf{R}$:

$$\text{(a)} \quad \dot{u} = u; \qquad \text{(b)} \quad \dot{u} = u^2; \qquad \text{(c)} \quad \dot{u} = -u + u^2.$$

The solution of (a) is

$$u(t) = ce^t.$$

This solution is defined globally in time and grows exponentially as $t \to +\infty$. Global existence and, at most, exponential growth are typical features of well-posed linear evolution equations. Equation (b) is called a Ricatti equation. The solution is

$$u(t) = \frac{1}{c - t}.$$

We have $u(t) \to \infty$ as $t \to c$. As this example shows, nonlinearities which grow super-linearly in u can lead to blow-up and a loss of global existence. The solution of (c) is

$$u(t) = \frac{1}{1 - ce^t}.$$

If $c \leq 0$, which corresponds to $0 < u(0) < 1$, then the solution exists globally in time. If $0 < c < 1$, which corresponds to $u(0) > 1$, then the solution blows up at $t = \log(1/c)$. We thus have global existence of solutions with small initial data and local existence of solutions with large initial data. This type of behavior also occurs in many PDE's; for small initial data, linear damping terms can dominate the nonlinear terms, and one obtains global solutions. For large inital data, the nonlinear blow-up overwhelms the linear damping, and one only has local solutions.

For ODE's with a smooth vector field, the only way in which solutions can fail to exist is by becoming unbounded. For PDE's, the solution is redistributed in space, and there are many more ways in which blow-up can occur. The solution itself may become unbounded. This is "blow-up" in the narrow sense. Alternatively, spatial derivatives of the solution may become unbounded, due the formation of some kind of singularity in the solution. In some cases, it is possible to continue a smooth solution past the blow-up time by a weak solution; in other cases, blow-up signals a catastrophic breakdown in the ability of the PDE to model the original system.

One of the simplest nonlinear PDE's which exhibits blow-up is the **inviscid Burgers equation**,

$$u_t + \left(\frac{1}{2}u^2\right)_x = 0, \tag{1.2}$$
$$u(x,0) = u_0(x).$$

Here, $x \in \mathbf{R}$ and $u(x,t) \in \mathbf{R}$. We will show that (1.2) cannot have a global smooth solution if $u_0'(x) < 0$ at any point. The proof is by contradiction. Suppose that $u(x,t)$ is a smooth solution. Differentiation of (1.2) with respect to x implies that

$$v_t + uv_x + v^2 = 0,$$

where $v = u_x$. This equation can be written as

$$\dot{v} = -v^2,$$

where

$$\dot{} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial x}$$

is a derivative along the characteristic curves associated with u. If $v < 0$ at $t = 0$, then the solution of this Ricatti equation blows up at some positive time, and a global smooth solution cannot exist.

The blow-up time along a characteristic is $t = -1/v(0)$. Thus if u_0 is not a monotonic increasing function, a smooth solution of the PDE cannot exist past the time

$$t^* = -\frac{1}{\inf_{x\in\mathbf{R}} u_0'(x)}.$$

One has to be a little careful in interpreting the conclusion of this type of argument. The proof shows that global smooth solutions of the inviscid Burgers equation cannot exist. As it stands, it does not prove that u_x blows up or that the life-span of smooth solutions is actually t^*, although this is in fact true. The reason that this conclusion does not follow from the the argument above is that some other quantity (for instance, u itself) could conceivably blow up first. In the case of the inviscid Burgers equation, the method of characteristics shows that for smooth initial data, there is a smooth solution in the the interval $0 \le t < t^*$.

There are a number of proofs of the lack of global existence of smooth solutions for nonlinear evolution equations which show that a certain quantity would have to blow-up for smooth solutions even though the quantity in question never actually blows up.

It is possible to obtain global weak solutions of the inviscid Burgers equation which are defined for all $0 \leq t < \infty$. These solutions are of bounded variation but they are not continuously differentiable since they typically contain jump discontinuities, or shocks.

Existence results can thus be classified in various different ways:

- local existence vs. global existence;
- smooth (or classical) solutions vs. weak solutions;
- small initial data vs. large initial data.

Nonlinear equations are usually difficult to analyze. Local existence can be established by standard arguments for most reasonable PDE's. Global existence is often much harder to establish and typically requires sufficiently strong *a priori* estimates for solutions of the PDE. The basic global existence and uniqueness questions are not resolved even for some of the most important nonlinear evolution equations arising in classical physics. We will briefly discuss what is known for the Navier-Stokes and Euler equations in fluid mechanics and for the Einstein field equations in general relativity.

1.3 Fluid mechanics

The motion of a viscous incompressible fluid with velocity vector $\mathbf{u}(\mathbf{x}, t) \in \mathbf{R}^3$ and pressure $p(\mathbf{x}, t) \in \mathbf{R}$ is modelled by the **incompressible Navier-Stokes** equations,

$$\begin{aligned} &\mathbf{u}_t + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p = \nu \Delta \mathbf{u}, \\ &\nabla \cdot \mathbf{u} = 0. \end{aligned} \tag{1.3}$$

Here $\nu > 0$ is a viscosity coefficient. The component form of (1.3) is

$$\begin{aligned} &\frac{\partial u^i}{\partial t} + u^j \frac{\partial u^i}{\partial x^j} + \frac{\partial p}{\partial x^i} = \nu \frac{\partial^2 u^i}{\partial x^j \partial x^j}, \\ &\frac{\partial u^j}{\partial x^j} = 0, \end{aligned}$$

where $\mathbf{u} = (u^1, u^2, u^3)$, $\mathbf{x} = (x^1, x^2, x^3)$, and we sum over repeated indices. Leray (1934) and Hopf (1951) proved that there are global weak solutions of the incompressible Navier-Stokes equations. Leray's paper is one of the foundational papers in the modern theory of PDE's. It is not known, however, if the Leray-Hopf weak solutions are smooth. Nor is it known if weak solutions are unique. Thus, despite the complete physical success of the Navier-Stokes equations in modelling fluid flows, there remains a fundamental gap in the mathematical theory of these equations. The possible lack of smoothness of solutions of (1.3) may appear surprising from a physical point of view since the viscous term $\nu\Delta\mathbf{u}$ is a smoothing term. However, it is not clear whether or not this linear viscous term can always overcome the quadratically nonlinear term $\mathbf{u}\cdot\nabla\mathbf{u}$ which conceivably could cause derivatives of $\mathbf{u}$ to blow-up.

For two-dimensional flows, with $\mathbf{x}\in\mathbf{R}^2$, the situation is much simpler and one has global existence and uniqueness of smooth solutions.

When $\nu = 0$, which corresponds to an inviscid fluid, the incompressible Navier-Stokes equations reduce to the **incompressible Euler equations**,

$$\begin{aligned}&\mathbf{u}_t + \mathbf{u}\cdot\nabla\mathbf{u} + \nabla p = 0,\\&\nabla\cdot\mathbf{u} = 0.\end{aligned} \tag{1.4}$$

These equations are not as well founded from a physical point of view as the Navier-Stokes equations. The limit $\nu\to 0$ is a singular limit, and it is not clear that the Euler equations always provide a correct description of fluid flows in this limit. For example, in turbulent flows the rate at which energy is dissipated seems to tend to a non-zero value as $\nu\to 0$. This non-zero limit is inconsistent with the conservative nature of the Euler equations.

Smooth solutions of the Euler equations exist for short times. For three-dimensional flows, it is not known whether or not smooth solutions blow-up in finite time. Nor is it known whether or not global weak or smooth solutions exist. The crucial quantity which controls the blow-up is the vorticity,

$$\omega = \operatorname{curl}\mathbf{u}.$$

Beale, Kato, and Majda (1984) proved that if blow-up occurs at time $t = t^*$, then the supremum of the vorticity must have a nonintegrable singularity as a function of time, meaning that

$$\int_0^t \|\omega\|_\infty(s)\,ds \uparrow \infty \qquad \text{as } t\uparrow t^*.$$

The existence theory for the equations which describe compressible fluid flows is in even worse shape. Inviscid compressible flows of an inviscid fluid are described by the **compressible Euler equations**,

$$\begin{aligned}
&\rho_t + \nabla \cdot (\rho \mathbf{u}) = 0, \\
&(\rho \mathbf{u})_t + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u} + pI) = 0, \\
&\left[\rho \left(e + \frac{1}{2}\mathbf{u} \cdot \mathbf{u}\right)\right]_t + \nabla \cdot \left[\rho \left(e + \frac{1}{2}\mathbf{u} \cdot \mathbf{u}\right) \mathbf{u} + p\mathbf{u}\right] = 0.
\end{aligned} \tag{1.5}$$

For an ideal gas with constant ratio of specific heats γ, the internal energy e is given in terms of the density ρ and the pressure p by

$$e(p, \rho) = \frac{1}{\gamma - 1} \frac{p}{\rho}.$$

Thus (1.5) is a system of equations for $(\rho, \mathbf{u}, p)$.

The compressible Euler equations are the fundamental physical example of a hyperbolic systems of conservation laws. The general form of a hyperbolic system of conservation laws is

$$\mathbf{u}_t + \nabla \cdot \mathbf{F}(\mathbf{u}) = 0 \tag{1.6}$$

where $\mathbf{u} = (u^1, \dots, u^m)$ is a vector of conserved quantities and $\mathbf{F} = (f^{ij})$ is a flux tensor. In component notation, with space variables $\mathbf{x} = (x^1, \dots, x^d)$, equation (1.6) is

$$u^i_t + f^{ij}_{x^j} = 0.$$

The simplest example of a hyperbolic conservation law is the inviscid Burgers equation (1.2). The breakdown of smooth solutions and the subsequent formation of shocks introduce special difficulties in the mathematical theory of such equations.

In the case of one space dimension, the conservation law (1.6) reduces to the equation

$$\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = 0, \tag{1.7}$$

where $\mathbf{f} : \mathbf{R}^m \to \mathbf{R}^m$. In a fundamental paper, Glimm (1965) proved that, under suitable hypotheses, the initial value problem for equation (1.7) has a global weak solution for small initial data. The uniqueness of the solutions constructed by Glimm's method has only recently been shown by Bressan. Global existence of weak solutions of (1.5) or (1.7) with large initial data in one space dimension remains open.

Glimm's method of proof is essentially restricted to the case of one space dimension. Local existence of smooth solutions in more than one space dimension follows from a standard theory of symmetric hyperbolic PDE's. However, smooth solutions typically break down in finite time. The are no global existence results for (1.5) or (1.6) in two or three space dimensions.

The unsatisfactory state of the existence and uniqueness theory for these fluid equations is remarkable, particularly in light of their physical importance and the enormous effort that has been devoted to their study.

1.4 General relativity

One of the most fundamental nonlinear field theories is the theory of general relativity proposed by Einstein in 1916. According to general relativity, space-time is a four dimensional (one time and three space dimensions) Lorentzian manifold $(\mathcal{M}, \mathbf{g})$. The curvature of space-time $\mathcal{M}$ is described by a fourth order tensor called the Riemann tensor, $\mathbf{Riem}(\mathbf{g})$. The Ricci tensor $\mathbf{Ric}(\mathbf{g})$ is a second order tensor obtained by contraction of the Riemann tensor and the scalar curvature $R(\mathbf{g})$ is obtained by contraction from the Ricci tensor. The Einstein tensor $\mathbf{G}(\mathbf{g})$ is defined by

$$\mathbf{G} = \mathbf{Ric} - \frac{1}{2} R\mathbf{g}.$$

The Einstein field equations, in suitable units, are then

$$\mathbf{G} = 8\pi\mathbf{T},$$

where the energy-momentum tensor $\mathbf{T}$ describes the distribution of matter and radiation fields in space-time. If space-time is empty then $\mathbf{T} = 0$ and we obtain the Einstein vacuum equations,

$$\mathbf{G} = 0,$$

which are equivalent to

$$\mathbf{Ric} = 0. \tag{1.8}$$

In order to carry out a concrete analysis of (1.8), one has to introduce a local coordinate system $\{x^\alpha : \alpha = 0, \ldots, 3\}$ in $\mathcal{M}$. The metric tensor and the Ricci tensor are then given by

$$\begin{aligned} \mathbf{g} &= g_{\alpha\beta}(x) dx^\alpha \otimes dx^\beta, \\ \mathbf{Ric} &= R_{\alpha\beta}(x) dx^\alpha \otimes dx^\beta. \end{aligned}$$

Since the metric tensor is symmetric, $\mathbf{g}$ is determined by ten independent components $\{g_{\alpha\beta} : 0 \leq \alpha \leq \beta \leq 3\}$. The components of the Ricci tensor are given in terms of the metric components by

$$R_{\alpha\beta} = \frac{\partial \Gamma^{\lambda}_{\alpha\beta}}{\partial x^{\lambda}} - \frac{\partial \Gamma^{\lambda}_{\beta\lambda}}{\partial x^{\alpha}} + \Gamma^{\lambda}_{\alpha\beta}\Gamma^{\mu}_{\lambda\mu} - \Gamma^{\mu}_{\alpha\lambda}\Gamma^{\lambda}_{\beta\mu}$$

where $\Gamma^{\alpha}_{\beta\gamma}$ are the connection coefficients,

$$\Gamma^{\alpha}_{\beta\gamma} = \frac{1}{2} g^{\alpha\lambda} \left[\frac{\partial g_{\beta\lambda}}{\partial x^{\gamma}} + \frac{\partial g_{\gamma\lambda}}{\partial x^{\beta}} - \frac{\partial g_{\beta\gamma}}{\partial x^{\lambda}} \right].$$

The component form of (1.8) is thus

$$R_{\alpha\beta} = 0,$$

which is a system of ten second order equations for the components of $\mathbf{g}$.

If we choose our coordinates so that the initial surface is given locally by the equation $x^0 = 0$, where x^0 is a time-like variable, then appropriate initial data for the Einstein vacuum equations is

$$g_{\alpha\beta} = \bar{g}_{\alpha\beta}, \qquad \frac{\partial g_{\alpha\beta}}{\partial x^0} = \bar{h}_{\alpha\beta} \qquad \text{on } x^0 = 0. \tag{1.9}$$

The resulting initial value problem is very degenerate. Any initial surface is characteristic. As a result, the Einstein field equations imply that the initial data must satisfy certain constraints. Moreover, if these constraints are satisfied, the problem does not have a unique solution. These features are related to the gauge invariance of the Einstein field equations under arbitrary changes of the coordinate system in $\mathcal{M}$.

One way to obtain a more standard set of PDE's is to write the field equations in a special class of coordinate systems. Wave (or harmonic) coordinates are characterized by the requirement that

$$g^{\alpha\beta}\Gamma^{\mu}_{\alpha\beta} = 0. \tag{1.10}$$

In wave coordinates, the Einstein-vacuum equations become

$$g^{\alpha\beta} \frac{\partial^2 g_{\mu\nu}}{\partial x^{\alpha} \partial x^{\beta}} + Q_{\mu\nu}(\mathbf{g}) \cdot (\partial \mathbf{g}, \partial \mathbf{g}) = 0. \tag{1.11}$$

The lower order term is a quadratic function of the first derivatives of $\mathbf{g}$ with coefficients depending on $\mathbf{g}$. Explicitly,

$$Q_{\mu\nu} = \frac{\partial g^{\alpha\beta}}{\partial x^{\mu}} \frac{\partial g_{\beta\nu}}{\partial x^{\alpha}} + \frac{\partial g^{\alpha\beta}}{\partial x^{\nu}} \frac{\partial g_{\beta\mu}}{\partial x^{\alpha}} + 2\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}.$$

The form of (1.11) is similar to a standard wave equation since $g^{\alpha\beta}\partial_\alpha\partial_\beta$ is a D'Alembertian operator associated with $g^{\alpha\beta}$. The problem is then to solve (1.11) subject to initial data (1.9) which satisfies the constraint (1.10). The constraints (1.10) are automatically preserved by the evolution; if the initial data satisfies these constraints, then a smooth solution also satisfies the constraints at later times. Choquet-Bruhat (1952) used this formulation to prove local existence of smooth solutions of the Einstein vacuum equations.

Singularity theorems of Penrose show that, at least for certain classes of initial data, smooth solutions of the Einstein vaccum equations must blow-up due to the formation of singularities inside black holes. The definition of a singularity in general relativity is rather subtle, since some singularities may be coordinate singularities caused by the use of an inappropriate local coordinate system, rather than by an intrinsic singularity in space-time itself. For instance, the solution of the field equations written in wave coordinates typically develops coordinate singularities at later times. So wave coordinates are not well-suited to the study of global solutions. It seems likely that at least some singular solutions of the Einstein field equations cannot be continued by weak solutions. This fact suggests that general relativity does not provide a self-consistent description of gravity. Presumably, an as yet unknown theory of quantum gravity is required.

Recently, Christodoulou and Klainerman (1993) have proved that there exist global smooth solutions of the Einstein vaccuum equations with initial data which corresponds to a small, localized perturbation of flat Minkowski space-time. This result does not rule out singularity formation for large initial data.

1.5 Model equations

The examples described in the previous section are very complicated systems. Many ideas in the theory of nonlinear evolution equations are easier to understand in the context of simpler equations which model various aspects of these more complicated systems.

We will briefly mention a few examples of model nonlinear evolution equations. These equations are often physically important in their own right.

Examples.

Diffusion equations. There are many nonlinear generalizations of the linear diffusion equation $u_t = \Delta u$. Some example include

$$\begin{aligned}
&u_t = \Delta u + f(u), &&\text{reaction-diffusion equation}\\
&u_t + f(u)_x = \Delta u, &&\text{advection-diffusion equation}\\
&u_t = \nabla \cdot \left[D(u)\nabla u\right], &&\text{nonlinear diffusion matrix.}
\end{aligned}$$

Reaction-diffusion equations arise in modelling chemical reactions and in poulation biology. Advection-diffusion equations arise in modelling the transport of contaminants and are closely related to the equations which describe viscous fluid flows. Nonlinear diffusion matrices arise in many contexts, such as flow through porous media.

Wave equations. One of the simplest class of nonlinear hyperbolic equations is the nonlinear wave equation

$$u_{tt} - \Delta u + f(u) = 0.$$

This equation provides a simple model for nonlinear classical field theories.

Another important class of hyperbolic partial differential equations is the class of symmetric hyperbolic systems

$$\mathbf{u}_t + A^i(\mathbf{u})\mathbf{u}_{x^i} = 0,$$

where the A^i are symmetric matrices.

Schrödinger equations. The nonlinear Schrödinger equation is

$$iu_t = -\Delta u + f(|u|^2)u.$$

Here $u(x,t)$ is a complex valued function. The name comes from the fact that the equation has the same form as the linear Schrödinger equation in quantum mechanics,

$$iu_t = -\Delta u + V(x,t)u,$$

but with a potential function $V = f(|u|^2)$ which depends on the solution itself.

An important special case is the cubic nonlinear Schrödinger equation (cNLS)

$$iu_t = -\Delta u + \sigma|u|^2 u, \tag{1.12}$$

where $\sigma = \pm 1$. The choice of sign is important; for $\sigma = -1$ equation (1.12) is called the focussing cNLS and for $\sigma = +1$, it is called the defocussing cNLS. As we will see below, the blow-up behavior in these two cases is very different.

Equation (1.12) arises from very general nonlinear dispersive wave systems in a suitable asymptotic limit. For instance, the cNLS equation describes the propagation of a laser beam through a nonlinear optical medium such as a fibre optics cable.

In one space dimension (1.12) is a completely integrable PDE with soliton solutions and an associated linear scattering problem. The equation is not completely integrable in $d \geq 2$ space dimensions.

Chapter 2

The Contraction mapping Theorem and ODE's

2.1 The contraction mapping theorem

The contraction mapping theorem is a simple and widely applicable method of proving existence and uniqueness of solutions of nonlinear differential equations. The method is constructive, since solutions are obtained as a limit of a sequence of approximations. Many iterative methods for solving nonlinear equations can be formulated in terms of contraction mappings.

Theorem 2.1 *Suppose that $F \subset X$ is a closed subset of a Banach space X and*

$$T : F \to F$$

is a mapping on F such that

$$\|Tu - Tv\| \leq \alpha \|u - v\|$$

for all $u, v \in F$ and some constant $\alpha < 1$. Then T has a unique fixed point $\bar{u} \in F$ which satisfies

$$T\bar{u} = \bar{u}.$$

Proof. For any $u_0 \in X$, the iterates

$$u_n = T^n u_0$$

form a Cauchy sequence since

$$\|u_n - u_m\| \leq \sum_{k=m}^{n-1} \|u_{k+1} - u_k\| \leq \left(\sum_{k=m}^{n-1} \alpha^k \right) \|u_1 - u_0\|.$$

The limit of this sequence is the unique fixed point of T. QED

2.2 Lipschitz conditions

To use the contraction mapping theorem to prove existence and uniqueness of solutions of ODE's, we require a Lipschitz condition on the vector field.

Definition 2.1 *Let E be a subset of a Banach space X. A function $f : X \to X$ is said to be* **Lipschitz on** *E if there is a constant K such that*

$$\|f(u) - f(v)\| \leq K\|u - v\| \qquad \text{for all } u, v \in E.$$

If this condition holds for all $u, v \in X$, then f is said to be **globally Lipschitz**. *The constant K is called a* **Lipschitz constant** *for f*

Lipschitz functions are continuous but their graphs can have "corners," so they need not be continuously differentiable. Globally Lipschitz functions grow at most linearly in u as $u \to \infty$. Here a some simple examples of functions $f : \mathbf{R} \to \mathbf{R}$, which illustrate the definition.

Examples.

1. The function $|u|^{1/2}$ is continuous, but it is not Lipschitz on any neighborhood of $u = 0$.

2. The function $|u|$ is globally Lipschitz, but it is not differentiable at $u = 0$.

3. The functions u^2 and e^u are Lipschitz on any bounded set, but they are not globally Lipschitz, since they grow faster than $|u|$ as $u \to \infty$.

4. The function
$$f(u) = \frac{u^2}{(1 + u^2)^{1/2}}$$
is globally Lipschitz on $\mathbf{R}$.

A few immediate consequences of the definition are as follows.

Examples.

1. Any continuous affine function $f : X \to X$,

$$f(u) = Au + b,$$

where A is a bounded linear operator on X and $b \in X$, is globally Lipschitz with Lipschitz constant $K = \|A\|$.

2. If $f : X \to X$ is globally Lipschitz, there exists a constant C such that

$$\|f(u)\| \leq C\,(1 + \|u\|)$$

for all $u \in X$. If f is Lipschitz on a bounded set E, then f is bounded on E. To prove these statements, we pick a fixed u_0 in E (or X) and estimate

$$\|f(u)\| \leq \|f(u) - f(u_0)\| + \|f(u_0)\| \leq K\|u - u_0\| + \|f(u_0)\|.$$

3. Suppose that $f : X \to X$ is continuously differentiable in the ball

$$B_R(u_0) = \{u \in X : \|u - u_0\| < R\}$$

and f' is uniformly bounded, meaning that

$$\sup_{u \in B_R(u_0)} \|f'(u)\| = K < +\infty.$$

Then f is Lipschitz on $B_R(u_0)$ with Lipschitz constant K. In particular, if f' is uniformly bounded on X, then f is globally Lipschitz.

Proof. Using the fundamental theorem of calculus, we get that for any $u, v \in B_r(u_0)$

$$\begin{aligned} \|f(u) - f(v)\| &= \left\| \int_0^1 \frac{d}{dt} f\left(tu + (1-t)v\right)\, dt \right\| \\ &\leq \int_0^1 \left\| f'\left(tu + (1-t)v\right) \cdot (u - v) \right\| \, dt \\ &\leq K\|u - v\|. \qquad \text{QED} \end{aligned}$$

2.3 The Picard existence theorem

Two simple but fundamental examples illustrate the basic features of existence and uniqueness for the initial value problem for ODE's,

$$\dot{u} = f(u),$$
$$u(0) = u_0$$

The main points are:

- in order to guarantee uniqueness, f must be Lipschitz;
- if f is Lipschitz on bounded sets, the solution may blow up in finite time, and then we have only local existence;
- if f is globally Lipschitz, then the solution exists globally in time.

Examples.

1. Consider the initial value problem for $u : \mathbf{R} \to \mathbf{R}$,

$$\dot{u} = |u|^{1/2},$$
$$u(0) = 0.$$

Note that the vector field $f(u) = |u|^{1/2}$ is not Lipschitz in any neighborhood of the initial data $u = 0$. This initial value problem has many solutions. For example, two solutions are

$$u(t) = 0,$$
$$u(t) = \begin{cases} 0 & \text{if } t \leq 0 \\ t^2/4 & \text{if } t > 0 \end{cases}$$

2. Consider the initial value problem for $u : \mathbf{R} \to \mathbf{R}$,

$$\dot{u} = u^2,$$
$$u(0) = u_0.$$

Note that the vector field $f(u) = u^2$ is Lipschitz on any bounded set (it's continuously differentiable), but it is not globally Lipschitz. The solution is $u(t) = 0$ if $u_0 = 0$, and

$$u(t) = \frac{1}{1/u_0 - t}$$

otherwise. The solution is not defined for all t since $|u(t)| \to \infty$ as $t \to 1/u_0$. This example is the simplest example of blow up.

In most applications, the vector field f is smooth so that a local Lipschitz condition is satisfied. A global Lipschitz condition requires that $f(u)$ grows linearly in u as $\|u\| \to \infty$ and this condition is too restrictive for most applications.

Theorem 2.2 *(Picard existence theorem.) Let X be a Banach space. Suppose that $f : X \to X$ is Lipschitz on a closed ball $\bar{B}_R(u_0) \subset X$, where $R > 0$ and $u_0 \in X$. Let*

$$M = \sup_{u \in \bar{B}_R(u_0)} \|f(u)\| < \infty.$$

The initial value problem

$$\begin{aligned} \dot{u} &= f(u), \\ u(0) &= u_0 \end{aligned}$$

has a unique continuously differentiable local solution $u(t)$. This solution is defined in the time interval $-\delta < t < \delta$, where

$$\delta = \frac{R}{M}.$$

Proof. We rewrite the initial value problem as an integral equation

$$u = Tu,$$

where

$$Tu(t) = u_0 + \int_0^t f(u(s))\, ds.$$

For $0 < \eta < R/M$ we define

$$Y = C\left([-\eta, \eta]; \bar{B}_R(u_0)\right).$$

We will show that $T : Y \to Y$ is a contraction when η is sufficiently small.

First, note that if $u \in Y$ then

$$\|Tu(t) - u_0\| = \left\| \int_0^t f(u(s))\, ds \right\| \le M\eta < R.$$

Hence $Tu \in Y$ so that $T : Y \to Y$.

Second, we estimate

$$\begin{aligned}\|Tu - Tv\|_Y &= \sup_{|t|\le\eta}\left\|\int_0^t [f(u(s)) - f(v(s))]\, ds\right\| \\ &\le K\eta\|u - v\|_Y,\end{aligned}$$

where K is a Lipschitz constant for f on $\bar{B}_R(u_0)$. Hence if we choose $\eta = K/2$ then T is a contraction on Y and it has a unique fixed point.

Since η depends only on the Lipschitz constant of f and on the distance R of the initial data from the boundary of $\bar{B}_R(u_0)$, repeated application of this result gives a unique local solution defined for $|t| < R/M$. QED

An important feature of this result is that the existence time only depends on the norm of u. Thus, the only way in which the solution of an ODE can fail to exist (assuming that the vector field f is Lipschitz continuous on any ball) is if $\|u(t)\|$ becomes unbounded.

Suppose we can prove an priori estimate of the form

$$\|u(t)\| \le C \qquad \text{for all } t \ge 0.$$

Then the local existence theorem implies that for any initial data at $t = t_0$ with $\|u(t_0)\| \le C$ we have a local solution defined in an interval $|t - t_0| < \delta$ where $\delta = C/M$ depends only on C and the supremum M of $f(u)$ over $\{u : \|u\| \le C\}$. Thus the local solution can be extended outside any finite interval to give a global solution $u : [0, \infty) \to X$.

One way to obtain an a priori estimate is by means of Liapounov functions. Another way is to use differential inequalities.

2.4 Gronwall's inequality

First we give a linear differential inequality called Gronwall's inequality.

Theorem 2.3 *Suppose that $u(t) \ge 0$ and $\varphi(t) \ge 0$ are continuous nonnegative real-valued functions defined on the interval $0 \le t \le T$ and $u_0 \ge 0$ is a nonnegative constant. If u satisfies the inequality*

$$u(t) \le u_0 + \int_0^t \varphi(s)u(s)\, ds, \qquad t \in [0, T],$$

then

$$u(t) \le u_0 \exp\left(\int_0^t \varphi(s)\, ds\right), \qquad t \in [0, T].$$

In particular, if $u_0 = 0$ then $u(t) = 0$.

Proof. Suppose first that $u_0 > 0$. Let

$$U(t) = u_0 + \int_0^t \varphi(s)u(s)\,ds.$$

Then, since

$$u(t) \le U(t),$$

we have that

$$\begin{aligned} \dot{U} &= \varphi u \le \varphi U, \\ U(0) &= u_0. \end{aligned}$$

Since $U(t) > 0$, it follows that

$$\frac{d}{dt}\log U = \frac{\dot{U}}{U} \le \varphi.$$

Hence

$$\log U(t) \le \log u_0 + \int_0^t \varphi(s)\,ds,$$

so

$$u(t) \le U(t) \le u_0 \exp\left(\int_0^t \varphi(s)\,ds\right).$$

If $u_0 = 0$ then the above estimate holds for all $u_0 > 0$. Taking the limit as $u_0 \to 0$ we conclude that $u(t) = 0$. QED

Here is a nonlinear generalization of Gronwall's inequality.

Theorem 2.4 *Suppose that $f(t,u)$ is a continuous function which is Lipschitz continuous and monotone increasing in u. Suppose that $u_0 \le v_0$ are constants. Let $u(t)$ be a continuous function such that*

$$u(t) \le u_0 + \int_0^t f(s,u(s))\,ds, \qquad t \in [0,T],$$

and let $v(t)$ be the solution of

$$\begin{aligned} \dot{v} &= f(t,v), \\ v(0) &= v_0, \end{aligned}$$

Then

$$u(t) \le v(t) \qquad t \in [0,T].$$

.

Proof. Let
$$U(t) = u_0 + \int_0^t f(s, u(s))\, ds.$$
Then since $u(t) \leq U(t)$ and f is monotone, we have that
$$\dot{U} = f(t, u(t)) \leq f(t, U(t)).$$
Moreover,
$$U(0) = u_0 \leq v_0.$$
The theorem is proved once we show that $U(t) \leq v(t)$ on $[0, T]$.

Given $\varepsilon > 0$, let $v_\varepsilon(t)$ be the solution of
$$\begin{aligned} \dot{v}_\varepsilon &= f(t, v_\varepsilon) + \varepsilon, \\ v_\varepsilon(0) &= v_0. \end{aligned}$$
Continuous dependence results for ODE's imply that $v_\varepsilon \to v$ uniformly on $[0, T]$ as $\varepsilon \to 0$. It is therefore sufficient to prove that
$$U(t) \leq v_\varepsilon(t)$$
for $\varepsilon > 0$. Suppose this inequality is false. Then there exists $t_1 \in [0, T]$ such that $U(t_1) > v_\varepsilon(t_1)$. Let
$$t_2 = \sup_{t \in [0, t_1)} \{t : U(t) = v_\varepsilon(t)\}.$$
Then $U(t) > v_\varepsilon(t)$ for $t \in (t_2, t_1]$. However, $U(t_2) = v_\varepsilon(t_2)$ and $\dot{U}(t_2) < \dot{v}_\varepsilon(t_2)$. Hence we must have $v_\varepsilon(t) > U(t)$ for some $t_2 < t < t_1$. This contradiction proves the result. QED

One immediate consequence of Gronwall's inequality is the global existence of solutions of ODE's with globally Lipschitz vector fields. In particular, linear systems of ODE's have global solution.

Theorem 2.5 *Let $f : X \to X$ be a globally Lipschitz vector field. Then there is a unique global solution $u \in C^1(\mathbf{R}; X)$ of the initial value problem*
$$\begin{aligned} \dot{u} &= f(u), \\ u(0) &= u_0 \in X. \end{aligned}$$

Proof. In view of the local existence theorem, the result follows once we show that any local solution satisfies an priori estimate $\|u(t)\| \leq C(T)$ on any time interval $[0, T]$. Changing t to $-t$, we get global existence backwards in time in an identical fashion. Integration of the ODE implies that

$$u(t) = u_0 + \int_0^t f(u(s))\, ds.$$

Since f is globally Lipschitz, there exists a constant K such that

$$\|f(u)\| \leq K(1 + \|u\|).$$

It follows that

$$\|u(t)\| \leq \|u_0\| + Kt + K \int_0^t \|u(s)\|\, ds.$$

The solution of

$$\begin{aligned} v(t) &= v_0 + Kt + \int_0^t v(s)\, ds, \\ v(0) &= v_0, \end{aligned}$$

is given by

$$v(t) = (v_0 + 1)\, e^{Kt} - 1.$$

Hence, by Gronwall's inequality, we have

$$\|u(t)\| \leq (\|u_0\| + 1)\, e^{Kt} - 1.$$

This estimate proves the theorem. QED

This global existence result can also be proved directly by use of the contraction mapping theorem. If f is globally Lipschitz, the contraction mapping argument gives local existence of solutions in a fixed time interval which depends only on the Lipschitz constant of f. It follows that the local solution can be extended to a global solution.

2.5 Linear evolution equations and semigroups

The local existence theorem for ODE's is not restricted to the case when X is finite dimensional. This fact might suggest that the existence theorem is widely applicable to PDE's as well as ODE's. However, there is a fundamental obstacle in applying the theorem to PDE's: *differential operators are not continuous.* As a result the vector fields associated with PDE's are typically defined only on a dense subset of the state space X. Nevertheless, there is a rather general theory for linear evolution equations. This theory is called semigroup theory since the solution operators form a semigroup. It is possible to define semigroups for some classes of nonlinear equations as well, but there is no general theory of nonlinear semigroups. We begin by summarizing some basic facts about linear operators.

2.5.1 Linear operators

A linear operator $A : X \to X$ is said to be **bounded** if its operator-norm,

$$\|A\| = \sup_{\substack{u \in X \\ u \neq 0}} \frac{\|Au\|}{\|u\|},$$

is finite. A linear operator is continuous iff it is bounded. Moreover, if a linear operator is continuous then it is automatically Lipschitz continuous since

$$\|Au - Av\| = \|A(u - v)\| \leq \|A\|\|u - v\|.$$

Unbounded operators that arise in applications cannot be defined on the whole space X. Instead, their domain is a dense subspace $D(A) \subset X$. For instance, consider the Laplacian operator Δ acting on L^2. We only have $\Delta u \in L^2$ if $u \in H^2$. If we take all of H^2 as the domain, we get the Laplacian operator,

$$A_1 = \Delta : H^2(\Omega) \subset L^2(\Omega) \to L^2(\Omega).$$

The domain of an unbounded operator is a crucial part of its definition. In particular, boundary conditions are built into the domain. For instance, we define the Dirichlet Laplacian operator by

$$A_2 = \Delta : H^2(\Omega) \cap H_0^1(\Omega) \subset L^2(\Omega) \to L^2(\Omega).$$

The operators A_1 and A_2 are different since they have different domains.

Although differential operators are unbounded they do possess the following weaker continuity property.

Definition 2.2 *An operator $A : D(A) \subset X \to X$ is said to be* **closed** *if whenever $u_n \in D(A)$, $u_n \to u$, and $Au_n \to v$ then $u \in D(A)$ and $Au = v$.*

Note that the definition does not assert that $\{Au_n\}$ converges if $\{u_n\}$ converges, as would be the case if A were bounded. All that is asserted is that *if* both $\{u_n\}$ and $\{Au_n\}$ converge then the limits u and v are related by $v = Au$. The fact that differential operators, with properly defined domains, are closed follows from the continuity of distributional derivatives.

As an example, consider the Laplacian operator A_3 defined on C^2-functions,

$$A_3 = \Delta : C^2(\bar{\Omega}) \subset L^2(\Omega) \to L^2(\Omega).$$

This operator is not closed. To see why, suppose that $C^2 \ni u_n \to u \in H^2$ and $C^0 \ni \Delta u_n \to v = \Delta u \in L^2$ where $u \in H^2 \setminus C^2$ and the convergence is in the L^2-sense. Then $u \notin D(A_3)$, so the operator is not closed. We can enlarge the domain of A_3 to obtain the closure $\bar{A}_3 = A_2$. Both A_1 and A_2 are closed.

A detailed understanding of linear operators can be obtained by the use of spectral theory. We will describe a few basic facts which we use below.

Suppose that A is a bounded or unbounded linear operator on the Banach space X,

$$A : D(A) \subset X \to X.$$

We say that $\lambda \in \mathbf{C}$ belongs to the **resolvent set** of A if $\lambda I - A$ has a bounded inverse with domain all of X,

$$R(\lambda) = (\lambda I - A)^{-1} : X \to X.$$

The operator $R(\lambda)$ is called the **resolvent operator** of A. The **spectrum** of A is the complement of the resolvent set in the complex plane. We denote the resolvent set of A by $\rho(A) \subset \mathbf{C}$ and the spectrum of A by $\sigma(A) = \mathbf{C} \setminus \rho(A)$.

If X is finite dimensional then the spectrum of A is a finite set which consists of the eigenvalues of A. If X is infinite dimensional then A may have a continuous spectrum in addition to its eigenvalues. This can occur even if A is bounded. However, the spectrum of a bounded operator A is contained in the closed disc $\{\lambda \in \mathbf{C} : |\lambda| \leq \|A\|\}$.

Finally we define the adjoint of an operator acting in a Hilbert space. We denote the Hilbert space inner product by $(\cdot, \cdot)$.

Definition 2.3 *Suppose that A is a linear operator in a Hilbert space H,*

$$A : D(A) \subset H \to H,$$

with domain $D(A)$ dense in H. We say that $u \in D(A^*)$ if there exists $w \in H$ such that

$$(u, Av) = (w, v) \qquad \textit{for all } v \in D(A).$$

In that case, we set $w = A^*u$. This defines a linear operator

$$A^* : D(A^*) \subset H \to H.$$

We call A^* the **adjoint** *of A. We say that A is* **self-adjoint** *if $D(A) = D(A^*)$ and $A = A^*$.*

The characteristic property of the adjoint is that

$$(u, Av) = (A^*u, v) \qquad \text{for all } u \in D(A^*) \text{ and } v \in D(A).$$

For example, the Laplacian operator A_2 is self-adjoint but neither A_1 nor A_3 are self-adjoint. The problem with A_1 is that the boundary conditions are not self-adjoint so that

$$D(A_1^*) = H_0^2(\Omega) \neq D(A_1).$$

The problem with A_3 is that it is not closed; the operator is self-adjoint after closure.

Self-adjoint operators are very important in applications (such as quantum mechanics) and have many nice properties. For example, the spectrum of a self adjoint operator is real.

2.5.2 Linear evolution equations

Now let us consider the linear evolution equation

$$\begin{aligned} \dot{u} &= Au, \\ u(0) &= u_0 \in X. \end{aligned} \tag{2.1}$$

If A is bounded, this equation has a unique global solution

$$u \in C^1(\mathbf{R}; X).$$

The solution can be written as

$$u(t) = e^{At} u_0,$$

where the exponential is defined by an operator-norm convergent power series,

$$e^{At} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n.$$

Using the inequality

$$\left\| e^{At} \right\| \leq \sum_{n=0}^{\infty} \frac{t^n}{n!} \|A\|^n \leq e^{\|A\| t}$$

we see that

$$\|u(t)\| \leq e^{\|A\| t} \|u_0\|. \tag{2.2}$$

Thus the norm of the solution grows at most exponentially in time.

This result can be looked at from the point of view of spectral theory. Suppose λ is an eigenvalue of A and $\varphi \in X$ is an eigenvector, so that

$$A\varphi = \lambda \varphi.$$

A particular solution of the linear evolution equation (2.1) is then

$$u(t) = e^{\lambda t} \varphi.$$

If A is bounded, then $|\lambda| \leq \|A\|$ and this solution obviously satisfies (2.2).

If A is unbounded, the spectrum of A contains complex numbers with arbitrarily large magnitude. In that case, (2.1) may have solutions which grow at an arbitrarily fast exponential rate when the initial value problem is ill-posed. The initial value problem is only well-posed forwards in time if the real part of the spectrum of A is bounded above so that there is a limit on the growth rate of solutions.

The Hille-Yoshida theorem provides precise necessary and sufficient conditions for a linear operator A to generate a well-posed evolution equation with solution operators $S(t) = e^{At}$. The main condition is a condition on the resolvent operator of A which limits the growth rate of solutions. For completeness we state the theorem here. First we give an abstract definition of a semigroup.

Definition 2.4 *A* C^0 **semigroup** *on a Banach space* X *is a family of bounded linear operators* $S(t) : X \to X$*, defined for* $t \geq 0$*, such that:*

(*a*) $S(t_1)S(t_2) = S(t_1 + t_2) \qquad t_1, t_2 \geq 0;$

(*b*) $S(0) = I;$

(*c*) $\|S(t+h)u - S(t)u\| \to 0 \qquad$ *as* $h \to 0$ *for all* $t \geq 0$.

The semigroup is said to be a C^0 **contraction semigroup** *if, in addition,*

$$(d) \quad \|S(t)\| \leq 1 \qquad \textit{for all } t \geq 0.$$

The term C^0 refers to the condition in (c) which states that for each $u_0 \in X$, the function $u(t) = S(t)u_0$ is a continuous (C^0) function of time with values in X. Note that the family of operators $S(t)$ is *not* required to be continuous in the operator norm. Thus, although

$$S(t)u_0 \to u_0 \qquad \text{as } t \downarrow 0$$

for each $u_0 \in X$, it need not be true that $S(t) \to I$ in the operator norm as $t \downarrow 0$. In fact this uniform convergence only happens when A is a bounded operator.

We can go from the solution operators S to the vector field A by means of the following definition.

Definition 2.5 *Let $S(t)$ be a C^0 semigroup on X. We define an operator $A : D(A) \subset X \to X$ by*

$$Au = \lim_{h \to 0+} \left(\frac{S(h)u - u}{h} \right),$$

where $D(A)$ is the set of $u \in X$ such that this limit exists. We call A the **generator** *of S.*

Roughly speaking, contraction semigroups correspond to equations in which there is no growth of solutions. There is no essential loss of generality in restricting our attention to contraction semigroups. If $S(t)$ is any C^0 semigroup with generator A, then there exists a constant $\omega \geq 0$ and an equivalent norm on X such that

$$\|S(t)\| \leq e^{\omega t}.$$

It follows that

$$\bar{S}(t) = e^{-\omega t} S(t)$$

is a contraction semigroup with generator $\bar{A} = A - \omega I$. Properties of $S(t)$ can be recovered from properties of the contraction semigroup $\bar{S}(t)$.

The basic result in semigroup theory is the following characterization of those operators A which generate a contraction semigroup S.

Theorem 2.6 *(Hille-Yoshida.) An operator $A : D(A) \subset X \to X$ is the generator of a C^0 contraction semigroup if and only if it satisfies the following conditions.*

1. *$D(A)$ is dense in X and A is closed.*
2. *Any $\lambda \in \mathbf{R}$ with $\lambda > 0$ belongs to the resolvent set of A.*
3. *For all $\lambda > 0$,*
$$\|\lambda I - A\| \leq \frac{1}{\lambda}.$$

For general operators A the conditions in the Hille-Yoshida theorem are not simple to check. The following result gives simple sufficient conditions in the case of Hilbert space operators.

Theorem 2.7 *Let $A : D(A) \subset H \to H$ be a closed, densely defined operator acting in a Hilbert space H. If*

$$\begin{aligned} (Au, u) &\leq 0, \qquad \textit{for all } u \in D(A) \\ (A^*v, v) &\leq 0, \qquad \textit{for all } v \in D(A^*) \end{aligned}$$

then A generates a C^0 contraction semigroup on H.

Two important special cases of this result are negative self-adjoint operators and skew-adjoint operators,

$$\begin{aligned} &\text{(a)} \quad A \leq 0, \qquad A^* = A, \\ &\text{(b)} \quad A^* = -A. \end{aligned}$$

In (a), $A \leq 0$ means that $(Au, u) \leq 0$ for all $u \in D(A)$. The simplest example of (a) is the diffusion equation and the simplest example of (b) is the Schrödinger equation (see below). In the case when A is skew-adjoint, the solution operator $U(t) = e^{At}$ is unitary (meaning that $U^*(t) = U^{-1}(t)$) and is defined for all $t \in \mathbf{R}$, not just for $t \geq 0$. In fact, Stone's theorem asserts that A generates a C^0 unitary group $U(t)$ on a Hilbert space H if and only if A is skew-adjoint. Note that skew-adjointness includes the condition that $D(A) = D(A^*)$.

Examples.

1. Any linear operator on a finite dimensional space $A : \mathbf{R}^m \to \mathbf{R}^m$ is bounded. Thus the finite dimensional system of ODE's $\dot{u} = Au$ has a unique global solution $u(t) = e^{At}u(0)$.

2. Suppose $k : \Omega \times \Omega \to \mathbf{R}$ satisfies

$$\int_{\Omega\times\Omega} |k(x,y)|^2 \, dxdy < \infty.$$

   Then the integral operator $K : L^2(\Omega) \to L^2(\Omega)$ defined by

$$Ku(x) = \int_\Omega k(x,y)u(y) \, dy$$

   is a bounded linear operator. Hence the evolution equation

$$\begin{aligned} \dot{u}(x,t) &= \int_\Omega k(x,y)u(y,t) \, dy, \\ u(x,0) &= u_0(x) \in L^2(\Omega), \end{aligned}$$

   has a unique global solution $u(t) = e^{Kt}u_0$ with

$$u \in C^1\left(\mathbf{R}; L^2(\Omega)\right).$$

3. Consider the diffusion equation,

$$\begin{aligned} &\dot{u} = \Delta u, \\ &u = 0, \qquad \text{on } \partial\Omega, \\ &u(0) = u_0 \in L^2(\Omega). \end{aligned}$$

   Suppose we use the state space $H = L^2(\Omega)$. We define

$$A = \Delta : H^2(\Omega) \cap H_0^1(\Omega) \subset H \to H.$$

   This operator is negative and self-adjoint, so it generates a contraction semigroup

$$S(t) = e^{\Delta t} : H \to H, \qquad t \geq 0.$$

4. The backwards diffusion equation is

$$\begin{aligned} &\dot{u} = -\Delta u, \\ &u = 0, \qquad \text{on } \partial\Omega, \\ &u(0) = u_0 \in L^2(\Omega). \end{aligned}$$

The eigenvalues λ_n of $-\Delta$ are real and $\lambda_n \to +\infty$ as $n \to \infty$. The initial value problem is therefore ill-posed.

Note that the diffusion equation differs in an essential way from an ODE because we can only solve in one time direction for general data $u_0 \in L^2$.

5. The linear Schrödinger equation

$$\begin{aligned} &\dot{u} = i\Delta u, \\ &u = 0, \qquad \text{on } \partial\Omega, \\ &u(0) = u_0 \in L^2(\Omega). \end{aligned}$$

has a unitary solution operator

$$U(t) = e^{i\Delta t}.$$

This equation can be solved forwards and backwards in time because the eigenvalues of $i\Delta$ are pure imaginary.

6. Semi-linear PDE's

$$\dot{u} = Au + f(u)$$

cannot be solved directly by the ODE existence theorem. However, we can remove the unbounded linear part of the vector field and reformulate the PDE as an integral equation

$$u(t) = u(0) + \int_0^t e^{A(t-s)} f\left(u(s)\right)\, ds.$$

Under suitable conditions on A and f, this equation can be solved by the contraction mapping principle (see below).

2.6 A nonlinear PDE

Let us consider an example of a nonlinear PDE which can be treated directly by means of the ODE theory. The PDE is a nonlocal, nonlinear Schrödinger equation for complex-valued functions $u(x,t)$, $v(x,t)$,

$$\begin{aligned} iu_t &= |u|^2 u + v, \\ -\Delta v + v &= u. \end{aligned}$$

We suppose that $x \in \mathbf{R}^d$.

If $u \in L^2(\mathbf{R}^d)$, say, we can solve the second equation uniquely for $v \in H^2(\mathbf{R}^d)$. We write the solution as

$$v = Au, \qquad A = (I - \Delta)^{-1}.$$

Thus u satisfies a nonlinear integro-differential equation,

$$\begin{aligned} iu_t &= |u|^2 u + Au, \\ u(0) &= u_0. \end{aligned} \tag{2.3}$$

For long waves which vary slowly in x we have the formal approximation

$$(I - \Delta)^{-1} \sim I + \Delta.$$

Thus the long wave limit of (2.3) is a Nonlinear Schrödinger equation equation,

$$iu_t = \Delta u + \left(1 + |u|^2\right) u.$$

We want to regard (2.3) as an evolution equation

$$\dot{u} = f(u)$$

where

$$f(u) = -i\left(|u|^2 u + Au\right)$$

is a Lipschitz continuous vector field on a space X.

The nonlinear term causes some difficulties here. For example, if we suppose that $u \in L^p$ then $|u|^2 u \in L^{p/3}$. Thus we do not have $f : L^p \to L^p$ for any $p < \infty$. One way around this difficulty is to consider smooth solutions with $u(t) \in H^k(\mathbf{R}^d)$. When $k > d/2$, $H^k(\mathbf{R}^d)$ is an algebra imbedded in $C_0(\mathbf{R}^d)$. If $\|u\|, \|v\| \leq R$, where $\|\cdot\|$ denotes the H^k-norm, then we have

$$\begin{aligned} &\||u|^2 u - |v|^2 v\| \leq C(R)\|u - v\|, \\ &\|Au - Av\| \leq \|u - v\|. \end{aligned}$$

Thus, $f : H^k(\mathbf{R}^d) \to H^k(\mathbf{R}^d)$ and f is Lipschitz on bounded sets. It follows from the ODE existence theorem that equation (2.3) has a unique local solution

$$u \in C^1\left(-\delta, \delta; H^k(\mathbf{R}^d)\right).$$

The existence time δ depends only on the H^k-norm of the initial data. The global existence of H^k-valued solutions would follow from global a priori estimates of the H^k norm of local solutions.

Chapter 3

Sobolev Spaces and Laplace's Equation

Sobolev spaces provide the simplest and most useful setting for the application of functional analytic methods to the theory of partial differential equations. Sobolev spaces consist of functions with integrable derivatives.

Once the notion of weak solutions of Laplace's equation is formulated in terms of Sobolev spaces, the proof of the existence and uniqueness of weak solutions is almost trivial.

3.1 L^p spaces

Let Ω be an open set in $\mathbf{R}^d$. In particular, we could have $\Omega = \mathbf{R}^d$. We say that a function $f : \Omega \to \mathbf{R}^d$ is **locally integrable** if it is Lebesgue measurable and if

$$\int_K f(x)\, dx$$

is finite for all compact sets $K \subset \Omega$. We denote the space of locally integrable functions by $L^1_{\text{loc}}(\Omega)$.

For $1 \le p < \infty$ we define the space

$$L^p(\Omega) = \left\{ f : \Omega \to \mathbf{R} | \int_\Omega |f(x)|^p \, dx < \infty \right\}.$$

For $p = \infty$ we define the space $L^\infty(\Omega)$ of essentially bounded functions. Sometimes, we will consider complex-valued functions $f : \Omega \to \mathbf{C}$ or vector-valued functions $f : \Omega \to \mathbf{R}^m$. In that case, $|\cdot|$ denotes the absolute value of a complex number or any convenient norm on $\mathbf{R}^m$. It should be clear from the context when this is the case.

Two functions are regarded as the same element of L^p if they differ on a set of measure zero.

For $1 \le p < \infty$, we define a norm on $L^p(\Omega)$ by

$$\|f\|_{L^p(\Omega)} = \left(\int_\Omega |f(x)|^p \, dx \right)^{1/p}.$$

For $p = \infty$ we define

$$\|f\|_{L^\infty(\Omega)} = \sup_{x \in \Omega} |f(x)|$$

Here and below, sup denotes the essential supremum, that is the infimum of the supremum of functions which are equal to f almost everywhere. We will sometimes use the shorter notation

$$\|f\|_p = \|f\|_{L^p(\Omega)}.$$

Theorem 3.1 *The space L^p is a Banach space for all $1 \le p \le \infty$.*

For $1 \le p \le \infty$, we define the **Hölder conjugate** $1 \le p' \le \infty$ of p by

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

If $f \in L^p$ and $g \in L^{p'}$ then we have Hölder's inequality

$$\left| \int_\Omega f g \, dx \right| \le \|f\|_p \|g\|_{p'}.$$

If $1 \le p < \infty$, then we can identify $L^{p'}$ with the dual space of L^p. Note that although L^∞ is the dual space of L^1, it is not true that L^1 is the dual space of L^∞. Thus L^1 and L^∞ are not reflexive. If $p = 2$, then $p' = 2$ and L^2 is self-dual. In fact, L^2 is a Hilbert space. For $p = 2$, Hölder's inequality reduces to the Cauchy-Schwartz inequality.

A useful generalization of Hölder's inequality is the following interpolation inequality. Suppose that $p \le r \le q$. Define $0 \le \theta \le 1$ by the equation

$$\frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{q}.$$

If $f \in L^p(\Omega) \cap L^q(\Omega)$, then $f \in L^r(\Omega)$ and

$$\|f\|_{L^r} \le \|f\|_{L^p}^{\theta} \|f\|_{L^q}^{1-\theta}.$$

The space of locally integrable functions $L^1_{\mathrm{loc}}(\Omega)$ is the "largest" space and contains $L^p(\Omega)$ for all $1 \le p \le \infty$. If Ω has finite measure (for example, if Ω is bounded), and if $1 \le p \le q \le \infty$ we have the inclusions

$$L^1(\Omega) \supset L^p(\Omega) \supset L^q(\Omega) \supset L^\infty(\Omega).$$

This fact follows immediately from Hölder's inequality, since

$$\|f\|_p = \left(\int_\Omega 1 \cdot |f|^p \, dx\right)^{1/p} \le \left(\int_\Omega 1^r \, dx\right)^{1/r} \left(\int_\Omega |f|^q \, dx\right)^{p/q} \le |\Omega|^{1/r} \|f\|_q^p.$$

Here $|\Omega|$ is the Lebesgue measure of Ω and

$$\frac{p}{q} + \frac{1}{r} = 1.$$

If the measure of Ω is infinite (e.g. if $\Omega = \mathbf{R}^d$) then this inclusion does not hold.

Examples.

1. Consider the function
$$|x|^{-a} : \mathbf{R}^d \to \mathbf{R}$$
where
$$|x| = \left(x_1^2 + \ldots + x_d^2\right)^{1/2}.$$
Let $\Omega = B_1(0)$ be the unit ball,
$$B_1(0) = \left\{x \in \mathbf{R}^d : |x| < 1\right\}.$$
Then $|x|^{-a}$ belongs to $L^1(\Omega)$ for $a < d$ since then
$$\int_\Omega |x|^{-a} \, dx = \omega_d \int_0^1 |x|^{d-a-1} \, d|x| < \infty.$$
Here, ω_d is the surface area of the unit sphere in $\mathbf{R}^d$. This area is given by
$$\omega_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}.$$
Note that $\Gamma(1) = 1$ and $\Gamma(3/2) = \sqrt{\pi}/2$ so this formula reduces to the elementary formula for the circumference of the unit circle when $d = 2$ and the area of the unit sphere when $d = 3$.
More generally, $|x|^{-a} \in L^1(\Omega)$ if $a < d$ for any bounded set Ω.

2. The function $|x|^{-a}$ is locally integrable on $\mathbf{R}^d$ for $a < d$, but it does not belong to $L^1(\mathbf{R}^d)$ for any value of a since if $a \geq d$ the function is not locally integrable in any neighborhood of the origin, while if $a \leq d$ the integral does not converge at infinity.

3. If Ω is a bounded set, then $|x|^{-a} \in L^p(\Omega)$ if $a < d/p$.

3.2 Distributional derivatives

The theory of partial differential equations is greatly simplified by the use of distributional or weak derivatives rather than classical derivatives which are defined pointwise. Distributions are dual to test functions and they possess distributional derivatives of all orders. Distributions which correspond to locally integrable functions are called regular distributions. If the distributional derivative g of a regular distribution f turns out to be a regular distribution, then we say that f is weakly differentiable with weak derivative g. This notion provides an extension of the classical pointwise derivative with much better functional analytic properties. Sobolev spaces consist of functions whose weak derivatives belong to L^p spaces.

Definition 3.1 *A* **test funtion** $\varphi : \Omega \to \mathbf{R}$ *is a function with continuous partial derivatives of all orders whose support (that is, the closure of the set where* $\varphi(x) \neq 0$*) is a compact subset of* Ω*. We denote the set of test functions on* Ω *by* $C_c^\infty(\Omega)$*. A* **distribution** T *on* Ω *is a continuous linear map*

$$T : C_c^\infty(\Omega) \to \mathbf{R}.$$

We denote the value of a distribution T acting on a test function φ by $\langle T, \varphi \rangle$. The condition that T is continuous requires the introduction of an appropriate topology on the space of test functions. We omit a discussion of this topology since we will not need to use it.

Examples.

1. If $f \in L^1_{\mathrm{loc}}(\Omega)$ then the map $F : C_c^\infty(\Omega) \to \mathbf{R}$ given by

$$\langle F, \varphi \rangle = \int_\Omega f(x)\varphi(x)\, dx$$

is a distribution on Ω. Conversely, given the map F, we can recover the pointwise values of the function f almost everywhere in Ω. We may therefore identify F with f. Any distribution which is associated with a locally integrable function in this way is called a **regular distribution**.

2. The simplest example of a distribution which is not a regular distribution is the **delta function**. (Strictly speaking, we should say "delta distribution.") For any $\xi \in \Omega$ we define the delta function δ_ξ supported at ξ by
$$\langle \delta_\xi, \varphi \rangle = \varphi(\xi).$$

General distributions are used extensively in the theory of linear PDE's. However, there is no consistent way to define the product of general distributions (e.g. δ^2 doesn't make any sense). Thus the use of general distributions in the theory of nonlinear PDE's is severely curtailed. It is possible to define the product of regular distributions by taking the pointwise product of the associated functions.

If $f : \Omega \to \mathbf{R}$ is smooth funtion, then an integration by parts shows that
$$\int_\Omega f_{x_i} \varphi \, dx = -\int_\Omega f \varphi_{x_i} \, dx.$$
The boundary terms vanish because φ has compact support in Ω. This result motivates the following definition of the distributional derivative.

Definition 3.2 *The* **distributional derivative** T_{x_i} *of a distribution* T *with respect to* x_i *is defined by*
$$\langle T_{x_i}, \varphi \rangle = -\langle T, \varphi_{x_i} \rangle.$$

Our main interest is in the case when T and T_{x_i} are regular distributions which are associated with functions f and g_i, respectively. This leads to the following definition of the weak derivative.

Definition 3.3 *Suppose that* $f, g_i \in L^1_{\mathrm{loc}}(\Omega)$ *are such that*
$$\int_\Omega g_i \varphi \, dx = -\int_\Omega f \varphi_{x_i} \, dx$$
for all test functions $\varphi \in C_c^\infty(\Omega)$. *Then we say that* f *is weakly differentiable with respect to* x_i *and we call* $g_i = f_{x_i}$ *the* **weak derivative** *of* f *with respect to* x_i.

Note that the weak derivative is only defined pointwise up to a set of measure zero. In the future we will use the weak derivative as our primary notion of derivative, so we will often drop the qualification "weak."

Examples.

1. The weak derivative of the function

$$f(x) = \frac{1}{|x|^a}$$

with respect to x_i is given by

$$g_i = -a\frac{x_i}{|x|}\frac{1}{|x|^{a+1}}$$

provided that g_i is locally integrable. For $x \in \mathbf{R}^d$ this is the case when

$$a < d - 1.$$

For example, in one space dimension, if $Df \in L^p$ for some $p > 1$, then $a < 0$ so f is continuous. This turns out to be a general result — see the Sobolev imbedding theorem below. In higher space dimensions, the function can be unbounded near the origin, yet still be weakly differentiable. The strength of an allowable singularity in a weakly differentiable function increases with the number of space dimensions d.

2. We define the step function $H : \mathbf{R} \to \mathbf{R}$ by

$$H(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

The distributional derivative of H is the delta function supported at 0. Thus, although H is a regular distribution, its derivative is not.

3.3 Sobolev spaces

Sobolev spaces are spaces of functions whose weak derivatives belong to L^p. We use the following compact notation for partial derivatives. Let

$$x = (x_1, \ldots, x_d) \in \mathbf{R}^d$$

and let

$$\alpha = (\alpha_1, \ldots \alpha_d) \in \mathbf{N}^d$$

be a multi-index of non-negative integers. We define

$$|\alpha| = \alpha_1 + \ldots + \alpha_d.$$

Given a multi-index α, we define the corresponding partial derivative of order $|\alpha|$ by

$$D^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_d}\right)^{\alpha_d}.$$

Definition 3.4 *Let k be a positive integer and $1 \leq p \leq \infty$. Let Ω be an open subset of $\mathbf{R}^d$. The* **Sobolev space** *$W^{k,p}(\Omega)$ consists of functions $f : \Omega \to \mathbf{R}$ such that $D^\alpha f \in L^p(\Omega)$ for all partial derivatives of order $0 \leq |\alpha| \leq k$. We define a norm on $W^{k,p}(\Omega)$ by*

$$\|f\|_{W^{k,p}(\Omega)} = \left(\sum_{0 \leq |\alpha| \leq k} \int_\Omega |D^\alpha f(x)|^p \, dx\right)^{1/p}$$

when $1 \leq p < \infty$, and by

$$\|f\|_{W^{k,\infty}(\Omega)} = \max_{0 \leq |\alpha| \leq k} \left\{\sup_{x \in \Omega} |D^\alpha f(x)|\right\}$$

when $p = \infty$.

When $p = 2$, corresponding to the case of square integrable functions, we write $W^{k,2}(\Omega) = H^k(\Omega)$. We define an inner product on $H^k(\Omega)$ by

$$(f, g)_{H^k(\Omega)} = \sum_{0 \leq |\alpha| \leq k} \int_\Omega D^\alpha f(x) D^\alpha g(x) \, dx.$$

The space $W^{k,p}(\Omega)$ is a Banach space and $H^k(\Omega)$ is a Hilbert space. We will sometimes use the abbreviation

$$\|f\|_{k,p} = \|f\|_{W^{k,p}(\Omega)}$$

Next, we define a space of Sobolev functions which "vanish on the boundary of Ω."

Definition 3.5 *The closure of $C_c^\infty(\Omega)$ in $W^{k,p}(\Omega)$ is denoted by*

$$W_0^{k,p}(\Omega) = \overline{C_c^\infty(\Omega)}.$$

We also define

$$H_0^k(\Omega) = W_0^{k,2}(\Omega).$$

Informally, we can think of $W_0^{k,p}(\Omega)$ as the $W^{k,p}(\Omega)$-functions whose derivatives of order less than or equal to $k-1$ vanish on the boundary $\partial\Omega$ of Ω.

Compactly supported functions are dense in $W^{k,p}(\mathbf{R}^d)$, so that there is no distinction between $W_0^{k,p}(\mathbf{R}^d)$ and $W^{k,p}(\mathbf{R}^d)$.

The definition of the boundary values of Sobolev functions which do not vanish on the boundary is non-trivial. The boundary of a smooth set has measure zero, but Sobolev functions are not necessarily continuous and they are are defined pointwise only up to sets of measure zero. The Trace Theorem below gives a way to assign boundary values to Sobolev functions.

Examples.

1. For $u \in W^{1,p}(\Omega)$ the Dirichlet boundary condition

   $$u = 0 \qquad \text{on } \partial\Omega$$

   can be formulated precisely by requiring that

   $$u \in W_0^{1,p}(\Omega).$$

2. If $u \in W_0^{2,p}(\Omega)$, then both u and its normal derivative,

   $$u = \frac{\partial u}{\partial n} = 0,$$

   vanish on the boundary $\partial\Omega$.

3. The right way to formulate the boundary condition $u = 0$ on $\partial\Omega$ for $u \in W^{2,p}(\Omega)$ is by requiring that

   $$u \in W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega).$$

Next, we define Sobolev spaces of negative integer orders.

Definition 3.6 *For any positive integer k and any $1 \le p < \infty$ we define the negative Sobolev space $W^{-k,p'}(\Omega)$ to be the dual space of $W_0^{k,p}(\Omega)$. Here, p' is the Hölder conjugate of p. That is, $f \in W^{-k,p'}(\Omega)$ is a continuous linear map*

$$f : W_0^{k,p}(\Omega) \to \mathbf{R}, \qquad f : u \mapsto \langle f, u \rangle.$$

We define a norm in $W^{-k,p'}(\Omega)$ by

$$\|f\|_{W^{-k,p'}} = \sup_{\substack{u \in W_0^{k,p} \\ u \neq 0}} \frac{\langle f, u \rangle}{\|u\|}.$$

Elements of $W^{-k,p'}(\Omega)$ are distributions which can be extended continuously from test functions to functions in $W_0^{k,p}(\Omega)$. The dual space of $W^{k,p}(\Omega)$ is not a space of distributions because a continuous linear functional on $W^{k,p}(\Omega)$ is not completely determined by its values on test functions.

Any distribution $f \in W^{-k,p'}(\Omega)$ can be written non-uniquely in the form

$$f = \sum_{|\alpha| \le k} D^\alpha g_\alpha$$

for some functions $g_\alpha \in L^{p'}(\Omega)$. The action of f on a $W_0^{k,p}$-function u is given by

$$\langle f, u \rangle = \sum_{|\alpha| \le k} (-1)^{|\alpha|} \int_\Omega g_\alpha D^\alpha u \, dx.$$

More generally, it is possible to define Sobolev spaces $W^{s,p}$ of fractional order for any $s \in \mathbf{R}$ and $p \in [1, \infty]$. These spaces arise naturally in connection with the trace theorem below, but we will not give any details here.

3.4 Properties of Sobolev spaces

In this section, we summarize the main facts about Sobolev spaces without proof. These facts involve the approximation of Sobolev functions by smooth functions (density theorems), the integrability or continuity properties of Sobolev functions (imbedding theorems), compactness conditions (the Rellich-Kondrachov theorem), and boundary values of Sobolev functions (trace theorems). We also state some other useful inequalities.

We use $C(\bar{\Omega})$ to denote the space of uniformly continuous functions on Ω with the sup-norm. In the case of $\mathbf{R}^d$ we use $C_0(\mathbf{R}^d)$ to denote the space of continuous functions which tend to zero as $x \to \infty$. This space is the closure of $C_c^\infty(\mathbf{R}^d)$ in $L^\infty(\mathbf{R}^d)$. The space $C^k(\bar{\Omega})$ consists of functions whose partial derivatives of order less than or equal to k are uniformly continuous, and $C^\infty(\bar{\Omega})$ consists of functions with uniformly continuous derivatives of all orders in Ω.

In the theorems stated below, we consider two types of domains: (a) $\Omega = \mathbf{R}^d$; (b) Ω is a bounded, open subset of $\mathbf{R}^d$ with a smooth boundary $\partial\Omega$. It is frequently possible to consider more general domains, but this complicates the statements of the theorems.

The order k is a positive integer and $1 \le p \le \infty$, unless stated otherwise.

3.4.1 Density Theorem

Theorem 3.2 *(a) The space of test functions $C_c^\infty(\mathbf{R}^d)$ is dense in $W^{k,p}(\mathbf{R}^d)$; (b) the space of test functions $C_c^\infty(\Omega)$ is dense in $W_0^{k,p}(\Omega)$; (c) the space of smooth functions $C^\infty(\bar{\Omega})$ is dense in $W^{k,p}(\Omega)$.*

Part (c) requires some smoothness of the domain Ω. For general domains Meyers and Serrin proved that $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$. Note that functions in $C^\infty(\Omega)$ are not required to be smooth up to the boundary. For instance, $1/x \in C^\infty(0,1)$ although it does not belong to $C^\infty[0,1]$ or to $L^1(0,1)$.

3.4.2 Imbedding Theorem

To motivate the imbedding theorem, we consider functions $u : \mathbf{R}^d \to \mathbf{R}$ and ask when it is possible to have an estimate of the form

$$\|u\|_{L^q} \le C\|Du\|_{L^p}. \tag{3.1}$$

Let $\lambda > 0$. We define the rescaled function

$$u_\lambda(x) = u(\lambda x).$$

A simple calculation shows that

$$\begin{aligned} \|u_\lambda\|_{L^q} &= \lambda^{-d/q}\|u\|_{L^q}, \\ \|Du_\lambda\|_{L^p} &= \lambda^{1-d/p}\|Du\|_{L^p}. \end{aligned}$$

These norms have to scale according to the same exponent if the estimate in (3.1) is to hold. This occurs only if $p < d$ and $q = p^*$ where

$$p^* = \frac{dp}{d-p}.$$

We call p^* the **Sobolev conjugate** of p. Note that

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{d}.$$

The inequality (3.1) is, in fact, true in this case. It follows that every $W^{1,p}$ function is an L^{p^*} function, and the imbedding $J : W^{1,p} \to L^{p^*}$ is continuous.

The proof of this fact depends on a clever application of Hölder's inequality. First one proves the result for test functions. The inequality then follows by density for arbitrary Sobolev functions.

Theorem 3.3 *(Sobolev-Gagliardo-Nirenberg.) Suppose that $p < d$ and p^* is the Sobolev conjugate of p.*
(a) If $u \in W^{1,p}(\mathbf{R}^d)$, then $u \in L^{p^*}(\mathbf{R}^d)$ and there exists a constant $C = C(p,d)$ such that

$$\|u\|_{L^{p^*}(\mathbf{R}^d)} \leq C\|Du\|_{L^p(\mathbf{R}^d)}.$$

(b) If $u \in W^{1,p}(\Omega)$ and $1 \leq q \leq p^*$ then $u \in L^q(\Omega)$ and there exists a constant $C = C(p,q,d,\Omega)$ such that

$$\|u\|_{L^q(\Omega)} \leq C\|u\|_{W^{1,p}(\Omega)}.$$

Note that for $\mathbf{R}^d$, the estimate only involves the L^p-norm of the derivative of u and not the norm of u itself. For a bounded domain, once $u \in L^{p^*}$ it follows that $u \in L^q$ for all $1 \leq q \leq p^*$. In the case of $\mathbf{R}^d$, it follows by interpolation that $u \in L^q$ for $p \leq q \leq p^*$:

Corollary 3.1 *Let $1 \leq p < d$ and $p \leq q \leq p^*$. If $u \in W^{1,p}(\mathbf{R}^d)$, then $u \in L^q(\mathbf{R}^d)$ and there exists a constant $C = C(p,q,d)$ such that*

$$\|u\|_{L^q} \leq C\|u\|_{W^{1,p}}.$$

Proof. Choose $0 \leq \theta \leq 1$ such that

$$\frac{1}{q} = \frac{\theta}{p} + \frac{1-\theta}{p^*}.$$

Then by the interpolated Hölder inequality and Young's inequality we have

$$\|u\|_{L_q} \leq \|u\|_{L^p}^{\theta}\|u\|_{L^{p^*}}^{1-\theta} \leq \|u\|_{L^p} + \|u\|_{L^{p^*}}.$$

The result then follows from the Sobolev imbedding theorem. QED

In the case when $d < p < \infty$, functions in $W^{1,p}$ are continuous and we can estimate their sup-norm by their $W^{1,p}$-norm.

Theorem 3.4 *Suppose that $d < p < \infty$.*
(a) If $u \in W^{1,p}(\mathbf{R}^d)$, then $u \in C_0(\mathbf{R}^d)$ and there exists a constant $C = C(p,d)$ such that
$$\|u\|_{L^\infty(\mathbf{R}^d)} \leq C\|Du\|_{L^p(\mathbf{R}^d)}.$$
(b) If $u \in W^{1,p}(\Omega)$ then $u \in C(\bar{\Omega})$ and there exists a constant $C = C(p,d,\Omega)$ such that
$$\|u\|_{L^\infty(\Omega)} \leq C\|u\|_{W^{1,p}(\Omega)}.$$

This theorem gives the simplest conclusion of the imbedding theorem for $p > d$. In fact, a stronger result is true: the functions are Hölder continuous with exponent $\beta = 1 - d/p$. Before stating this result, we briefly summarize the definition of Hölder spaces.

Let Ω be an open subset of $\mathbf{R}^d$ with closure $\bar{\Omega}$. We say that $u : \Omega \to \mathbf{R}$ is **Hölder continuous** on Ω with exponent $0 < \beta \leq 1$ if

$$\sup_{\substack{x,y\in\Omega \\ x\neq y}} \frac{|u(x) - u(y)|}{|x-y|^\beta} < \infty.$$

If $\beta = 1$ then u is Lipschitz continuous in Ω. Any Hölder continuous function is continuous, but not conversely. The Banach space $C^{0,\beta}(\bar{\Omega})$ consists of all bounded Hölder continuous functions on Ω with the norm

$$\|u\|_{C^{0,\beta}(\bar{\Omega})} = \sup_{x\in\Omega} |u(x)| + \sup_{\substack{x,y\in\Omega \\ x\neq y}} \frac{|u(x) - u(y)|}{|x-y|^\beta}.$$

If k is a positive integer and $0 < \beta \leq 1$, we define $C^{k,\beta}(\bar{\Omega})$ to be the space of functions which are k times continuously differentiable in Ω, with bounded derivatives, and whose kth-order derivatives are Hölder continuous with exponent β. This space is a Banach space with the norm

$$\|u\|_{C^{k,\beta}(\bar{\Omega})} = \max_{0\leq|\alpha|\leq k} \left\{ \sup_{x\in\Omega} |D^\alpha u(x)| \right\} + \sup_{\substack{x,y\in\Omega \\ x\neq y, |\alpha|=k}} \frac{|D^\alpha u(x) - D^\alpha u(y)|}{|x-y|^\beta}.$$

Theorem 3.5 *(Morrey.) Suppose that $d < p < \infty$. Let*

$$\beta = 1 - \frac{d}{p}.$$

(a) If $u \in W^{1,p}(\mathbf{R}^d)$, then $u \in C^{0,\beta}(\mathbf{R}^d)$ and there exists a constant $C = C(p,d)$ such that

$$\|u\|_{C^{0,\beta}(\mathbf{R}^d)} \leq C\|Du\|_{L^p(\mathbf{R}^d)}.$$

(b) If $u \in W^{1,p}(\Omega)$ then $u \in C^{0,\beta}(\Omega)$ and there exists a constant $C = C(p,d,\Omega)$ such that

$$\|u\|_{C^{0,\beta}(\Omega)} \leq C\|u\|_{W^{1,p}(\Omega)}.$$

A general Sobolev imbedding theorem follows by repeated application of these results.

Theorem 3.6 *(a) Suppose $kp < d$ and let*

$$q = \frac{dp}{d-kp}.$$

Then

$$\|u\|_{L^q} \leq C\|u\|_{W^{k,p}}.$$

(b) Suppose $kp > d$ and let

$$m = k - \left[\frac{d}{p}\right] - 1.$$

If d/p is not an integer, let

$$\beta = \left[\frac{d}{p}\right] - \frac{d}{p} + 1$$

be the fractional part of d/p. Here the square brackets denote the integer part. If d/p is an integer, let

$$\beta = 1 - \varepsilon$$

for any $0 < \varepsilon < 1$. Then

$$\|u\|_{C^{m,\beta}} \leq C\|u\|_{W^{k,p}}.$$

3.4.3 Compactness theorems

It is a general principle that a set of functions whose derivatives are uniformly bounded is compact. The Sobolev-space version of this principle is the Rellich-Kondrachov theorem which states that $W^{k,p}(\Omega)$ is compactly imbedded in $L^q(\Omega)$ for $q < p^*$. The boundedness of the domain Ω and the condition that q is strictly less than p^* are both essential for compactness. In the critical case $q = p^*$ the imbedding is continuous but not compact.

Theorem 3.7 *Let Ω be a smooth bounded domain in $\mathbf{R}^d$. (a) Suppose that $1 \leq p < d$ and $1 \leq q < p^*$. Then bounded sets in $W^{1,p}(\Omega)$ are precompact in $L^q(\Omega)$. (b) Suppose $p > d$. Then bounded sets in $W^{1,p}(\Omega)$ are precompact in $C(\bar{\Omega})$.*

In particular, suppose that $\{u_n\}_{n=1}^{\infty}$ is a sequence of functions in $W^{1,p}(\Omega)$ such that

$$\|u_n\|_{W^{1,p}} \leq C$$

for some constant C which is independent of n. If $p < d$ and $1 \leq q < p^*$ then there is a subsequence $\{u_{n_j}\}_{j=1}^{\infty}$ which converges strongly in $L^q(\Omega)$. If $p > d$, then there is a uniformly convergent subsequence.

If $p > d$ and $0 < \beta < 1 - d/p$ then the imbedding of $W^{1,p}(\Omega)$ into $C^{0,\beta}(\bar{\Omega})$ is compact.

General compactness theorems follow by repeated application of this result. For example, if $kp < d$ then $W^{k,p}(\Omega)$ is compactly imbedded in $L^q(\Omega)$ for any $1 \leq q < dp/(d - kp)$, while if $kp > d$ then $W^{k,p}(\Omega)$ is compactly imbedded in $C(\bar{\Omega})$.

3.4.4 Trace theorems

There is no sensible way to assign boundary values $u|_{\partial\Omega}$ to a general function $u \in L^p(\Omega)$. Functions in L^p are defined pointwise only almost everywhere, and the boundary $\partial\Omega$ has measure zero. The situation is different for Sobolev functions. If $u \in W^{k,p}(\Omega)$ then one can assign boundary values to the derivatives of u of order less than or equal to $k - 1$. It is not possible to define boundary values of kth order derivatives, however, since they are just L^p functions.

Theorem 3.8 *There is a surjective bounded linear operator*

$$\gamma : W^{1,p}(\Omega) \to W^{1-1/p,p}(\partial\Omega)$$

such that
$$\gamma u = u|_{\partial\Omega} \qquad \textit{if } u \in W^{1,p}(\Omega) \cup C(\bar{\Omega}).$$

Note the "loss of $1/p$ derivatives" in restricting a Sobolev function to the boundary. For example, the boundary values of a function in $H^1(\Omega)$ belong to $H^{1/2}(\partial\Omega)$. Conversely given any element of $H^{1/2}(\partial\Omega)$, there is a function in $H^1(\Omega)$ which takes those boundary values.

3.4.5 Poincaré's inequality

There are many variants of the Poincaré inequality. The common theme is that after removing nonzero constants one can estimate the L^p-norm of a function in terms of the L^p-norm of its derivative.

We denote the L^p-norm of the derivative of u by

$$\|Du\|_{L^p} = \left(\sum_{i=1}^{d} |D_i u|^p\right)^{1/p}.$$

Here,
$$D_i u = \frac{\partial u}{\partial x_i}.$$

If k is a positive integer, we similarly define the L^p norm of the kth order derivatives by
$$\|D^k u\|_{L^p} = \left(\sum_{|\alpha|=k} |D^\alpha u|^p\right)^{1/p}.$$

Theorem 3.9 *Suppose that Ω is a bounded domain (or, more generally, that Ω is bounded in one direction). Then there exists a constant C such that*
$$\|u\|_{L^p} \leq C\|Du\|_{L^p}$$
for all $u \in W_0^{1,p}(\Omega)$.

Note that the Poincaré estimate is false for nonzero constants, so the assumption that $u \in W_0^{1,p}$, rather than $u \in W^{1,p}$, is essential.

One useful consequence of this estimate is that $\|Du\|_{L^p}$ provides an equivalent norm on $W_0^{1,p}$. In a similar way, $\|D^k u\|_{L^p}$ provides a norm on $W_0^{k,p}$. When $p = 2$, it also follows that we can use

$$(u, v) = \int_\Omega Du(x) \cdot Dv(x)\, dx$$

as an inner product on $H_0^1(\Omega)$.

Here is another Poincaré-type inequality.

Theorem 3.10 *Suppose that Ω is a smooth connected bounded domain. There exists a constant C such that*

$$\|u - \bar{u}\|_{L^p} \leq C\|Du\|_{L^p}$$

for all $u \in W^{1,p}(\Omega)$. Here $\bar{u}$ is the mean of u over Ω,

$$\bar{u} = \frac{1}{|\Omega|}\int_\Omega u(x)\, dx.$$

The same results hold for periodic functions defined on the d-dimensional torus $\mathbf{T}^d$.

3.4.6 Gagliardo-Nirenberg inequalities

Special cases of the following **Gagliardo-Nirenberg** interpolation inequalities are frequently useful in nonlinear problems. They allow us to estimate L^r norms of intermediate derivatives in terms of L^p norms of higher derivatives and an L^q norm of the function itself.

Theorem 3.11 *Suppose that k is a positive integer and $1 \leq p, q \leq \infty$. Let j be any integer such that $0 \leq j < k$ and d the number of space dimensions. If $(k - j - d/p)$ is a nonnegative integer let*

$$\frac{j}{k} \leq \theta < 1;$$

otherwise, let

$$\frac{j}{k} \leq \theta \leq 1.$$

Define r by

$$\frac{1}{r} = \frac{j}{d} + \theta\left(\frac{1}{p} - \frac{k}{d}\right) + (1-\theta)\frac{1}{q}.$$

(a) If $D^k u \in L^p(\mathbf{R}^d)$ and $u \in L^q(\mathbf{R}^d)$ then $D^j u \in L^r(\mathbf{R}^d)$ and there exists a constant $C = C(k, p, q, d, \theta)$ such that

$$\|D^j u\|_{L^r} \leq C\|D^k u\|_{L^p}^\theta \|u\|_{L^q}^{1-\theta}.$$

(b) If Ω is a smooth bounded domain and $u \in W^{k,p}(\Omega) \cap L^q(\Omega)$ then $u \in W^{j,r}(\Omega)$ and there exists a constant $C = C(k, p, q, d, \theta, \Omega)$ such that

$$\|u\|_{W^{j,r}} \leq C\|u\|_{W^{k,p}}^\theta \|u\|_{L^q}^{1-\theta}.$$

The expression for r can be derived by a simple scaling argument applied to the function $u_\lambda(x) = u(\lambda x)$ just as in the case of the Sobolev imbedding theorem.

Note that the limiting case $\theta = 1$ gives the Sobolev imbedding theorem. For instance, if $j = 0$ and $k = 1$ we get that $r = p^*$ and the Gagliardo-Nirenberg inequality reduces to

$$\|u\|_{L^{p^*}} \leq C\|Du\|_{L^p}.$$

One useful consequence of these estimates is that $W^{k,p}$ forms an algebra for $kp > d$. This is the range in which $W^{k,p}$-functions are continuous. A function space is an algebra if the product of functions in the space also belongs to the space. Note that the L^p-spaces are not algebras when $p < \infty$. For example if $u, v \in L^2$, then all we can say about the product is that $uv \in L^1$.

Theorem 3.12 *If $kp > d$ then $W^{k,p}$ is an algebra. There is a constant C such that*

$$\|uv\|_{W^{k,p}} \leq C\|u\|_{W^{k,p}}\|v\|_{W^{k,p}}.$$

Many useful inequalities can be obtained as special cases of the above Gagliardo-Nirenberg inequality. Here are a few examples.

Examples.

1. If $u \in L^p \cap W^{2,r}$ then $u \in W^{1,q}$ where q is the harmonic mean of p and r,
$$\frac{1}{q} = \frac{1}{2}\left(\frac{1}{p} + \frac{1}{r}\right).$$
Moreover,
$$\|Du\|_{L^q} \leq C\|u\|_{W^{2,r}}^{1/2}\|u\|_{L^p}^{1/2}.$$

2. Taking $p = \infty$, which gives $q = 2r$, we get
$$\|Du\|_{L^{2r}} \leq C\|u\|_{W^{2,r}}^{1/2}\|u\|_{L^\infty}^{1/2}.$$
It follows from this inequality that $X = L^\infty \cap W^{2,r}$ is an algebra.

3. Taking $p = q = r$, we get that

$$\|Du\|_{L^p} \leq C\|u\|_{W^{2,p}}^{1/2}\|u\|_{L^p}^{1/2}.$$

In particular, this inequality implies that for any $\varepsilon > 0$ there exists a constant C_ε such that

$$\|Du\|_{L^p} \leq \varepsilon\|D^2u\|_{L^p} + C_\varepsilon\|u\|_{L^p}.$$

So the L^p-norm of Du is controlled by the L^p-norms of u and D^2u.

4. Let $1 \leq q \leq p < \infty$. Then

$$\|u\|_{L^p} \leq C\|u\|_{L^q}^{1-\theta}\|u\|_{W^{1,d}}^{\theta},$$

where

$$\theta = 1 - \frac{q}{p}.$$

3.5 Poisson's equation

Let us consider the Dirichlet problem for the Laplacian,

$$\begin{aligned} -\Delta u &= f \qquad x \in \Omega, \\ u(x) &= 0 \qquad x \in \partial\Omega. \end{aligned} \tag{3.2}$$

Here, $f : \Omega \to \mathbf{R}$ is a given function (or distribution) and we assume that Ω is a smooth bounded open set in $\mathbf{R}^d$.

To formulate any PDE problem in a precise way, we have to specify what function space solutions should belong to. We also have to specify how the derivatives are defined and in what sense the solution satisfies the boundary conditions and any other side conditions. There is often a great deal of choice in how this is done.

A **classical solution** of (3.2) is a function $u \in C^2(\bar{\Omega})$ which statisfies the PDE and the boundary condition pointwise. Classical solutions can be constructed directly by means of potential theory or by maximum principle estimates.

In many ways it is simpler to analyze (3.2) by means of variational methods in which we look for a weak solution which is not required to be continuously differentiable. In this case, the derivatives are understood in a distributional sense. To motivate the definition of weak solutions, suppose that u is a smooth solution. Let $\varphi \in C_c^\infty(\Omega)$ be any test function. Then multiplication of (3.2) by φ and an integration by parts imply that

$$\int_\Omega \nabla u(x) \cdot \nabla \varphi(x)\, dx = \int_\Omega f(x)\varphi(x)\, dx. \tag{3.3}$$

Conversely, if u is a smooth function which vanishes on $\partial\Omega$ and which satisfies (3.3) for all test functions φ, then u is a classical solution of the original boundary value problem.

Let us suppose that the solution u and the test functions φ belong to the same space. Then ∇u and $\nabla \varphi$ must both be square integrable so we must look for solutions in the space $H_0^1(\Omega)$. Since $\varphi \in H_0^1(\Omega)$, we can make sense of the right hand side provided that $f \in H^{-1}(\Omega)$. This leads to the following definition.

Definition 3.7 *Given any $f \in H^{-1}(\Omega)$ we say that u is a* **weak solution** *of (3.2) if $u \in H_0^1(\Omega)$ and if*

$$\int_\Omega \nabla u \cdot \nabla \varphi\, dx = \langle f, \varphi \rangle$$

for all $\varphi \in H_0^1(\Omega)$.

Remarks.

1. Define a functional $J : H_0^1(\Omega) \to \mathbf{R}$ by

   $$J(u) = \frac{1}{2} \int_\Omega |\nabla u|^2\, dx - \langle f, u \rangle.$$

   Any function which minimizes J is a weak solution of (3.2). This fact explains the connection between the present approach and variational methods.

2. Note that the boundary condition $u = 0$ on $\partial\Omega$ is replaced by the condition that $u \in H_0^1(\Omega)$. The precise sense in which weak solutions satisfy boundary conditions or initial conditions often requires careful attention.

3. This definition is not the most general definition of weak solutions. For example, we could consider distributional solutions of (3.2) in which $f \notin H^{-1}(\Omega)$. The definition given is the "natural" one for the existence proof below.

The existence and uniqueness of weak solutions follows trivially from the Poincaré inequality and the Riesz representation theorem.

Theorem 3.13 *There is a unique weak solution* $u \in H_0^1(\Omega)$ *of (3.2) for every* $f \in H^{-1}(\Omega)$*. There exists a constant* $C = C(d, \Omega)$ *such that*

$$\|u\|_{H^1} \leq C \|f\|_{H^{-1}}.$$

Proof. By the Poincaré inequality, we can use

$$(u, v) = \int_\Omega \nabla u \cdot \nabla v \, dx$$

as an inner product on $H_0^1(\Omega)$. Since $f \in H^{-1}(\Omega) = H_0^1(\Omega)'$, and $H_0^1(\Omega)$ is a Hilbert space, the Riesz representation theorem implies that there exists a unique $u \in H_0^1(\Omega)$ such that

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

for all $\varphi \in H_0^1(\Omega)$. This function u is the unique weak solution of (3.2). Moreover, we have $\|u\|_{H_0^1} = \|f\|_{H^{-1}}$. QED

Remarks.

1. This argument can be generalized to other strongly elliptic linear operators which needn't be self-adjoint (the Lax-Milgram theorem). Our main interest is in nonlinear problems, so we will not consider these generalizations here.

2. It follows from the theorem that

   $$-\Delta : H_0^1(\Omega) \to H^{-1}(\Omega)$$

   is a Hilbert space isomorphism — in fact it is just the isomorphism between $H_0^1(\Omega)$ and its dual space in which the dual space is represented concretely as a space of distributions.

3. The Riesz representation theorem also shows that there is a unique solution $u \in H^1(\mathbf{R}^d)$ of the PDE

   $$-\Delta u + u = f(x), \qquad x \in \mathbf{R}^d$$

for any $f \in H^1(\mathbf{R}^d)$ and therefore

$$(-\Delta + I) : H^1(\mathbf{R}^d) \to H^{-1}(\mathbf{R}^d)$$

is an isomorphism. The lower order term proportional to u is essential here because the Poincaré inequality does not hold in $\mathbf{R}^d$. Classically, one would impose a decay condition at infinity, such as $u(x) \to 0$ as $x \to \infty$, in order to obtain a unique solution. This condition is replaced by the integrability condition $u \in H^1(\mathbf{R}^d)$ in the case of weak solutions.

4. The proof gives a solution u of (3.2) which belongs to H^1. This is the best regularity one can hope for in the case of general right hand side $f \in H^{-1}$. However, if $f \in H^k$ is smooth, elliptic regularity theory shows that the solution $u \in H^{k+2}$ and that

$$\|u\|_{H^{k+2}} \leq C\|f\|_{H^k}.$$

This gain of derivatives is typical of elliptic equations. The main point is that one can estimate the L^2-norm of *all* second derivatives of $u \in H_0^1$ in terms of the L^2-norm of the single combination of second derivatives Δu. In particular, if $f \in H^k$ with $k > d/2$ then it follows from the Sobolev imbedding theorem that $u \in H^{k+2} \subset C^2$ is a classical solution.

5. Rather suprisingly, if $f \in C(\bar{\Omega})$ it is *not* always possible to find a solution of (3.2) with $u \in C^2(\bar{\Omega})$ or even with $u \in W^{2,\infty}(\bar{\Omega})$. The use of Hölder continuous functions gives a much better results. If $0 < \beta \leq 1$ and k is a non-negative integer, then there is a unique Hölder continuous function $u \in C^{k+2,\beta}(\bar{\Omega})$ for every Hölder continuous right hand side $f \in C^{k,\beta}(\bar{\Omega})$. One can also estimate the Hölder norms of the second derivatives of u in terms of the Hölder norm of f — these estimates are called Schauder estimates.

6. Analogous existence, uniqueness, and regularity results are true for $1 < p < \infty$. If Ω is a smooth bounded set and $f \in W^{k,p}(\Omega)$, then there is a unique solution of (3.2) with $u \in W^{k+2,p}(\Omega) \cap W_0^{1,p}(\Omega)$. Furthermore,

$$\|u\|_{W^{k+2,p}} \leq C\|f\|_{W^{k,p}}.$$

The proof of this result in the case $p \neq 2$ (Agmon-Douglis-Nirenberg, 1959) is much more difficult than in the case $p = 2$, and it requires delicate estimates for singular integral operators which are due to Calderón and Zygmund.

7. This last regularity result fails in the cases $p = 1$ and $p = \infty$. Thus, $C(\bar{\Omega})$, $L^1(\Omega)$, and $L^\infty(\Omega)$ are "bad" spaces for the study of elliptic PDE's and are best avoided if at all possible.

3.6 Weak convergence

Suppose we have constructed a sequence of approximate solutions of a PDE. To construct an exact solution we need to show two things: (a) the sequence, or at least some subsequence, has a limit; (b) the limit is a solution. Part (a) follows if we know that the sequence is contained in a compact set. Part (b) follows if we can pass to the limit in all the terms appearing in the PDE. This passage to the limit depends on the continuity of the terms in the PDE with respect to the convergence established in (a). The two most useful types of convergence are strong convergence and weak convergence. These types of convergence have compensating advantages and disadvantages: for (a), strong compactness is harder to establish than weak compactness; but for (b), strong continuity is easier to establish than weak continuity.

The basic theorem about weak compactness is that norm-bounded sets are weakly precompact. This usually makes it easy to prove weak compactness. Linear functions are weakly continuous if and only if they are strongly continuous. Thus, weak convergence methods are well-suited to the study of linear problems. Strongly continuous nonlinear functions are never weakly continuous. Thus, in nonlinear problems one either has to establish a suitable strong compactness result, or one has to use the specific structure of the nonlinearity to show that the nonlinear terms converge. For example, a convex function is weakly lower semi-continuous if and only if it is strongly lower semi-continuous. Many nonlinear convex minimization problems in the calculus of variations can be treated by weak convergence methods using this result. A sizable part of recent research in the calculus of variations has been concerned with weakening the requirement of convexity. For instance, Ball (1977) proved the existence of static solutions of the nonlinear elasticity equations. These solutions minimize the elastic energy functional. The main difficulty in this problem is that the energy is not convex, although it does possess a weaker "polyconvexity" property which Ball showed was sufficient to obtain existence.

In this section we summarize some basic definitions and facts about weak convergence. First, we define the dual space of a Banach space X. To be definite, we assume that X is a real Banach space; complex Banach spaces require the introduction of complex conjugates in the appropriate places.

Definition 3.8 *Let X be a Banach space. The* **dual space** X' *consists of continuous linear functionals* $\omega : X \to \mathbf{R}$*. We denote the action of* $\omega \in X'$ *on* $u \in X$ *by* $\langle \omega, u \rangle$*. The dual space is a Banach space with norm*

$$\|\omega\| = \sup_{\substack{u \in X \\ u \neq 0}} \frac{\langle \omega, u \rangle}{\|u\|}$$

A Banach space is said to be **reflexive** if $X'' = X$, meaning that every continuous linear functional $F : X' \to \mathbf{R}$ is of the form

$$\langle F, \omega \rangle = \langle \omega, u \rangle$$

for some $u \in X$.

If $X = H$ is a Hilbert space, any continuous linear functional $\omega : H \to \mathbf{R}$ has the form

$$\langle \omega, u \rangle = (w, u)$$

for some $w \in H$, so we can identify H' with H.

Examples.

1. For $1 \leq p < \infty$, we can identify $(L^p)'$ with $L^{p'}$ where p' is the Hölder conjugate of p. In that case, for $f \in L^{p'}$ and $g \in L^p$ we have

   $$\langle f, g \rangle = \int f(x) g(x)\, dx.$$

   In particular, $(L^2)' = L^2$.

2. Although L^∞ is the dual space of L^1, it is not true that L^1 is the dual space of L^∞. Thus L^1 and L^∞ are not reflexive.

We now define weak convergence.

Definition 3.9 *Let X be a Banach space. We say that a sequence* $\{u_n\} \subset X$ **converges weakly** *to* $u \in X$ *if*

$$\langle \omega, u_n \rangle \to \langle \omega, u \rangle \qquad \text{as } n \to \infty$$

for all $\omega \in X'$*. We denote weak convergence by a "harpoon,"*

$$u_n \rightharpoonup u.$$

Weak convergence can be thought of as componentwise convergence. The sequence u_n converges weakly iff each sequence of components $\langle \omega, u_n \rangle$ converges.

We have $u_n \rightharpoonup u$ in a Hilbert space H iff

$$(u_n, v) \to (u, v) \qquad \text{for all } v \in H.$$

We use a full arrow, $u_n \to u$, to denotes strong convergence, meaning that

$$\|u_n - u\| \to 0 \qquad \text{as } n \to \infty.$$

On a finite dimensional space, weak and strong convergence are equivalent. On an infinite dimensional space, strong convergence implies weak convergence but not conversely. The following examples illustrate the two basic ways in which a weakly convergent sequences fails to be strongly convergent, namely oscillations and concentrations.

Examples.

1. The sequence

   $$u_n(x) = \sin nx$$

   is weakly convergent to 0 in $L^2(0, 2\pi)$ but it is not strongly convergent. To prove the weak convergence we first observe that if $\varphi \in C_c^\infty(0, 2\pi)$ is a test function, then an integration by parts shows that

   $$(\sin nx, \varphi) = \frac{1}{n}(\cos nx, \varphi') \to 0 \qquad \text{as } n \to \infty.$$

   The result for general $\varphi \in L^2$ follows from the density of test functions.

   The failure of strong convergence follows immediately from the fact that the u_n are orthogonal functions with norm π.

   The problem here is that the functions u_n oscillate more and more rapidly as $n \to \infty$. Thus averages of the functions converge to zero although the functions themselves do not converge to zero.

   Note that although

   $$u_n \rightharpoonup 0,$$

   we have

   $$u_n^2 \rightharpoonup \frac{1}{2}.$$

   This result shows that the nonlinear function u^2 is not weakly continuous.

2. The sequence

$$u_n(x) = \begin{cases} n & \text{if } 0 < x < 1/n^2 \\ 0 & \text{otherwise} \end{cases}$$

is weakly convergent to 0 in $L^2(\mathbf{R})$ but it is not strongly convergent. The problem here is that the functions u_n concentrate more and more strongly at the origin as $n \to \infty$.

Although the sequence u_n^2 does not converge weakly in L^2, it does converge in an even weaker distributional sense,

$$u_n^2 \rightharpoonup \delta,$$

meaning that

$$\langle u_n, \varphi \rangle \to \langle \delta, \varphi \rangle$$

for all test functions $\varphi \in C_c^\infty(\mathbf{R})$. Again, this result shows that the nonlinear function u^2 is not weakly continuous.

If X' is the dual of a Banach space, we can define another type of weak convegence called weak-$*$ convergence.

Definition 3.10 *Let X' be the dual of a Banach space X. We say $\omega \in X'$ is the* **weak-$*$** *limit of a sequence $\{\omega_n\} \subset X'$ if*

$$\langle \omega_n, u \rangle \to \langle \omega, u \rangle \qquad \text{as } n \to \infty$$

for all $u \in X$. We denote weak-$*$ convergence by

$$\omega_n \overset{*}{\rightharpoonup} \omega.$$

By contrast, weak convergence in X' means that

$$\langle F, \omega_n \rangle \to \langle F, \omega \rangle \qquad \text{for all } F \in X''.$$

If X is reflexive, then weak convergence and weak-$*$ convergence in X' are equivalent and we can ignore the distinction. If X' is the dual space of a nonreflexive space X, then it is nearly always better to use weak-$*$ convergence in X' instead of weak convergence. The most commonly occuring nonreflexive spaces are L^∞, L^1, and the associated Sobolev spaces $W^{k,\infty}$, $W^{k,1}$.

Example.

1. We have $(L^1)' = L^\infty$. Hence

$$u_n \rightharpoonup u \qquad \text{in } L^1$$

iff

$$\int u_n(x)v(x)\,dx \to \int u(x)v(x)\,dx \qquad \text{for all } v \in L^\infty.$$

2. Since $L^\infty = (L^1)'$, we have

$$u_n \overset{*}{\rightharpoonup} u \qquad \text{in } L^\infty$$

iff

$$\int u_n(x)v(x)\,dx \to \int u(x)v(x)\,dx \qquad \text{for all } v \in L^1.$$

Here is the basic compactness result for weak-$*$ convergence, sometimes called Alaoglu's theorem.

Theorem 3.14 *Let X' be the dual space of a Banach space X. A bounded set in X' is weak-$*$ precompact.*

If X is reflexive then it follows that bounded sets in X are weakly precompact. In particular, suppose that $\{u_n\}$ is a sequence in a reflexive Banach space X which is uniformly bounded with respect to n,

$$\|u_n\| \leq C.$$

Then there exists a weakly convergent subsequence, $\{u_{n_k}\}$, such that

$$u_{n_k} \rightharpoonup u \qquad \text{as } k \to \infty.$$

Examples.

1. Suppose that $u_n \in L^2(\Omega)$ is a sequence of functions such that

$$\int_\Omega |u_n(x)|^2\,dx \leq 1.$$

   Then Alaoglu's theorem implies that there is a subsequence which converges weakly in $L^2(\Omega)$.

2. Suppose that $u_n \in H_0^1(\Omega)$ is a sequence of functions such that

$$\int_\Omega |\nabla u_n(x)|^2 \, dx \leq 1.$$

Then, if Ω is a bounded domain, the Rellich-Kondrachov theorem implies that there is a subsequence which converges strongly in $L^2(\Omega)$.

Thus, if we only have bounds on the functions, we get a weakly convergent subsequence, but if we have bounds on their derivatives, we get a strongly convergent subsequence.

Chapter 4

The Diffusion Equation

Suppose that $u(x,t)$ is a solution of a partial differential evolution equation. such as the diffusion equation. We want to regard $u(t) = u(\cdot, t)$ as a function of time which takes values in a Banach space X of functions of x. In the case of a system of ODE's, the vector space $X = \mathbf{R}^m$ is finite dimensional, and we can regard a vector-valued function as a finite collection of scalar-valued functions. In the case of PDE's, the vector space X is infinite dimensional and we need to proceed more abstractly. Even in the finite-dimensional case this more geometric approach helps to clarify ideas.

We begin with a brief description of vector-valued functions. Then we formulate the definition of weak solutions of the diffusion equation and use the Galerkin method to prove the existence and uniqueness of weak solutions.

4.1 Time-dependent Sobolev spaces

Let X be a Banach space, and $(a,b) \subset \mathbf{R}$ a time interval. For $1 \leq p < \infty$, we define the L^p-space of functions $u : (a,b) \to X$ by

$$L^p(a,b;X) = \left\{ u : u \text{ is measurable and } \textstyle\int_a^b \|u(t)\|^p \, dt < \infty \right\},$$

$$\|u\|_{L^p(a,b;X)} = \left(\int_a^b \|u(t)\|^p \, dt \right)^{1/p}.$$

For $p = \infty$ we define

$$L^\infty(a,b;X) = \left\{ u : u \text{ is measurable and } \|u(t)\| \text{ is essentially bounded} \right\},$$

$$\|u\|_{L^\infty(a,b;X)} = \sup_{t \in (a,b)} \|u(t)\|.$$

These spaces are Banach spaces.

For a compact interval $[a, b] \subset \mathbf{R}$ we define the space of strongly continuous functions

$$\begin{aligned} C([a,b];X) &= \{u : u \text{ is strongly continuous on } [a,b]\}, \\ \|u\|_{C([a,b];X)} &= \sup_{t \in [a,b]} \|u(t)\|. \end{aligned}$$

Here, strong continuity means that

$$\|u(t+h) - u(t)\| \to 0 \qquad \text{as } h \to 0.$$

One can also define the space $C_w([a, b]; X)$ of weakly continuous functions $u : [a, b] \to X$. Weak continuity means that

$$\langle \omega, u(t+h) \rangle \to \langle \omega, u(t) \rangle \qquad \text{as } h \to 0$$

for all $\omega \in X'$. If X is finite dimensional, then there is no distinction between weak and strong continuity. If X is infinite dimensional, weakly continuous functions need not be strongly continuous. For example, the function

$$u : \mathbf{R} \to L^2(0, 1)$$

defined by

$$u(x,t) = \begin{cases} \sin(x/t) & \text{if } t > 0 \\ 0 & \text{if } t \leq 0 \end{cases}$$

is weakly continuous at $t = 0$ but it is not strongly continuous at $t = 0$.

Integrals and distributional time derivatives of vector-valued functions $u : (a, b) \to X$ can be defined in essentially the same way as for scalar valued functions $u : (a, b) \to \mathbf{R}$. The basic definitions are described briefly below.

We define the space $C^1([a, b]; X)$ of continuously differentiable functions $u : [a, b] \to X$ with the norm

$$\|u\|_{C^1([a,b];X)} = \sup_{t \in [a,b]} \|u(t)\| + \sup_{t \in [a,b]} \|\dot{u}(t)\|.$$

We also define the Sobolev spaces $W^{1,p}(a, b; X)$ consisting of functions in $L^p(a, b; X)$ whose distributional time derivative belongs to $L^p(a, b; X)$. The associated norm is

$$\|u\|_{W^{1,p}(a,b;X)} = \left(\int_a^b \left[\|u(t)\|^p + \|\dot{u}(t)\|^p \right] \, dt \right)^{1/p}.$$

Examples.

1. Suppose $X = L^2(\Omega)$. Then

$$\begin{aligned}
\|u\|_{L^2(a,b;X)} &= \left(\int_a^b \int_\Omega |u(x,t)|^2 \, dxdt \right)^{1/2}, \\
\|u\|_{L^\infty(a,b;X)} &= \sup_{t \in (a,b)} \left(\int_\Omega |u(x,t)|^2 \, dx \right)^{1/2}.
\end{aligned}$$

2. Suppose $X = H_0^1(\Omega)$. Then

$$\|u\|_{L^2(a,b;X)} = \left(\int_a^b \int_\Omega |Du(x,t)|^2 \, dxdt \right)^{1/2}.$$

   If $X = H^{-1}(\Omega)$, then

$$\|u\|_{L^2(a,b;X)} = \left(\int_a^b \sup_{\substack{v \in H_0^1 \\ v \neq 0}} \frac{\langle u, v \rangle^2}{\|v\|^2} \, dt \right)^{1/2}$$

3. The space $H^1(a, b; L^2(\Omega))$ has the norm

$$\|u\|_{H^1(a,b;L^2)} = \left(\int_a^b \int_\Omega \left(|u(x,t)|^2 + |\dot{u}(x,t)|^2 \right) \, dxdt \right)^{1/2}.$$

4. Be careful not to confuse regularity in t with regularity in x. For example, any $f \in L^p(\mathbf{R}^d)$, with $p < \infty$, is continuous in the L^p-sense, meaning that

$$f(x+h) \to f(x) \qquad \text{in } L^p \text{ as } h \to 0.$$

   It follows that the function $u : \mathbf{R} \to L^p(\mathbf{R}^d)$ defined by $u(x,t) = f(x+t)$ is a continuous function of t, that is $u \in C(\mathbf{R}; L^p)$. However, if $f(x)$ is not a continuous function of x, then $u(x,t)$ is not a continuous function of x for any value of t.

4.1.1 Vector-valued integrals

In this section, we sketch the basic ideas of integration of vector-valued functions. When the target space X is infinite dimensional it has many different topologies (for example, the strong and weak topologies) and the integral of vector-valued functions can be defined in different, non-equivalent, ways. The simplest and most useful definition is the Bochner integral. This definition is a straightforward generalization of the usual Lebesgue integral of real valued functions.

A function $u : (a, b) \to X$ is said to be **countably-valued** if it assumes at most countably many values in X and if each value is assumed on a Lebesgue measurable set in (a, b). We say that u is **measurable** if it is the pointwise almost everywhere limit of countably valued functions. All limits in X are with respect to the strong topology (i.e. norm convergence) unless stated otherwise. We say that u is **integrable** if it is measurable and if

$$\int_a^b \|u(t)\| \, dt < \infty.$$

Note that if $u : (a, b) \to X$ is measurable, then it follows that $\|u\| : (a, b) \to \mathbf{R}$ is measurable, so this condition makes sense.

We define the Banach space

$$\begin{aligned} L^1(a, b; X) &= \{u : (a, b) \to X | u \text{ is integrable}\}, \\ \|u\|_{L^1(a,b;X)} &= \int_a^b \|u(t)\| \, dt. \end{aligned}$$

The space $L^1_{\text{loc}}(a, b; X)$ of locally integrable functions consists of all measurable functions such that

$$\int_K \|u(t)\| \, dt < \infty.$$

for every compact set $K \subset (a, b)$.

Suppose that u is an integrable, countably-valued function. We can write u in the form

$$\begin{aligned} u(t) &= \sum_{i=1}^{\infty} c_i \chi_{A_i}(t), \\ \|u\|_{L^1} &= \sum_{i=1}^{\infty} \|c_i\| |A_i| < \infty \end{aligned}$$

where $c_i \in X$, $A_i \subset (a, b)$ is the set of times t where $u(t) = c_i$, χ_{A_i} is the indicator function of A_i, and $|A_i|$ is the one-dimensional Lebesgue measure of A_i.

We then define
$$\int_a^b u(t)\,dt = \sum_{i=1}^{\infty} c_i |A_i| \in X.$$

For general $u \in L^1(a,b;X)$ it can be shown that there exists a sequence $\{u_n\}$ of countably-valued functions which converges to u in $L^1(a,b;X)$. The proof is of this fact is almost identical to the proof in the case of real-valued functions. We then define

$$\int_a^b u(t)\,dt = \lim_{n\to\infty} \int_a^b u_n(t)\,dt \in X.$$

This limit exists and is independent of the sequence of countably-valued functions which is used to approximate u.

Most of the usual properties of scalar-valued integrals remain true. For example:

1. It follows immediately from the definition that

$$\left\| \int_a^b u(t)\,dt \right\| \le \int_a^b \|u(t)\|\,dt.$$

2. **Lebesgue dominated convergence theorem.** Suppose that $u_n \in L^1(a,b;X)$ and $u_n(t) \to u(t)$ as $n \to \infty$ pointwise a.e. in t. Suppose that there exists a function $F : (a,b) \to \mathbf{R}$ such that

$$\int_a^b F(t)\,dt < \infty$$

and for all n,
$$\|u_n(t)\| \le F(t).$$

Then $u \in L^1(a,b;X)$ and

$$\int_a^b u(t)\,dt = \lim_{n\to\infty} \int_a^b u_n(t)\,dt.$$

3. If $A : X \to Y$ is a continuous linear map between Banach spaces, then

$$A \int_a^b u(t)\,dt = \int_a^b Au(t)\,dt.$$

4.1.2 Vector-valued distributions

To give a precise mathematical formulation of an evolution equation

$$\dot{u} = f(u)$$

where $u(t) \in X$, we need to define the time derivative of a vector-valued function u. If $u \in C^1([a, b]; X)$ is a continuously differentiable function of time, then the strong derivative $\dot{u}$ of u can be defined pointwise by

$$\lim_{h \to 0} \left\| \frac{u(t+h) - u(t)}{h} - \dot{u}(t) \right\| = 0.$$

In many cases we cannot assume that u is continuously differentiable, and the time derivative must be interpreted as a distributional derivative. The definition of vector-valued distributions is very similar to the definition of real-valued distributions.

Let

$$C_c^\infty(a, b) = \{\text{test functions } \varphi : (a, b) \to \mathbf{R}\}.$$

Let X be a Banach space. An X-valued distribution T is a continuous linear map

$$T : C_c^\infty(a, b) \to X.$$

We denote the value of T acting on a test function φ by $\langle T, \varphi \rangle \in X$. This element of X can be thought of as the time average of T weighted by the test function φ.

We define the distributional derivative $\dot{T}$ of a distribution T by

$$\langle \dot{T}, \varphi \rangle = -\langle T, \dot{\varphi} \rangle.$$

Any function $u \in L^1_{\text{loc}}(a, b; X)$ defines a regular distribution by

$$\langle u, \varphi \rangle = \int_a^b u(t)\varphi(t)\, dt.$$

Theorem 4.1 *Let $u, v : (a, b) \to X$ be locally integrable functions. Any of the following three conditions are equivalent to the condition that $\dot{u} = v$ in the sense of distributions.*

(a) $$\int_a^b u(t)\dot{\varphi}(t)\, dt = -\int_a^b v(t)\varphi(t)\, dt$$
for all $\varphi \in C_c^\infty(a, b)$;

(b) $u(t) = u_0 + \int_{t_0}^{t} v(s)\,ds$

a.e. in $t \in (a,b)$ *for some* $t_0 \in (a,b)$, $u_0 \in X$;

(c) $\frac{d}{dt}\langle \omega, u(t)\rangle = \langle \omega, v(t)\rangle$

in the real-valued distributional sense for all $\omega \in X'$.

4.2 The diffusion equation

The basic example of a linear parabolic equation is the diffusion equation. Let us consider the following initial boundary value problem for $u(x,t)$,

$$\begin{aligned} \dot{u} &= \Delta u + f(x,t), && x \in \Omega,\ t > 0, \\ u &= 0, && x \in \Omega, \\ u &= u_0(x) && t = 0. \end{aligned} \tag{4.1}$$

Here, as usual, $\dot{}$ denotes the time derivative and we assume that Ω is a smooth bounded open set in $\mathbf{R}^d$.

4.2.1 A priori estimates

As we saw in Chapter 2, the "natural" spaces for the study of the Poisson equation $-\Delta u = f$ by variational methods are $u \in H_0^1$ and $f \in H^{-1}$. This suggests that we seek solutions $u(t)$ of the diffusion equation (4.1) in which

$$\begin{aligned} u &: (0,T) \to H_0^1(\Omega), \\ \dot{u} &: (0,T) \to H^{-1}(\Omega). \end{aligned}$$

We still need to specify the regularity of u and $\dot{u}$ with respect to t. In order to discover the appropriate function space setting for the diffusion equation, we derive a priori energy estimates which are satisfied by smooth solutions. We will then show that it is possible to construct solutions which satisfy these estimates.

Multiplication of the heat equation by u implies that

$$\left(\frac{1}{2}u^2\right)_t = \nabla \cdot (u \nabla u) - \nabla u \cdot \nabla u + fu.$$

Integration of this equation over Ω, and use of the boundary condition $u = 0$ gives

$$\frac{1}{2}\frac{d}{dt}\|u\|_{L^2}^2 + \|u\|_{H_0^1}^2 = \int_\Omega fu\,dx.$$

Here,

$$\|u\|_{H_0^1} = \|\nabla u\|_{L^2}.$$

Integration with respect to t over the time interval $(0, T)$ then implies that

$$\frac{1}{2}\|u(T)\|_{L^2}^2 + \int_0^T \|u\|_{H_0^1}^2(t)\,dt = \frac{1}{2}\|u(0)\|_{L^2}^2 + \int_0^T \int_\Omega f u\,dxdt. \tag{4.2}$$

From the definition of the H^{-1}-norm and the Cauchy-Schwartz inequality we have that

$$\begin{aligned}
\left|\int_0^T \int_\Omega f u\,dxdt\right| &\leq \int_0^T \|f\|_{H^{-1}}\|u\|_{H_0^1}\,dt \\
&\leq \left(\int_0^T \|f\|_{H^{-1}}^2\,dt\right)^{1/2} \left(\int_0^T \|u\|_{H_0^1}^2\,dt\right)^{1/2} \\
&\leq \|f\|_{L^2(0,T;H^{-1})}\|u\|_{L^2(0,T;H_0^1)}.
\end{aligned}$$

Thus, neglecting the term $\|u(T)\|_{L^2}^2$ on the left-hand side of (4.2), we see that

$$\|u\|_{L^2(0,T;H_0^1)}^2 \leq \frac{1}{2}\|u(0)\|_{L^2}^2 + \|f\|_{L^2(0,T;H^{-1})}\|u\|_{L^2(0,T;H_0^1)}. \tag{4.3}$$

By Young's inequality,

$$ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2,$$

we have

$$\|f\|_{L^2(0,T;H^{-1})}\|u\|_{L^2(0,T;H_0^1)} \leq \frac{1}{2}\|f\|_{L^2(0,T;H^{-1})}^2 + \frac{1}{2}\|u\|_{L^2(0,T;H_0^1)}^2.$$

Use of this inequality in (4.3) gives

$$\|u\|_{L^2(0,T;H_0^1)}^2 \leq \|f\|_{L^2(0,T;H^{-1})}^2 + \|u(0)\|_{L^2}^2.$$

Thus, the $L^2(0,T;H_0^1)$-norm of the solution is uniformly bounded by the L^2-norm of the initial data and the $L^2(0,T;H^{-1})$-norm of the forcing term. The equation $\dot{u} = \Delta u + f$ then implies that

$$\|\dot{u}\|_{L^2(0,T;H^{-1})} \leq C\left(\|f\|_{L^2(0,T;H^{-1})} + \|u(0)\|_{L^2}\right).$$

These results suggest that the “natural” solution spaces are

$$\begin{aligned} u &\in L^2(0,T;H_0^1), \\ \dot{u} &\in L^2(0,T;H^{-1}). \end{aligned}$$

These spaces have the additional advantage of being Hilbert spaces.

Neglecting the term $\int_0^T \|u\|_{H_0^1}^2 \, dt$ on the left-hand side of (4.2), we also conclude that

$$\frac{1}{2}\|u(T)\|_{L^2}^2 \le \frac{1}{2}\|u(0)\|_{L^2}^2 + \|f\|_{L^2(0,T;H^{-1})}\|u\|_{L^2(0,T;H_0^1)}.$$

Thus, since T is arbitrary, we see that

$$\|u\|_{L^\infty(0,T;L^2)}^2 \le \|f\|_{L^2(0,T;H^{-1})}^2 + \|u(0)\|_{L^2}^2.$$

The boundary condition $u = 0$ on $\partial\Omega$ is formulated in a weak sense by requiring that $u(t) \in H_0^1(\Omega)$. It is not immediately clear that the initial condition makes sense because we cannot assign a boundary value $u(0)$ to a function $u \in L^2(0,T;X)$. However, the following proposition shows that we can make sense of the initial condition for any $u_0 \in L^2(\Omega)$.

Proposition 4.1 *If*

$$u \in L^2(0,T;H_0^1(\Omega)) \quad \textit{and} \quad \dot{u} \in L^2(0,T;H^{-1}(\Omega))$$

then

$$u \in C([0,T];L^2(\Omega)).$$

Proof. Suppose that $v \in C_c^\infty(\mathbf{R};H_0^1(\Omega))$ is an H_0^1-valued test function. Then

$$\frac{d}{dt}\|v(t)\|_{L^2}^2 = 2\langle \dot{v}, v\rangle,$$

and

$$\|v(t)\|_{L^2}^2 = 2\int_{-\infty}^t \langle \dot{v}(s), v(s)\rangle \, ds.$$

Using the Cauchy-Schwartz inequality to estimate the right hand side of this equation and taking the sup of the result with respect to t, we obtain that

$$\|v(t)\|_{L^\infty(\mathbf{R};L^2)}^2 \le 2\|\dot{v}\|_{L^2(\mathbf{R};H^{-1})}\|v\|_{L^2(\mathbf{R};H_0^1)}. \tag{4.4}$$

Any $u \in L^2(0,T;H_0^1(\Omega))$ can be approximated by a a sequence of test functions $u_n \in C_c^\infty(\mathbf{R};H_0^1(\Omega))$ such that

$$\begin{aligned} u_n &\to u && \text{in } L^2(0,T;H_0^1(\Omega)), \\ \dot{u}_n &\to \dot{u} && \text{in } L^2(0,T;H^{-1}(\Omega)). \end{aligned}$$

To prove this assertion, we extend u by zero outside the interval $a \le t \le b$ and mollify the extended function with a smooth compactly supported kernel. We omit the details

Applying (4.4) to $v = u_n - u_m$, we see that $\{u_n\}$ is a Cauchy sequence in $L^\infty(\mathbf{R};L^2)$ so it has a limit $\bar{u}$. Since $u_n \in C(\mathbf{R};L^2)$ the limit $\bar{u} \in C(\mathbf{R};L^2)$ is continuous. The restriction of $\bar{u}$ to $[0,T]$ is u, so u is continuous with values in $L^2(\Omega)$. QED

4.2.2 Weak solutions

The above estimates suggest the following definition of a weak solution of the diffusion equation.

Definition 4.1 *Given $f \in L^2(0,T;H^{-1}(\Omega))$ and $u_0 \in L^2(\Omega)$, we say that u is a* **weak solution** *of (4.1) if*

$$\begin{aligned} & u \in L^2\left(0,T;H_0^1(\Omega)\right) \quad \textit{and} \quad \dot{u} \in L^2\left(0,T;H^{-1}(\Omega)\right); \\ & \dot{u} = \Delta u + f \qquad \textit{in the sense of } H^{-1}(\Omega)\textit{-valued distributions}; \\ & u(0) = u_0 \quad \textit{in } C\left([0,T];L^2(\Omega)\right). \end{aligned}$$

The condition that $u(t)$ is an $H^{-1}(\Omega)$-distributional solution is equivalent to the condition that

$$\langle \dot{u}(t), v\rangle + (\nabla u(t), \nabla v) = \langle f(t), v\rangle \tag{4.5}$$

for all $v \in H_0^1(\Omega)$, with time derivatives interpreted in the sense of real-valued distributional derivatives. In (4.5), $(\cdot,\cdot)$ denotes the L^2-inner product and $\langle\cdot,\cdot\rangle$ denotes the H^{-1}-H_0^1 duality pairing. For $u \in L^2 \subset H^{-1}$ and $v \in H_0^1 \subset L^2$, we have

$$\langle u, v\rangle = \int_\Omega u(x)v(x)\,dx = (u,v).$$

The condition that u is a distributional solution can also be written using time dependent test functions. An equivalent formulation is that

$$\int_0^T \left[\langle \dot{u}(t), v(t)\rangle + (\nabla u(t), \nabla v(t))\right]\,dt = \int_0^T \langle f(t), v(t)\rangle\,dt \tag{4.6}$$

for all test functions $v \in L^2(0,T;H_0^1(\Omega))$.

4.3 Existence of weak solutions

Our proof of the existence of weak solutions of the diffusion equation makes use of the Galerkin method in which the PDE is approximated by a system of ODE's. The Galerkin method is also very useful in the analysis of nonlinear PDE's. We will prove that the Galerkin approximations converge weakly to a solution of the diffusion equation.

Theorem 4.2 *There is a unique weak solution of (4.1) for every*

$$u_0 \in L^2(\Omega), \qquad f \in L^2\left(0,T;H^{-1}(\Omega)\right).$$

Proof. The uniqueness of weak solutions follows immediately from energy estimates. Suppose that u_1 and u_2 are two solutions. Let

$$w = u_1 - u_2.$$

Then w satisfies (4.6) with $f = 0$. Setting $u = v = w$ in (4.6) we get

$$\frac{d}{dt}\|w(t)\|_{L^2}^2 = -2\left(\nabla w(t), \nabla w(t)\right) \le 0.$$

We also have the initial condition

$$\|w(0)\|^2 = 0.$$

Gronwall's inequality implies that $\|w(t)\|^2 = 0$ so that $u_1 = u_2$.

We split the existence proof into three steps.

1. Construction of finite dimensional approximations u^n.
2. A priori estimates for the approximations.
3. Convergence of the approximations u^n to a solution u as $n \to \infty$.

Step 1. Finite dimensional approximations. We choose functions $v_k(x)$, $k = 1, 2, \ldots$ such that $v_k \in H_0^1(\Omega)$ and

$$\{v_k : k = 1, 2, \ldots\} \quad \text{is an orthonormal basis of } L^2(\Omega).$$

For example, the eigenfunctions of the Laplacian

$$\begin{aligned} -\Delta v_k &= \lambda_k v_k, \\ v_k(x) &= 0 \qquad x \in \partial\Omega, \end{aligned}$$

form such a basis.

Let V^n be the finite dimensional subspace of L^2 (and H_0^1) which is spanned by $\{v_k : k = 1, 2, \ldots, n\}$. We introduce the orthogonal projection $P^n : L^2(\Omega) \to V^n$. We look for an approximate solution $u^n \in V^n$. This approximation is required to satisfy the equation which is obtained by projection of the diffusion equation (4.5) onto V^n, namely

$$\begin{aligned} \dot{u}^n &= P^n \left(\Delta u^n + f\right), \\ u^n(0) &= P^n u_0. \end{aligned}$$

In terms of components we have

$$\begin{aligned} u^n(t) &= \sum_{k=1}^{n} c_k^n(t) v_k, \\ c_k^n(t) &= \left(u^n(t), v_k\right). \end{aligned}$$

The equation for u^n is equivalent to the system

$$\begin{aligned} &\left(\dot{u}^n, v_k\right) = -\left(\nabla u^n, \nabla v_k\right) + \langle f, v_k\rangle \qquad k = 1, \ldots, n, \\ &u^n(0) = u_0^n. \end{aligned} \tag{4.7}$$

Here, u_0^n is the projection of u_0 onto V^n,

$$u_0^n = \sum_{k=1}^{n} \left(u_0, v_k\right) v_k.$$

The corresponding system of ODE's for the coefficients $c_k^n(t)$is

$$\begin{aligned} \dot{c}_k^n(t) &= -\sum_{\ell=1}^{n} A_{k\ell} c_\ell^n(t) + f_k(t), \qquad k = 1, \ldots, n, \\ c_k^n(0) &= c_{0k}. \end{aligned}$$

Here

$$\begin{aligned} A_{k\ell} &= \left(\nabla \varphi_k, \nabla \varphi_\ell\right), \\ f_k(t) &= \langle f(t), \varphi_k\rangle, \\ c_{0k} &= \left(u_0, \varphi_k\right). \end{aligned}$$

This system of ODE's for

$$\mathbf{c} = (c_1, \ldots, c_n)^t$$

is an $n \times n$ linear system of the form

$$\begin{aligned}\dot{\mathbf{c}} &= A\mathbf{c} + \mathbf{f},\\ \mathbf{c}(0) &= \mathbf{c}_0.\end{aligned}$$

It therefore has a unique solution $\mathbf{c} : \mathbf{R} \to \mathbf{R}^n$. Thus, for any $T > 0$, we have constructed an approximate solution

$$u^n \in C\left([0, T]; V^n\right)$$

with

$$\dot{u}^n \in L^2\left(0, T; V^n\right).$$

Step 2. A priori estimates. In order to pass to the limit $n \to \infty$, we need estimates on the approximations u^n which are uniform in n. These estimates are completely analogous to the a priori estimates for the diffusion equation.

We multiply (4.7) by c_k^n and sum the result over $1 \le k \le n$. We get that

$$\frac{1}{2}\frac{d}{dt}\|u^n\|_{L^2}^2 + \|\nabla u^n\|_{L^2}^2 = \langle f, u^n\rangle.$$

Integration of this equation with respect to t over the time interval $(0, T)$ implies that

$$\frac{1}{2}\|u^n(T)\|_{L^2}^2 + \int_0^T \|u^n\|_{H_0^1}^2(t)\, dt = \frac{1}{2}\|u_0^n\|_{L^2}^2 + \int_0^T \int_\Omega f u \, dxdt.$$

Neglecting the term $\|u^n(T)\|_{L^2}^2$ on the left-hand side and applying Young's inequality as before we conclude that

$$\|u^n\|_{L^2(0,T;H_0^1)}^2 \le \|f\|_{L^2(0,T;H^{-1})}^2 + \|u_0\|_{L^2}^2. \tag{4.8}$$

Here, we have used the fact that

$$\|u_0^n\|_{L^2} \le \|u_0\|_{L^2}.$$

Thus, the $L^2(0, T; H_0^1)$-norms of the approximate solutions are uniformly bounded. The equation $\dot{u}^n = P^n\left(\Delta u + f\right)$ then implies that

$$\|\dot{u}\|_{L^2(0,T;H^{-1})} \le C\left(\|f\|_{L^2(0,T;H^{-1})} + \|u(0)\|_{L^2}\right). \tag{4.9}$$

Step 3. Convergence. From (4.8) and (4.9), the sequences $\{u^n\}$ and $\{\dot{u}^n\}$ are uniformly bounded. We can therefore extract a weakly convergent subsequence, which we still denote by $\{u^n\}$, such that as $n \to \infty$

$$\begin{aligned} u^n &\rightharpoonup u && \text{in } L^2(0,T;H_0^1), \\ \dot{u}^n &\rightharpoonup \dot{u} && \text{in } L^2(0,T;H^{-1}). \end{aligned}$$

We want to prove that the limit u is a weak solution of the diffusion equation.

Consider any test function with values in V^m for some $m \in \mathbf{N}$,

$$v \in C^\infty\left([0,T];V^m\right).$$

If $n \geq m$ then integration of the Galerkin equations for u^n implies that

$$\int_0^T \left\{\langle \dot{u}^n(t), v(t)\rangle + (\nabla u^n(t), \nabla v(t))\right\}\, dt = \int_0^T \langle f(t), v(t)\rangle\, dt.$$

We can take the weak limit as $n \to \infty$ in this equation to conclude that

$$\int_0^T \left\{\langle \dot{u}(t), v(t)\rangle + (\nabla u(t), \nabla v(t))\right\}\, dt = \int_0^T \langle f(t), v(t)\rangle\, dt.$$

Since V^m-valued test functions are dense in $L^2(0,T;H_0^1)$ this equation holds for all $v \in L^2(0,T;H_0^1)$. Hence, u is a weak solution of the diffusion equation.

A small additional argument is required to show that $u(0) = u_0$ in $C([0,T];L^2(\Omega))$. We omit the details. QED

Bibliography

[1] L. C. Evans, *Partial Differential Equations*, Berkeley Mathematics Lecture Notes, Vol. 3A–3B, Berkeley, 1993.

[2] H. Fujita, On the blowing up of solutions of the Cauchy problem for $u_t = \Delta u + u^{1+\alpha}$, *J. Fac. Sci. Tokyo, Sect. IA, Math.*, 13 (1966), 109-124.

[3] J. Goodman, Stability of the Kuramoto-Sivashinsky and related systems, *Commun. Pure Appl. Phys.*, *Comm. Pure Appl. Math.*, 47 (1994), 293-306.

[4] J. M. Hyman, and B. Nicolaenko, The Kuramoto-Sivashinsky eqation: a bridge between pde's and dynamical systems, *Physica D*, 18 (1986), 113-126.

[5] F. John, *Nonlinear Wave Equations, Formation of Singularities*, University Lecture Series, Vol. 2, Amer. Math. Soc., Providence, 1990.

[6] T. Kato, On the Cauchy problem for the (generalized) Korteweg-de-Vries equation, *Adv. in Math. Suppl. Studies, Studies in Appl. Math.*, 8 (1983), 93-128.

[7] S. Klainerman, Mathematical theory of classical fields and general relativity, in *Mathematical Physics X*, ed. K. Schmüdgen, Springer-Verlag, New York (1992), 213-236.

[8] H.-O. Kreiss and J. Lorentz, *Initial-Boundary Value Problems and the Navier-Stokes Equations*, Academic Press, San Diego, 1989.

[9] P. D. Lax, *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 11, SIAM, Philadelphia, 1973.

[10] A. J. Majda, *Compressible Fluid Flow and Conservation Laws in Several Space Dimensions*, Appl. Math. Sci. Vol. 53, Springer-Verlag, New York, 1984.

[11] B. Nicolaenko, B. Scheurer, and R. Temam, Some global dynamical properties of the Kuramoto-Sivashinsky equations: Nonlinear stability and attractors, *Physica D*, 16 (1985), 155-183.

[12] J. Rauch, *Partial Differential Equations*, Springer-Verlag, New York, 1991.

[13] M. Renardy, and R. C. Rogers, *An Introduction to Partial Differential Equations*, Texts in Applied Mathematics, Vol. 13, Springer-Verlag, New York, 1993.

[14] J.-C. Saut, Generalized Kadomtsev-Petviashvili equations, *Indiana Univ. Math. J.*, 42 (1993), 1011-1026.

[15] J. Smoller, *Shock Waves and Reaction-Diffusion Equations*, Springer-Verlag, New York, 1983.

[16] W. Strauss, *Nonlinear Wave Equations*, Conference Board of the Mathematical Sciences, AMS, Philadelphia, 1989.

[17] L. Tartar, Compensated compactness and applications to PDEs, *Nonlinear Analysis and Mechanics*, Herriot-Watt symposium, Pitman (1979).