Applied Functional Analysis and Partial Differential Equations

Milan Miklavčič

World Scientific

Applied Functional Analysis and Partial Differential Equations

Applied Functional Analysis and Partial Differential Equations

Milan Miklavčič
Michigan State University


World Scientific
Singapore • New Jersey • London • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.

P O Box 128, Farrer Road, Singapore 912805

USA office: Suite 1B, 1060 Main Street, River Edge, NJ 07661

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

First published 1998
Reprinted 2001

APPLIED FUNCTIONAL ANALYSIS AND PARTIAL DIFFERENTIAL EQUATIONS

Copyright © 1998 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN 981-02-3535-6

Printed in Singapore by Uto-Print

Contents

Preface

1 Linear Operators in Banach Spaces **1**
1.1 Metric Spaces 1
1.2 Vector Spaces 6
1.3 Banach Spaces 8
1.4 Linear Operators 11
1.5 Duals 16
1.6 Spectrum 23
1.7 Compact Linear Operators 29
1.8 Boundary Value Problems for Linear ODEs 39
1.9 Exercises 43

2 Linear Operators in Hilbert Spaces **47**
2.1 Orthonormal Sets 47
2.2 Adjoints 54
2.3 Accretive Operators 58
2.4 Weak Solutions 62
2.5 Example: Constant Coefficient PDEs 68
2.6 Self-adjoint Operators 70
2.7 Example: Sturm-Liouville Problem 76
2.8 Sectorial Forms 80
2.9 Example: Harmonic Oscillator and Hermite Functions 88
2.10 Example: Completeness of Bessel Functions 92
2.11 Example: Finite Element Method 96
2.12 Friedrichs Extension 99
2.13 Exercises 104

3 Sobolev Spaces **109**
3.1 Introduction 109
3.2 Fourier Transform 114
3.3 Distributions 122

3.4 Weak Derivatives . . . 126
3.5 Definition and Basic Properties of Sobolev Spaces . . . 135
3.6 Imbeddings of $W^{m,p}(\Omega)$ . . . 140
3.7 Elliptic Problems . . . 153
3.8 Regularity of Weak Solutions . . . 160
3.9 Exercises . . . 163

4 Semigroups of Linear Operators **167**
4.1 Introduction . . . 167
4.2 Bochner Integral . . . 171
4.3 Basic Properties of Semigroups . . . 176
4.4 Example: Wave Equation . . . 187
4.5 Sectorial Operators and Analytic Semigroups . . . 192
4.6 Invariant Subspaces . . . 204
4.7 The Inhomogeneous Problem - Part I . . . 209
4.8 Exercises . . . 212

5 Weakly Nonlinear Evolution Equations **215**
5.1 Introduction . . . 215
5.2 Basic Theory . . . 218
5.3 Example: Nonlinear Heat Equation . . . 222
5.4 Approximation for Evolution Equations . . . 224
5.5 Example: Finite Difference Method . . . 233
5.6 Example: Galerkin Method for Parabolic Equations . . . 236
5.7 Example: Galerkin Method for the Wave Equation . . . 239
5.8 Friedrichs Extension and Galerkin Approximations . . . 243
5.9 Exercises . . . 245

6 Semilinear Parabolic Equations **247**
6.1 Fractional Powers of Operators . . . 247
6.2 The Inhomogeneous Problem - Part II . . . 254
6.3 Global Version . . . 257
6.4 Main Results . . . 259
6.5 Example: Navier-Stokes Equations . . . 267
6.6 Example: A Stability Problem . . . 271
6.7 Example: A Classical Solution . . . 273
6.8 Dynamical Systems . . . 276
6.9 Example: The Chafee-Infante Problem . . . 279
6.10 Exercises . . . 282

Bibliography **285**

List of Symbols **289**

Index **291**

Preface

This book is an introduction to partial differential equations (PDEs) and the relevant functional analysis tools which PDEs require. This material is intended for second year graduate students of mathematics and is based on a course taught at Michigan State University for a number of years. The purpose of the course, and of the book, is to give students a rapid and solid research-oriented foundation in areas of PDEs, like semilinear parabolic equations, that include studies of the stability of fluid flows and, more generally, of the dynamics generated by dissipative systems, numerical PDEs, elliptic and hyperbolic PDEs, and quantum mechanics. In other words, the book gives a complete introduction to and also covers significant portions of the material presented in such classics as *Partial Differential Equations* by Avner Friedman, *Geometric Theory of Semilinear Parabolic Equations* by Dan Henry, and *Semigroups of Linear Operators and Applications to PDEs* by A. Pazy.

The need for such a book is due to the fact that in order to study PDEs one needs to know some functional analysis, which requires a thorough knowledge of real analysis (Lebesgue integral). Therefore, if real analysis is studied in the first year of graduate school, and functional analysis in the second year, the student only begins with PDEs in the third year - and may even have to re-learn functional analysis if the prior instructor ignored unbounded operators (which sometimes happens).

The reader is expected to be comfortable with the Lebesgue integral; more specifically, with the material presented in Examples 1.3.4 and 1.5.2. The Cauchy Theorem is also used in a couple of places, with the most difficult version used in (4.44). These are the only real prerequisites for the whole book. Above this level, all theorems used are proved in the text. One may, and perhaps should, skip over some of the proofs. However, they are included in case they are needed.

With regard to the writing style, all formal statements, like Theorems, contain all assumptions except for those declared at the beginning of the section in which the statement appears. This should make it easy, even for a casual reader, to figure out what is actually assumed in a given statement. There is, however, one exception. Throughout Chapter 3 it is assumed, unless otherwise specified, that Ω is an arbitrary nonempty open set in $\mathbb{R}^n$, $n \in \mathbb{N}$.

In the first two chapters functional analysis tools are developed and differential operators are studied as examples. Sturm-Liouville operators are nice examples of self-

adjoint operators with compact resolvent and are reused in Chapter 4 as generators of strongly continuous semigroups. Hörmander's treatment of weak solutions of constant coefficient PDEs is also presented early on as an example. The foundation of elliptic, parabolic and wave equations, as well as of Galerkin approximations, is given in the section on Sectorial Forms. Throughout the text, completeness of a number of orthonormal systems is proven.

The Fourier transform and its applications to constant coefficient PDEs are presented in Chapter 3. We briefly touch upon distributions and fundamental solutions, and prove the Malgrange-Ehrenpreis Theorem. Most of Chapter 3 is devoted to study of Sobolev spaces. Many sharp results concerning existence and compactness of imbeddings, as well as interpolation inequalities, are proven. These results are applied to elliptic problems in the last two sections.

The study of evolution equations begins in Chapter 4 where the semigroup theory is introduced. The Hille-Yosida Theorem for strongly continuous semigroups and Hille's construction of analytic semigroups are presented. The semigroup theory and the results of the previous chapters enable us to discuss linear parabolic and wave equations. In preparation for studies of nonlinear evolution equations, the invariant subspaces associated with the semigroups and the inhomogeneous problem are also examined.

A dynamical systems approach to weakly nonlinear evolution equations is given in Chapter 5 with a nonlinear heat equation studied as an example. Trotter's approximation theory is adapted to such equations giving convergence of Galerkin and finite difference type approximations.

The chapter on semilinear parabolic equations begins with a very technical section on fractional powers of operators. Our main results contain existence, uniqueness, continuous dependence, maximal interval of existence, stability and instability results. These results are applied to the Navier-Stokes equations, to a stability problem in fluid mechanics, to showing how a classical solution can be obtained, and to the Chafee-Infante problem as an example of a gradient system.

I wish to thank S. N. Chow and D. R. Dunninger for their early encouragement and my wife Pam for checking the grammar.

Chapter 1

Linear Operators in Banach Spaces

1.1 Metric Spaces

A **metric space** is a set M in which a **distance** (or **metric**) d is defined, with the following properties:

(i) $0 \leq d(x,y) < \infty$ for all $x, y \in M$

(ii) $d(x,y) = 0$ if and only if $x = y$

(iii) $d(x,y) = d(y,x)$ for all $x, y \in M$

(iv) $d(x,y) \leq d(x,z) + d(z,y)$ for all $x, y, z \in M$ (triangle inequality).

The best known example of a metric space is $\mathbb{R}^n$, $n \in \mathbb{N}$, with the distance between two points $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ given by

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

For the remainder of this section, let M be a metric space with metric d. We shall review basic concepts associated with metric spaces mainly in order to standardize notation and terminology. It is assumed that the reader is somewhat familiar with metric spaces. Hence, our review will be brief.

When $A \subset M$, $B \subset M$ are not empty, define

$$\operatorname{dist}(A,B) = \inf\{d(x,y) \mid x \in A, y \in B\}.$$

Analogously, $\operatorname{dist}(x,A) = \operatorname{dist}(\{x\},A)$ for $x \in M$.

The **open ball** with center at $x \in M$ and radius r will be denoted by

$$B(x,r) = \{y \in M \mid d(x,y) < r\}.$$

A subset O of the metric space M is said to be **open** if for each $x \in O$ there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subset O$. Note that the empty set $\emptyset$ is open and that the intersection of any finite number of open sets is open.

$C \subset M$ is said to be **closed** if its **complement** $C^c \equiv \{x \in M | x \notin C\}$ is open.

The **closure** of a set A will be denoted by $\overline{A}$ and is defined as the intersection of all closed sets containing A. Note that $\overline{A}$ is a closed set and if A is closed, then $A = \overline{A}$.

A set A is said to be **dense** in M if $\overline{A} = M$, i.e. $A \cap B \neq \emptyset$ for every nonempty open set B. A metric space is **separable** if it contains a countable dense set.

A set $K \subset M$ is called **compact** if from any collection of open sets, whose union contains K, a finite number of sets can be chosen so that their union also contains K. Compact sets are closed in metric spaces. A set is said to be **relatively compact** if its closure is compact. Recall that a subset of $\mathbb{R}^n$, $n \in \mathbb{N}$, is compact if and only if it is closed and bounded.

A sequence $\{x_n\}_{n=1}^{\infty}$ in M is said to **converge** to $x \in M$ if $\lim_{n\to\infty} d(x_n, x) = 0$. Notations $x_n \to x$ or $\lim_{n\to\infty} x_n = x$ will be used in such a case. Note if $A \subset M$, then $x \in \overline{A}$ iff there exists a sequence of points of A converging to x. Observe also that a set $K \subset M$ is compact (relatively compact) iff for each sequence $\{y_n\}$ in K there exist $y \in K$ ($y \in M$, respectively) and integers $n_1 < n_2 < \cdots$ such that $\lim_{k\to\infty} y_{n_k} = y$.

A sequence of elements $\{x_n\}_{n=1}^{\infty}$ in M is said to be a **Cauchy sequence** if for each $\varepsilon > 0$ there exists an integer N such that $d(x_n, x_m) < \varepsilon$ whenever $n, m \geq N$. A metric space is said to be **complete** if every Cauchy sequence converges to some element in the space.

Theorem 1.1.1 (Baire) *If M is a complete metric space, the intersection of every countable collection of dense open subsets of M is dense in M.*

PROOF Suppose $V_1, V_2, V_3, \ldots$ are dense and open in M. Let W be any nonempty open set in M. It will be shown that $(\cap V_n) \cap W \neq \emptyset$.

Since V_1 is dense, $W \cap V_1$ is a nonempty open set. Hence we can find x_1 and r_1 such that

$$B(x_1, 2r_1) \subset W \cap V_1 \quad \text{and} \quad 0 < r_1 < 1. \tag{1.1}$$

If $n \geq 2$ and x_{n-1} and r_{n-1} are chosen, the denseness of V_n shows that $V_n \cap B(x_{n-1}, r_{n-1})$ is not empty, and we can therefore find x_n and r_n such that

$$B(x_n, 2r_n) \subset B(x_{n-1}, r_{n-1}) \cap V_n \quad \text{and} \quad 0 < r_n < \frac{1}{n}. \tag{1.2}$$

By induction, this process produces a sequence $\{x_n\}$ in M. If $i \geq n, j \geq n$ the construction shows that x_i and x_j both lie in $B(x_n, r_n)$. Hence $d(x_i, x_j) < 2r_n < 2/n$ and therefore $\{x_n\}$ is a Cauchy sequence. Since M is complete, $\lim_{n\to\infty} d(x_n, x) = 0$ for some $x \in M$. Now, for $m \geq n \geq 1$,

$$d(x_n, x) \leq d(x_n, x_m) + d(x_m, x) \leq r_n + d(x_m, x)$$

and hence $d(x_n, x) \leq r_n < 2r_n$, $x \in B(x_n, 2r_n)$. By (1.2), x belongs to each V_n, $n \geq 2$. Equation (1.1) implies $x \in W \cap V_1$. □

Corollary 1.1.2 *If M is a complete nonempty metric space and $M = \cup_{n=1}^{\infty} A_n$, then some $\overline{A_n}$ contains a nonempty open set.*

If $T : M \to M$ is such that for some $\varepsilon \in [0, 1)$ we have that $d(T(x), T(y)) \leq \varepsilon d(x, y)$ for all $x, y \in M$, then T is called a **contraction**. $x \in M$ is said to be a **fixed-point** of T if $T(x) = x$. The following Theorem implies that a contraction mapping on a complete, nonempty metric space always has a unique fixed-point. Try Exercise 2.

Theorem 1.1.3 (Contraction Mapping) *Suppose M is a complete, nonempty metric space with metric d and that $T : M \to M$ is such that for some $n \geq 1$ and some $\varepsilon < 1$ we have*

$$d(T^n(z), T^n(y)) \leq \varepsilon d(z, y) \quad \textit{for all} \quad z, y \in M.$$

Then there exists a unique $x \in M$ such that $T(x) = x$. Moreover,

$$d(x, T^m(y)) \leq \frac{\varepsilon^{[\frac{m}{n}]}}{1 - \varepsilon} \max_{0 \leq j \leq n-1} d(T^{n+j}(y), T^j(y)) \quad \textit{for all} \quad m \geq 0,\ y \in M.$$

PROOF Choose $y \in M$ and let $x_0 = y$, $x_{k+1} = T^n(x_k)$ for $k \geq 0$. By induction, $d(x_{k+1}, x_k) \leq \varepsilon^k d(x_1, x_0)$ for $k \geq 0$. Hence, for $m > k \geq 0$

$$\begin{aligned} d(x_m, x_k) &\leq d(x_m, x_{m-1}) + d(x_{m-1}, x_k) \\ &\leq d(x_m, x_{m-1}) + d(x_{m-1}, x_{m-2}) + \cdots + d(x_{k+1}, x_k) \\ &\leq \varepsilon^{m-1} d(x_1, x_0) + \cdots + \varepsilon^k d(x_1, x_0) \\ d(x_m, x_k) &\leq \varepsilon^k d(x_1, x_0)/(1 - \varepsilon). \end{aligned} \tag{1.3}$$

Thus $\{x_k\}$ forms a Cauchy sequence in M and therefore $\{x_k\}$ converges to some $x \in M$. Since $d(T^n(x), x_{k+1}) \leq \varepsilon d(x, x_k)$, we see that $\{x_k\}$ converges also to $T^n(x)$ and thus $T^n(x) = x$. If $T^n(z) = z$, then

$$d(x, z) = d(T^n(x), T^n(z)) \leq \varepsilon d(x, z).$$

Hence $x = z$. Since $T^n(T(x)) = T(x)$ we see that $T(x) = x$. If $T(w) = w$, then $T^n(w) = w$. Hence $w = x$. Taking $m \to \infty$ in (1.3) implies

$$d(x, T^{kn}(y)) \leq \varepsilon^k d(T^n(y), y)/(1 - \varepsilon).$$

So, if y is replaced by $T^j(y)$, $0 \leq j \leq n - 1$, the 'moreover' part follows. □

A function f from the metric space M into another metric space N with metric ρ is said to be **continuous at the point** $x \in M$ if for each $\varepsilon > 0$ there exists $\delta > 0$ such that $\rho(f(x), f(y)) < \varepsilon$ whenever $y \in M$ and $d(x, y) < \delta$. f is said to be **continuous** if it is continuous at each point. f is said to be **uniformly continuous** if for each $\varepsilon > 0$ there exists $\delta > 0$ such that $\rho(f(x), f(y)) < \varepsilon$ whenever $d(x, y) < \delta$. $C(M, N)$ will denote the collection of all continuous functions from M into N. Let $C_B(M, N)$ be the collection of those $f \in C(M, N)$ for which there exists $y \in N$ such that $\sup_{t \in M} \rho(f(t), y) < \infty$. Note that

$$d(f, g) \equiv \sup_{t \in M} \rho(f(t), g(t)) \quad \text{for} \quad f, g \in C_B(M, N)$$

defines a metric on $C_B(M, N)$.

Theorem 1.1.4 *Let M and N be nonempty metric spaces.*

(1) If M is compact, then $C_B(M, N) = C(M, N)$.

(2) If N is complete, then $C_B(M, N)$ is a complete metric space.

PROOF Choose $f \in C(M, N)$, $y \in N$ and let $A_k = \{t \in M \mid \rho(f(t), y) < k\}$. Note that $A_1, A_2, \ldots$ are open sets and that $M \subset \cup_{k=1}^{\infty} A_k$. So, if M is compact then $M \subset A_k$ for some k and, hence, $f \in C_B(M, N)$.

Assume that N is complete and that $\{f_n\}$ is a Cauchy sequence in $C_B(M, N)$. Thus, for each $\varepsilon > 0$ there exists k_ε such that

$$\rho(f_n(t), f_m(t)) < \varepsilon \quad \text{for all} \quad t \in M,\ n, m > k_\varepsilon. \tag{1.4}$$

Completeness of N implies that $\{f_n(t)\}$ converges to some $f(t)$ for every $t \in M$. If $\varepsilon > 0$ and $n > k_{\varepsilon/3}$, then (1.4) implies $\rho(f_n(t), f(t)) \leq \varepsilon/3$ for $t \in M$. Hence, $\rho(f(t), y) \leq \varepsilon/3 + \sup_{s \in M} \rho(f_n(s), y) < \infty$ for some $y \in N$, and since

$$\begin{aligned} \rho(f(t), f(s)) &\leq \rho(f(t), f_n(t)) + \rho(f_n(t), f_n(s)) + \rho(f_n(s), f(s)) \\ &\leq 2\varepsilon/3 + \rho(f_n(t), f_n(s)), \end{aligned}$$

we have that continuity of f_n implies continuity of f. Thus, $f \in C_B(M, N)$ and $d(f_n, f) \leq \varepsilon$ for $n > k_\varepsilon$. □

If N is another metric space, with metric ρ, and if $\nu \in (0, 1)$, then $C^\nu(M, N)$ denotes the set of all functions $f : M \to N$ for which there exists $\mathrm{c} \in (0, \infty)$ such that

$$\rho(f(x), f(y)) \leq \mathrm{c}\, d(x, y)^\nu \tag{1.5}$$

for all $x, y \in M$. Such a function f is called **Hölder continuous**. f is called **Lipschitz continuous** when (1.5) holds, with $\nu = 1$, for all $x, y \in M$.

$f : M \to N$ is said to be a **locally Hölder continuous** function if for each $z \in M$ there exist $r, c \in (0, \infty)$ and $\nu \in (0, 1)$ such that (1.5) holds for all $x, y \in B(z, r)$. The set of all such functions will be denoted by $C_H(M, N)$. Note that if $f \in C_H(M, N)$, then for each compact $K \subset M$ there exist $c \in (0, \infty)$ and $\nu \in (0, 1)$ such that (1.5) holds for all $x, y \in K$.

A set W is said to be **connected** if there do *not* exist two disjoint open sets A and B such that $W \subset A \cup B$ and both $W \cap A$ and $W \cap B$ are nonempty. If for each $x, y \in W$ there exists continuous $f : [0, 1] \to W$, such that $f(0) = x$ and $f(1) = y$, then W is said to be **arcwise-connected**. It can be easily shown that if W is arcwise-connected, then it is connected.

Theorem 1.1.5 (Arzela-Ascoli) *Suppose that M is a separable metric space, with metric d, and that $\{f_n\}_{n=1}^{\infty}$ is a sequence of complex valued functions on M such that*

(a) $\sup_n |f_n(x)| < \infty$ *for each* $x \in M$

(b) for each $\varepsilon > 0$, $x \in M$ *there exists* $\delta_{x\varepsilon} > 0$ *such that* $\sup_n |f_n(x) - f_n(y)| < \varepsilon$ *whenever* $y \in M$ *and* $d(x, y) < \delta_{x\varepsilon}$.

Then there exist integers $n_1 < n_2 < \cdots$ and a continuous complex valued function v on M such that

$$\lim_{k \to \infty} \sup_{x \in K} |f_{n_k}(x) - v(x)| = 0$$

for every compact subset K of M.

PROOF Let $S \equiv \{x_1, x_2, \ldots\}$ be a dense set in M. Let A_0 be the set of positive integers. For $m \geq 1$ let A_m be an infinite subset of A_{m-1} such that

$$\lim_{j \to \infty, j \in A_m} f_j(x_m) \quad \text{exists.}$$

Let $n_0 = 1$ and for $k \geq 1$ choose $n_k \in A_k$ such that $n_k > n_{k-1}$. Note that the

$$\lim_{i \to \infty} f_{n_i}(x) \quad \text{exists for all } x \in S.$$

Let K be a compact subset of M and choose $\varepsilon > 0$. Open balls $B(x, \delta_{x\varepsilon})$ with $x \in K$ cover K. Hence, a finite subcollection, say $B_1, \ldots, B_m$, also covers K. Choose $y_i \in S \cap B_i$ and let N be such that

$$|f_{n_i}(y_k) - f_{n_j}(y_k)| < \varepsilon \quad \text{for} \quad i, j \geq N, \ k = 1, \ldots, m.$$

Any $x \in K$ is contained in some $B_k = B(z_k, \delta_{z_k\varepsilon})$. Hence for $i, j > N$

$$\begin{aligned} |f_{n_i}(x) - f_{n_j}(x)| &\leq |f_{n_i}(x) - f_{n_i}(z_k)| + |f_{n_i}(z_k) - f_{n_i}(y_k)| \\ &\quad + |f_{n_i}(y_k) - f_{n_j}(y_k)| \\ &\quad + |f_{n_j}(y_k) - f_{n_j}(z_k)| + |f_{n_j}(z_k) - f_{n_j}(x)| \qquad (1.6) \\ &< 2\varepsilon + \varepsilon + 2\varepsilon = 5\varepsilon. \end{aligned}$$

This implies that $\{f_{n_i}(x)\}$ converges, say to $v(x)$, for each $x \in M$. Letting $j \to \infty$ in (1.6) gives that $\{f_{n_i}\}$ converges uniformly to v on every compact K.

To show that v is continuous, pick $x \in M$ and $\varepsilon > 0$. If $d(x,y) < \delta_{x\varepsilon}$, then

$$\begin{aligned} |v(x) - v(y)| &\leq |v(x) - f_{n_i}(x)| + |f_{n_i}(x) - f_{n_i}(y)| + |f_{n_i}(y) - v(y)| \\ &< |v(x) - f_{n_i}(x)| + \varepsilon + |f_{n_i}(y) - v(y)| \overset{i\to\infty}{\to} \varepsilon. \end{aligned}$$

□

1.2 Vector Spaces

The letter $\mathbb{K}$ will stand for either the field of real numbers $\mathbb{R}$ or the field of complex numbers $\mathbb{C}$. A **scalar** is a member of the **scalar field** $\mathbb{K}$. A **vector space over** $\mathbb{K}$ is a set X, whose elements are called vectors and in which two operations, **addition** and **scalar multiplication**, are defined as follows:

(a) For every pair of vectors x and y corresponds a vector $x + y$ in such a way that

$$x + y = y + x \qquad \text{and} \qquad x + (y + z) = (x + y) + z;$$

X contains a unique vector 0 (the **zero vector** or **origin** of X) such that $x + 0 = x$ for every $x \in X$; and for each $x \in X$ corresponds a unique vector $-x$ such that $x + (-x) = 0$.

(b) For every pair (α, x), with $\alpha \in \mathbb{K}$ and $x \in X$, corresponds a vector αx such that

$$1x = x, \qquad \alpha(\beta x) = (\alpha\beta)x,$$

and such that the two distributive laws

$$\alpha(x + y) = \alpha x + \alpha y, \qquad (\alpha + \beta)x = \alpha x + \beta x$$

hold.

The symbol 0 will, of course, also be used for the zero element of the scalar field.

A **real vector space** is one for which $\mathbb{K} = \mathbb{R}$; a **complex vector space** is one for which $\mathbb{K} = \mathbb{C}$. Any statement about vector spaces in which the scalar field is not explicitly mentioned is to be understood to apply to both of these cases.

When X is a vector space, $A \subset X, B \subset X, x \in X$ and $\alpha \in \mathbb{K}$, the following notation will be used:

$$x + A = \{x + y \,|\, y \in A\}$$

$$A + B = \{y + z \,|\, y \in A, z \in B\}$$

$$\alpha A = \{\alpha y \mid y \in A\}.$$

A nonempty subset Y of a vector space X is called a **subspace** of X if $\alpha x + \beta y \in Y$ for all $x, y \in Y$ and all scalars α, β. If $\alpha x + \beta y \in Y$ for all $x, y \in Y$ and for all real numbers α, β, then Y is called a **real subspace** of X.

A subset M of a vector space is said to be **convex** if

$$tx + (1-t)y \in M \quad \text{whenever} \quad t \in (0,1),\ x, y \in M.$$

Let $x_1, \ldots, x_n$ be elements of a vector space X. The set of all $\alpha_1 x_1 + \cdots + \alpha_n x_n$ with $\alpha_i \in \mathbb{K}$ is called the **span** of $x_1, \ldots, x_n$ and is denoted by $\operatorname{span}\{x_1, \ldots, x_n\}$. The elements $x_1, \ldots, x_n$ are said to be **linearly independent** if $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$ implies that $\alpha_i = 0$ for $1 \leq i \leq n$. A set M is said to be linearly independent if distinct elements of every finite subset of M are linearly independent.

$\dim X$ denotes the **dimension** of a vector space X and is either 0, a positive integer or ∞. If $X = \{0\}$, then $\dim X = 0$; if there exist $\{u_1, \ldots, u_n\}$ such that each $x \in X$ has a unique representation of the form

$$x = \alpha_1 u_1 + \cdots + \alpha_n u_n \qquad \text{with} \qquad \alpha_i \in \mathbb{K},$$

then $\dim X = n$ and $\{u_1, \ldots, u_n\}$ is a **basis** for X; in other cases $\dim X = \infty$.

EXAMPLE 1.2.1 If M is a nonempty metric space and X is a vector space, then $C(M, X)$ is also a vector space with the usual definitions of addition and scalar multiplication:

$$(f+g)(x) = f(x) + g(x) \quad \text{for} \quad f, g \in C(M, X),\ x \in M$$

$$(\alpha f)(x) = \alpha f(x) \quad \text{for} \quad \alpha \in \mathbb{K}, f \in C(M, X),\ x \in M.$$

$C(M, \mathbb{R})$, as well as $C(M, \mathbb{C})$, will frequently be denoted simply by $C(M)$.

The **support** of a scalar valued function f on the metric space M is the closure of the set $\{x \in M | f(x) \neq 0\}$ and is denoted by $\operatorname{supp}(f)$. Let $C_0(M)$ denotes the collection of all $f \in C(M)$ with compact $\operatorname{supp}(f)$.

EXAMPLE 1.2.2 Let Ω be a nonempty open set in $\mathbb{R}^n$, $n \geq 1$. For a scalar valued function f defined on Ω and a multi-index α (i.e. an ordered n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ of nonnegative integers α_i, $|\alpha| \equiv \alpha_1 + \cdots + \alpha_n$), define

$$D^\alpha f = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} f$$

when the indicated partial derivatives exist (in the indicated order) at each point in Ω; if $|\alpha| = 0$, then $D^\alpha f = f$. D_i will sometimes be used to denote $\frac{\partial}{\partial x_i}$.

For integers $m \geq 0$, let $C^m(\Omega)$ be the collection of all $f \in C(\Omega)$ such that $D^\alpha f \in C(\Omega)$ for every multi-index α with $|\alpha| \leq m$. Recall that if $f \in C^m(\Omega)$ and α is a multi-index with $|\alpha| \leq m$, then all partial derivatives of f, such that for each i the total number of differentiations with respect to x_i is equal to α_i, exist and are equal to $D^\alpha f$.

We write $f \in C^\infty(\Omega)$ iff $f \in C^m(\Omega)$ for all $m \geq 0$.

For $m \geq 0$, define $C_0^m(\Omega) = C_0(\Omega) \cap C^m(\Omega)$ and let $C_0^\infty(\Omega) = C_0(\Omega) \cap C^\infty(\Omega)$. $C_0^\infty(\Omega)$ is often denoted by $\mathcal{D}(\Omega)$ with its elements called test functions.

$C^m(\Omega)$, $C^\infty(\Omega)$, $C_0^m(\Omega)$, $C_0^\infty(\Omega)$ are subspaces of the vector space $C(\Omega)$.

When $\Omega = (a, b)$ we write $C_0^\infty(a, b)$ in place of $C_0^\infty((a, b))$ ...

For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$ we define $\alpha! = \alpha_1! \cdots \alpha_n!$. If β is also a multi-index, we write $\beta \leq \alpha$ provided that $\beta_i \leq \alpha_i$ for $1 \leq i \leq n$. For $\beta \leq \alpha$ define

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha - \beta)!} = \frac{\alpha_1! \cdots \alpha_n!}{\beta_1! \cdots \beta_n!(\alpha_1 - \beta_1)! \cdots (\alpha_n - \beta_n)!} = \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_n}{\beta_n}.$$

If $f, g \in C^{|\alpha|}(\Omega)$, then **derivative of the product** fg equals

$$D^\alpha(fg) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} (D^\beta f) D^{\alpha - \beta} g. \tag{1.7}$$

This can be easily verified, by induction, when differentiation with respect to only one variable is performed; the general case follows immediately from this special case.

Suppose $m \geq 1$, $f \in C^m(\Omega)$ and that $x + ty \in \Omega$ for $t \in [0, 1]$. For $t \in [0, 1]$, define $u(t) = f(x + ty)$. Induction on k gives

$$u^{(k)}(t) = \sum_{|\alpha| = k} \frac{k!}{\alpha!} y^\alpha (D^\alpha f)(x + ty) \quad \text{for} \quad t \in [0, 1],\ 0 \leq k \leq m$$

where $y^\alpha = y_1^{\alpha_1} \cdots y_n^{\alpha_n}$. The Taylor formula for u,

$$u(1) = \sum_{k=0}^{m-1} \frac{u^{(k)}(0)}{k!} + \int_0^1 (1 - t)^{m-1} \frac{u^{(m)}(t)}{(m-1)!} dt,$$

gives the **Taylor formula** for f:

$$f(x + y) = \sum_{|\alpha| < m} \frac{y^\alpha}{\alpha!} (D^\alpha f)(x) + m \sum_{|\alpha| = m} \frac{y^\alpha}{\alpha!} \int_0^1 (1 - t)^{m-1} (D^\alpha f)(x + ty) dt. \tag{1.8}$$

1.3 Banach Spaces

A vector space X is said to be a **normed space** if for every $x \in X$ there is associated a nonnegative real number $\|x\|$, called the **norm** of x, in such a way that

(a) $\|x + y\| \leq \|x\| + \|y\|$ for all x and y in X

(b) $\|\alpha x\| = |\alpha| \|x\|$ for all $x \in X$ and all scalars α

(c) $\|x\| > 0$ if $x \neq 0$.

A real-valued function $\|\cdot\|$, defined on a vector space X, is called a **seminorm** if it satisfies (a) and (b). Note if $\|\cdot\|$ is a seminorm on X, then

$$|\,\|x\| - \|y\|\,| \leq \|x - y\| \quad \text{for all} \quad x, y \in X$$

and hence if $\lim_{n\to\infty} \|x_n - x\| = 0$ then $\lim_{n\to\infty} \|x_n\| = \|x\|$.

A normed vector space will always be regarded as a metric space with the distance $d(x, y) = \|x - y\|$ and with the corresponding definitions of open sets, continuity, etc, introduced in Section 1.1. If a normed space is complete then it is called **Banach space**. More precisely, X is a real (complex) Banach space if X is a Banach space with real (complex) scalars.

A subset S of a normed space is said to be a **bounded set** if $S \subset B(0, r)$ for some $r \in (0, \infty)$.

Sometimes different norms are defined in the same vector space X. $\|\cdot\|$, $|\cdot|$ are said to be **equivalent norms** in X if there exist $c_1, c_2 \in (0, \infty)$ such that

$$\|x\| \leq c_1|x| \leq c_2\|x\| \quad \text{for all } x \in X.$$

Note that the property of a set to be open, closed or compact in a normed space is not affected if the norm is replaced by an equivalent norm.

EXAMPLE 1.3.1 $\mathbb{K}^n$, with the usual definitions of addition and scalar multiplication, is an n dimensional vector space. The usual euclidean norm,

$$\|x\| = (|x_1|^2 + \cdots + |x_n|^2)^{1/2}, \quad x = (x_1, \ldots, x_n),$$

makes it into a Banach space. This norm is often denoted by $|\cdot|$.

EXAMPLE 1.3.2 If X is a normed space and M is a nonempty metric space, then $C_B(M, X)$ is a collection of those $f \in C(M, X)$ which satisfy

$$\|f\|_\infty \equiv \sup_{t\in M} \|f(t)\| < \infty.$$

$C_B(M, X)$ is clearly a normed space with norm $\|\cdot\|_\infty$. If X is a Banach space, then Theorem 1.1.4 implies that $C_B(M, X)$ is a Banach space.

EXAMPLE 1.3.3 Let Ω be a nonempty open set in $\mathbb{R}^n$, $n \geq 1$, $m \geq 0$ and define $C_B^m(\Omega)$ to be the set of those $f \in C^m(\Omega)$ for which

$$\|f\|_{m,\infty} \equiv \max_{|\alpha|\leq m} \sup_{\Omega} |D^\alpha f| < \infty.$$

$C_B^m(\Omega)$ is clearly a normed space with the norm $\|\cdot\|_{m,\infty}$. Let us show that $C_B^m(\Omega)$ is complete and hence a Banach space. The case $m = 0$ is discussed above in Example 1.3.2. Assume that $C_B^{m-1}(\Omega)$ is complete for some $m \geq 1$ and that $\{f_k\}_{k=1}^\infty$ is a Cauchy sequence in $C_B^m(\Omega)$. Since $\{f_k\}_{k=1}^\infty$, $\{D_i f_k\}_{k=1}^\infty$ are Cauchy sequences in $C_B^{m-1}(\Omega)$, there

exist $f, g_i \in C_B^{m-1}(\Omega)$ such that $\|f_k - f\|_{m-1,\infty} \to 0$ and $\|D_i f_k - g_i\|_{m-1,\infty} \to 0$ as $k \to \infty$. Hence, if $x \in \Omega$ and e_i is the unit vector in $\mathbb{R}^n$ in the i-th direction, then

$$f_k(x + he_i) - f_k(x) = \int_0^h (D_i f_k)(x + se_i)ds$$

for small enough $h \in \mathbb{R}$. Letting $k \to \infty$ gives

$$f(x + he_i) - f(x) = \int_0^h g_i(x + se_i)ds.$$

Hence $D_i f = g_i$, which implies completeness.

$C_u^m(\Omega)$ denotes the set of those $f \in C_B^m(\Omega)$ for which $D^\alpha f$ is uniformly continuous for every multi-index α with $|\alpha| \leq m$. It can be easily seen that $C_u^m(\Omega)$ is a closed subspace of $C_B^m(\Omega)$ and hence also a Banach space with the norm $\|\cdot\|_{m,\infty}$. Note if $f \in C_u^m(\Omega)$, then each $D^\alpha f$, $|\alpha| \leq m$, has a unique uniformly continuous extension on $\overline{\Omega}$ which is defined for $x \in \overline{\Omega}\backslash\Omega$ by $(D^\alpha f)(x) = \lim_{n\to\infty}(D^\alpha f)(x_n)$ where $x_n \in \Omega$ are such that $x_n \to x$. $C_u^m(\Omega)$ is sometimes denoted by $C^m(\overline{\Omega})$ (note that $C^m(\overline{\mathbb{R}}) \neq C^m(\mathbb{R})$); when Ω is a bounded interval (a, b) it will be denoted by $C^m[a, b]$.

Let C_ℓ denote the collection of $f \in C(\mathbb{R}^n)$ such that $\lim_{|x|\to\infty} f(x)$ exists. The set of all $f \in C_\ell$, such that $\lim_{|x|\to\infty} f(x) = 0$, will be denoted by $C_{\ell 0}$. Observe that C_ℓ and $C_{\ell 0}$ are closed subspaces of $C_B(\mathbb{R}^n)$ and, hence, Banach spaces.

EXAMPLE 1.3.4 Let Ω be a nonempty (Lebesgue) measurable set in $\mathbb{R}^n$. For $p \in [1, \infty)$, we denote by $L^p(\Omega)$ the set of equivalence classes of measurable functions on Ω for which

$$\|f\|_p \equiv \left(\int_\Omega |f|^p\right)^{\frac{1}{p}} < \infty$$

(two functions belong to the same equivalence class, i.e. are equivalent, if they differ only on a set of measure 0). $L^\infty(\Omega)$ denotes the set of equivalence classes of measurable functions on Ω for which

$$\|f\|_\infty \equiv \text{ess-sup}_{x\in\Omega}|f(x)| < \infty.$$

$L^p(\Omega)$, $1 \leq p \leq \infty$, are Banach spaces with norm $\|\cdot\|_p$.

For convenience we also consider $L^p(\Omega)$ as a set of functions. With this convention in mind, we can assert that $C_0(\Omega) \subset L^p(\Omega)$. In fact, if $p \in [1, \infty)$, then $C_0(\Omega)$ is a dense subset of $L^p(\Omega)$. Using this basic fact let us prove the following statement. *If I is any nonempty open interval in $\mathbb{R}$, $m \geq 0$, then $C_0^m(I)$ is dense in $L^p(I)$, $p \in [1, \infty)$.* If true for some $m \geq 0$ and $f \in L^p(I)$ for some $p \in [1, \infty)$ and $\varepsilon > 0$, then $\|f - g\|_p < \varepsilon/2$ for some $g \in C_0^m(I)$. Define

$$h_\delta(x) = \frac{1}{\delta}\int_x^{x+\delta} g(t)dt$$

and note that $h_\delta \in C_0^{m+1}(I)$ for small δ. Since

$$h_\delta(x) - g(x) = \frac{1}{\delta}\int_x^{x+\delta} (g(t) - g(x))dt,$$

the uniform continuity of g implies $\|h_\delta - g\|_{0,\infty} \to 0$ as $\delta \to 0$, and since the support of h_δ lies in a fixed compact subset of I for small δ we can choose small $\delta > 0$ such that $\|h_\delta - g\|_p < \varepsilon/2$. Thus $\|f - h_\delta\|_p < \varepsilon$, showing that the statement is true for $m+1$ and hence for all m.

When $p, q, r \in [1, \infty]$, $p^{-1} + q^{-1} = r^{-1}$, Hölder's inequality implies that if $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$, then $fg \in L^r(\Omega)$ and

$$\|fg\|_r \leq \|f\|_p \|g\|_q. \tag{1.9}$$

This implies that if $r, p_1, ..., p_k \in [1, \infty]$, $r^{-1} = p_1^{-1} + \cdots + p_k^{-1}$ and $g_i \in L^{p_i}(\Omega)$, then $g_1 \cdots g_k \in L^r(\Omega)$ and

$$\|g_1 \cdots g_k\|_r \leq \|g_1\|_{p_1} \cdots \|g_k\|_{p_k}. \tag{1.10}$$

$L^p(\Omega)$ is also separable if $p \in [1, \infty)$. For $p \in [1, \infty]$, we write $g \in L^p_{loc}(\Omega)$ iff $g \in L^p(K)$ for each compact $K \subset \Omega$.

EXAMPLE 1.3.5 In the following definitions let $x = \{x_n\}_{n=1}^\infty$ be a sequence of scalars:

$$\ell^\infty = \{\, x \mid \|x\|_\infty \equiv \sup_n |x_n| < \infty \}$$

$$c_0 = \{\, x \mid \lim_{n\to\infty} x_n = 0 \}, \ \|x\| = \|x\|_\infty$$

$$\ell^p = \{\, x \mid \|x\|_p \equiv (\textstyle\sum_{n=1}^\infty |x_n|^p)^{1/p} < \infty \}, \quad 1 \leq p < \infty.$$

ℓ^∞, c_0, ℓ^p are nice, easy-to-play-with Banach spaces.

EXAMPLE 1.3.6 The Cartesian product $X \times Y$ is the set of all ordered pairs (x, y), where $x \in X$ and $y \in Y$. If X and Y are vector spaces with the same scalar field, then $X \times Y$ is a vector space under the following operations of addition and scalar multiplication:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$

$$\alpha(x, y) = (\alpha x, \alpha y).$$

If in addition, X and Y are normed spaces with respective norms $\|\cdot\|_X$, $\|\cdot\|_Y$, then $X \times Y$ is a normed space with the norm

$$\|(x, y)\| = \|x\|_X + \|y\|_Y.$$

Moreover, under this norm, $X \times Y$ is a Banach space provided X and Y are Banach spaces.

1.4 Linear Operators

Let X, Y be normed spaces with the same scalar field. A mapping T from a subspace of X, called the **domain** of T and denoted by $\mathcal{D}(T)$, into Y is said to be a **linear operator from X to Y** if

$$T(\alpha x + \beta y) = \alpha T x + \beta T y$$

for all $x, y \in \mathcal{D}(T)$ and all scalars α, β. Note that Tx rather than $T(x)$ is used when T is a linear operator. The **range**, $\mathcal{R}(T)$, of T is the set of all Tx with $x \in \mathcal{D}(T)$. The **null space**, $\mathcal{N}(T)$, of T is the set of all $x \in \mathcal{D}(T)$ such that $Tx = 0$. Thus, a linear operator T is one-to-one iff $\mathcal{N}(T) = \{0\}$. T is said to be **densely defined** if $\mathcal{D}(T)$ is dense in X. If $\mathcal{D}(T) = X$, then T is said to be **defined on** X. If $Y = X$, then T is said to be **defined in** X. If S is another linear operator from X to Y, such that $\mathcal{D}(S) \supset \mathcal{D}(T)$ and $Sx = Tx$ for $x \in \mathcal{D}(T)$, then S is called an **extension** of T and T is called a **restriction** of S. If S and T are linear operators from X to Y, then the domain of $T + S$ is $\mathcal{D}(T) \cap \mathcal{D}(S)$, and if R is a linear operator from Y to Z, then the domain of RS consists of those $x \in \mathcal{D}(S)$ for which $Sx \in \mathcal{D}(R)$.

A mapping $T : \mathcal{D}(T) \subset X \to Y$ is said to be **bounded** if it maps bounded sets into bounded sets. When T is linear, T is bounded iff the norm of T defined by

$$\|T\| = \sup\{\|Tx\| \mid x \in \mathcal{D}(T), \|x\| \leq 1\} \tag{1.11}$$

is finite. Equivalently, a linear operator T is bounded iff there exists $m \in [0, \infty)$ such that

$$\|Tx\| \leq m\|x\| \quad \text{for all} \quad x \in \mathcal{D}(T).$$

The smallest of such m is equal to $\|T\|$ when T is bounded.

EXAMPLE 1.4.1 Suppose $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^m$ and $t : B \times A \to \mathbb{C}$ are measurable, and that for some $M_1, M_2 \in (0, \infty)$ we have that

$$\int_A |t(x,y)|dy \leq M_1 \text{ for } x \in B, \qquad \int_B |t(x,y)|dx \leq M_2 \text{ for } y \in A.$$

Fix $p \in [1, \infty]$ and define an integral operator $T : L^p(A) \to L^p(B)$ by

$$(Tf)(x) = \int_A t(x,y)f(y)dy \quad \text{for} \quad f \in L^p(A),\ x \in B.$$

T is a bounded linear operator and $\|T\| \leq M_1^{1/q} M_2^{1/p}$ where $q \in [1, \infty]$ is such that $1/q + 1/p = 1$. To see this, when $p \in (1, \infty)$, note that Hölder's inequality implies

$$\begin{aligned} \left| \int_A t(x,y)f(y)dy \right|^p &\leq \left(\int_A |t(x,y)|dy \right)^{\frac{p}{q}} \int_A |t(x,y)||f(y)|^p dy \\ &\leq M_1^{\frac{p}{q}} \int_A |t(x,y)||f(y)|^p dy. \end{aligned}$$

Hence another integration and the Fubini Theorem imply the bound.

Theorem 1.4.2 *A linear operator is bounded iff it is continuous.*

PROOF If T is continuous at 0, then there exists $\delta > 0$ such that $\|Tx\| < 1$ for $x \in \mathcal{D}(T)$ with $\|x\| < \delta$. Hence $\|T\| \leq 1/\delta$. If T is bounded, then $\|Tx - Ty\| \leq \|T\|\|x - y\|$ implies continuity. □

Theorem 1.4.3 *Let $T : \mathcal{D}(T) \subset X \to Y$ be a densely defined bounded linear operator from a normed space X to a Banach space Y. Then there exists a unique continuous extension, say T^e, of T on X. Moreover, T^e is linear, bounded and $\|T\| = \|T^e\|$.*

PROOF Since T is densely defined, for each $x \in X$ there is a sequence $\{x_n\} \subset \mathcal{D}(T)$ with $x_n \to x$. Since T is bounded, $\{Tx_n\}$ is a Cauchy sequence in Y, and thus converges to some $y \in Y$. Moreover, it is easy to show that y is independent of the sequence used. Thus we can (and to obtain a continuous extension of T must) define an operator $T^e : X \to Y$ by $T^e x = \lim_{n\to\infty} Tx_n = y$. T^e is clearly a linear extension of T and thus (1.11) implies $\|T^e\| \geq \|T\|$. Since

$$\|T^e x\| = \lim_{n\to\infty} \|Tx_n\| \leq \|T\|\|x\|,$$

it follows that $\|T^e\| = \|T\|$. □

The collection of all bounded linear operators T from X to Y with $\mathcal{D}(T) = X$ will be denoted by $\mathfrak{B}(X,Y)$; define also $\mathfrak{B}(X) = \mathfrak{B}(X,X)$. Observe that

$$\|TS\| \leq \|T\|\|S\| \quad \text{if} \quad S \in \mathfrak{B}(X,Y),\ T \in \mathfrak{B}(Y,Z).$$

Theorem 1.4.4 *If X and Y are normed spaces, then $\mathfrak{B}(X,Y)$ is a normed space with norm defined by equation (1.11). If Y is a Banach space, then $\mathfrak{B}(X,Y)$ is also a Banach space.*

PROOF If $S, T \in \mathfrak{B}(X,Y)$, then

$$\|(S+T)x\| = \|Sx + Tx\| \leq \|Sx\| + \|Tx\| \leq \|S\| + \|T\|$$

for $x \in X$ with $\|x\| \leq 1$. This gives the triangle inequality. Clearly, if $\alpha \in \mathbb{K}$, then $\|\alpha T\| = |\alpha|\|T\|$. If $T \neq 0$, then $Tx \neq 0$ for some $x \in X$ and hence $\|T\| \neq 0$. Therefore $\mathfrak{B}(X,Y)$ is a normed space. Assume now that Y is complete and that $\{T_n\}$ is a Cauchy sequence in $\mathfrak{B}(X,Y)$. Since

$$\|T_n x - T_m x\| \leq \|T_n - T_m\|\|x\|, \tag{1.12}$$

we see that $\{T_n x\}$ is a Cauchy sequence in Y. We can therefore define a linear operator T by

$$Tx = \lim_{n\to\infty} T_n x \quad \text{for all} \quad x \in X.$$

If $\varepsilon > 0$, then the right side of (1.12) is smaller than $\varepsilon\|x\|$ provided that m and n are large enough. Thus

$$\|Tx - T_m x\| \leq \varepsilon\|x\| \quad \text{for all large enough } m.$$

Hence $\|Tx\| \leq (\|T_m\| + \varepsilon)\|x\|$, $T \in \mathfrak{B}(X,Y)$, $\lim_{n\to\infty} T_n = T$. □

The following Theorem is often referred to as the **uniform boundedness principle.**

Theorem 1.4.5 (Banach-Steinhaus) *Suppose X is a Banach space, Y is a normed space and $A \subset \mathfrak{B}(X,Y)$. Then either*

$$\sup_{T\in A} \|T\| < \infty$$

or there exists a dense set $E \subset X$ such that

$$\sup_{T\in A} \|Tx\| = \infty \quad \textit{for all} \quad x \in E.$$

PROOF Let $V_n = \{x \mid \|Tx\| > n \text{ for some } T \in A\}$ for $n = 1, 2, 3, \ldots$ Observe that each V_n is open.

If one of these sets, say V_N, fails to be dense in X, then $B(x_0, r) \cap V_N = \emptyset$ for some $x_0 \in X$ and some $r > 0$, i.e.

$$\|Tx\| \leq N \quad \text{for all} \quad T \in A,\ x \in B(x_0, r).$$

Therefore, if $\|x\| < r$, then

$$\|Tx\| = \|T(x_0 + x) - Tx_0\| \leq 2N \quad \text{for all} \quad T \in A$$

and $\|T\| \leq 2N/r$ for all $T \in A$.

The other possibility is that every V_n is dense in X. In that case the Baire Theorem 1.1.1 implies that $E = \cap V_n$ is dense in X. Now, if $x \in E$, then for each n there exists $T \in A$ such that $\|Tx\| > n$. □

An immediate consequence of the Theorem is:

Corollary 1.4.6 *If X is a Banach space, Y is a normed space and if $B_n \in \mathfrak{B}(X,Y)$, $n = 1, 2, \ldots$ are such that $\lim_{n\to\infty} B_n x$ exists for every $x \in X$, then there exists $B \in \mathfrak{B}(X,Y)$ such that $Bx = \lim_{n\to\infty} B_n x$ for every $x \in X$.*

Theorem 1.4.7 (Open Mapping) *If X and Y are Banach spaces and $T \in \mathfrak{B}(X,Y)$ is onto, then there exists $c < \infty$ such that for each $y \in Y$ there corresponds $x \in X$ with the properties $Tx = y$ and $\|x\| \leq c\|y\|$.*

PROOF Let $V_n = \{x \in X \mid \|x\| < n\}$ for $n = 1, 2, \ldots$ Since T is onto, $Y = \cup T(V_n)$. Corollary 1.1.2 of the Baire Theorem implies that some $\overline{T(V_N)}$ contains a nonempty, open W. Choose $y_0 \in W$ and $\eta > 0$ such that $\|y\| \leq \eta$ implies $y_0 + y \in W$. Since there exist $u_i, v_i \in V_N$ such that

$$Tu_i \to y_0, \quad Tv_i \to y_0 + y$$

we have $Tx_i = T(v_i - u_i) \to y$ and $\|x_i\| \leq 2N$. If $c = 4N/\eta$, then rescaling implies that *for each* $y \in Y$, $\varepsilon > 0$ *there exists* $x \in X$ *such that*

$$\|x\| \leq c\|y\|/2, \quad \|y - Tx\| < \varepsilon. \tag{1.13}$$

Now choose $y \in Y$ with $\|y\| \leq 1$. There exists an $x_1 \in X$ such that $\|x_1\| \leq c/2$, $\|y - Tx_1\| < 1/2$. Suppose $x_1, \ldots, x_n$ have been chosen so that

$$\|y - Tx_1 - \cdots - Tx_n\| < 2^{-n}. \tag{1.14}$$

Use (1.13), with y replaced by $y - Tx_1 - \cdots - Tx_n$, to obtain x_{n+1} such that (1.14) holds with n replaced by $n+1$ and

$$\|x_{n+1}\| \leq 2^{-n-1}c, \quad n = 0, 1, 2, \ldots. \tag{1.15}$$

Let $s_n = x_1 + \cdots + x_n$. (1.15) implies that $\{s_n\}$ is a Cauchy sequence hence it converges to some x. (1.15) implies $\|x\| \leq c$ and $Tx = y$ by (1.14). □

Corollary 1.4.8 *If X and Y are Banach spaces and if $T \in \mathfrak{B}(X, Y)$ is one-to-one and onto, then the inverse of T is bounded.*

Definition 1.4.9 A linear operator T from X to Y is said to be **closed** if for every sequence $\{x_n\}$ in $\mathcal{D}(T)$, which is such that the limits

$$x = \lim_{n\to\infty} x_n \quad \text{and} \quad y = \lim_{n\to\infty} Tx_n$$

exist, we have that $x \in \mathcal{D}(T)$ and $Tx = y$. T is said to be **closable** if there exists an extension of T that is closed.

EXAMPLE 1.4.10 Let $T : \mathcal{D}(T) \subset L^p(0,1) \to L^p(0,1)$, $p \in [1, \infty]$, be the differential operator given by $Tf = f'$ with domain

$$\mathcal{D}(T) = \{f \in AC[0,1] \mid f' \in L^p(0,1),\ f(0) = 0\}.$$

We are using $AC[a,b]$ to denote the set of all absolutely continuous scalar valued functions defined on the interval $[a,b]$.

If $g_n(x) = n^{1/p}x^n$, then $\|g_n\| \leq 1$ and $\|Tg_n\| \to \infty$ as $n \to \infty$. Thus T is not bounded. However, if $f_n \in \mathcal{D}(T)$ with $f_n \to f$ and $Tf_n \to g$, then

$$f_n(x) = \int_0^x f_n'(t)dt$$

and, by Hölder's inequality, the above integral converges uniformly to $\int_0^x g(t)dt$. Thus f_n converge uniformly to f. As a consequence, $f(x) = \int_0^x g(t)dt$, which implies that $f \in \mathcal{D}(T)$, $Tf = g$ and hence T is closed. Example 1.3.4 shows that T is densely defined if $p < \infty$.

Suppose that a linear operator T from X to Y has the following property: *if* $\{x_n\}$ *is a sequence in* $\mathcal{D}(T)$ *such that* $\lim_{n\to\infty} x_n = 0$ *and* $\{Tx_n\}$ *converges to some* $y \in Y$, *then* $y = 0$. Note that all closable operators have this property. For such T define **closure** of T, denoted by $\overline{T}$, as follows: $x \in \mathcal{D}(\overline{T})$ if and only if there exists a sequence $\{x_n\}$ in $\mathcal{D}(T)$ which is such that $\lim_{n\to\infty} x_n = x$ and $\{Tx_n\}$ converges to some $y \in Y$; for such x, y we define $\overline{T}x = y$. It is easy to check that $\overline{T}$ is an extension of T and that $\overline{T}$ is closed.

It is left to the reader to verify the following properties a linear operator T from X to Y:

(a) T is closed iff its graph $\{(x, Tx) \mid x \in \mathcal{D}(T)\}$ is closed in $X \times Y$.

(b) If T is continuous and $\mathcal{D}(T)$ is closed, then T is closed.

(c) If T is closed, then $\mathcal{N}(T)$ is closed.

(d) If T is one-to-one, then T is closed iff T^{-1} is closed.

(e) If T is continuous, then T is closable.

Theorem 1.4.11 (Closed Graph) *If* X *and* Y *are Banach spaces and* T *is a closed linear operator from* X *to* Y *with* $\mathcal{D}(T) = X$, *then* T *is bounded.*

PROOF Define $|x| = \|x\| + \|Tx\|$ for $x \in X$. One can easily check that $|\cdot|$ is also a norm on X. Closedness of T implies that X is also complete with this norm. Consider I, the identity map on X, as an operator from the Banach space X with norm $|\cdot|$ onto the Banach space X with norm $\|\cdot\|$. I is clearly bounded and the Open Mapping Theorem 1.4.7 implies that there exists $c < \infty$ such that $|x| \le c\|x\|$ for all $x \in X$. Hence $\|T\| \le c - 1$. □

1.5 Duals

When X is a normed space, the set $\mathfrak{B}(X, \mathbb{K})$ will be called the **dual space** of X and will be denoted by $X^\star$. $X^\star$ is also called the **adjoint space** or **conjugate space** of X. Elements of $X^\star$ are called **bounded linear functionals** on X. Theorem 1.4.4 implies that $X^\star$ is a Banach space.

EXAMPLE 1.5.1 For $f = \{f_n\} \in \ell^\infty$, define

$$F(x) = \sum_{n=1}^{\infty} f_n x_n \tag{1.16}$$

for all $x = \{x_n\} \in \ell^1$. Note that $F \in \ell^{1\star}$ and that $\|F\| = \|f\|_\infty$. Suppose now that $F \in \ell^{1\star}$ and let us show that equation (1.16) holds for some $f = \{f_n\} \in \ell^\infty$. For $i \ge 1$,

define $e^{(i)} = \{e_j^{(i)}\} \in \ell^1$ by $e_i^{(i)} = 1$, and let $e_j^{(i)} = 0$ if $j \neq i$. Define $f_i = F(e^{(i)})$ and note that $|f_i| \leq \|F\|\|e^{(i)}\|_1 = \|F\|$ and therefore $f = \{f_j\} \in \ell^\infty$. If $x \in \ell^1$, then

$$|F(x) - \sum_{i=1}^{n} f_i x_i| = |F(x) - F(\sum_{i=1}^{n} x_i e^{(i)})| \leq \|F\| \sum_{i=n+1}^{\infty} |x_i|.$$

Letting $n \to \infty$ proves the assertion. Thus one can, loosely, say that $\ell^\infty = \ell^{1\star}$. If $p \in [1, \infty)$, then one can show in a similar way that the linear functionals on ℓ^p can be represented by equation (1.16) with $f \in \ell^q$, where q is such that $1/p + 1/q = 1$ and therefore $\ell^q = \ell^{p\star}$.

EXAMPLE 1.5.2 Suppose $1 \leq p, q \leq \infty$ satisfy $1/p + 1/q = 1$ and let Ω be a nonempty Lebesgue measurable set in $\mathbb{R}^n$. For $g \in L^q(\Omega)$, define

$$G(f) = \int_\Omega gf \quad \text{for all} \quad f \in L^p(\Omega). \tag{1.17}$$

It is easy to see that G is a bounded linear functional on $L^p(\Omega)$ and that $\|G\| = \|g\|_q$. If $1 \leq p < \infty$, then one can show that for each $G \in L^p(\Omega)^\star$ there corresponds a unique $g \in L^q(\Omega)$ so that equation (1.17) holds. Hence, it is customary to identify $L^p(\Omega)^\star = L^q(\Omega)$. The case of $p = \infty$ is different. The dual of L^∞ is much larger than L^1.

Theorem 1.5.3 *Suppose $\Lambda_1, \Lambda_2, \ldots$ are bounded linear functionals on a separable Banach space X such that $\sup_i |\Lambda_i(x)| < \infty$ for all $x \in X$. Then $\sup_i \|\Lambda_i\| < \infty$ and there exist integers $n_1 < n_2 < \cdots$ and a bounded linear functional Λ on X such that*

$$\lim_{k\to\infty} \Lambda_{n_k}(x) = \Lambda(x) \quad \textit{for all} \quad x \in X. \tag{1.18}$$

PROOF The uniform boundedness principle implies $c \equiv \sup_i \|\Lambda_i\| < \infty$. Since $|\Lambda_i(x) - \Lambda_i(y)| \leq c\|x - y\|$, the Arzela-Ascoli Theorem 1.1.5 implies existence of $n_1 < n_2 < \cdots$ and $\Lambda \in C(X)$ such that (1.18) holds. It is easy to check that Λ is a bounded linear functional on X. □

Corollary 1.5.4 *Suppose $1 < p \leq \infty$, $1/p + 1/q = 1$ and let Ω be a nonempty Lebesgue measurable set in $\mathbb{R}^n$. If $f_1, f_2, \ldots$ is a bounded sequence in $L^p(\Omega)$, then there exist integers $n_1 < n_2 < \cdots$ and $f \in L^p(\Omega)$ such that*

$$\lim_{k\to\infty} \int_\Omega f_{n_k} g = \int_\Omega fg \quad \textit{for all} \quad g \in L^q(\Omega).$$

PROOF Let $\Lambda_i(g) = \int f_i g$ for $g \in X \equiv L^q(\Omega)$. Since X is separable, see Example 1.3.4, Theorem 1.5.3 gives integers $n_1 < n_2 < \cdots$ and $\Lambda \in X^\star$ such that $\Lambda_{n_k}(g) \to \Lambda(g)$ for all $g \in X$. In view of Example 1.5.2, there exists $f \in L^p(\Omega)$ such that $\Lambda(g) = \int fg$ for all $g \in X$. □

A sequence $x_1, x_2, \ldots$ in a Banach space X is said to be **weakly convergent** if there exists $x \in X$ such that $\lim_{n\to\infty} \Lambda(x_n) = \Lambda(x)$ for every $\Lambda \in X^\star$. In this case x is called a **weak limit** of the sequence and the notation $x_n \rightharpoonup x$ is used. Note that Corollary 1.5.4 implies that every bounded sequence in $L^p(\Omega)$, $1 < p < \infty$, has a weakly convergent subsequence.

In the proof of the following Hahn-Banach Theorem we shall need to make use of transfinite induction. Our argument will be based on the Hausdorff Maximality Theorem which we shall now briefly review.

A set $\mathcal{P}$ is said to be **partially ordered** by $\prec$ if

(1) $a \prec a$ for every $a \in \mathcal{P}$

(2) $a \prec b$ and $b \prec a$, then $a = b$

(3) $a \prec b$ and $b \prec \mathrm{c}$, then $a \prec \mathrm{c}$.

For example, any collection of subsets of a given set is partially ordered by $\subset$.

A subset $\mathcal{T}$ of a partially ordered set $\mathcal{P}$ is said to be **totally ordered** if every pair $a, b \in \mathcal{T}$ satisfies either $a \prec b$ or $b \prec a$. A totally ordered set $\mathcal{T}$ is said to be a **maximal totally ordered** subset of $\mathcal{P}$ if there does not exist a totally ordered set $\mathcal{T}' \subset \mathcal{P}$ such that $\mathcal{T} \subset \mathcal{T}'$ and $\mathcal{T} \neq \mathcal{T}'$.

Theorem 1.5.5 (Hausdorff Maximality) *Every nonempty, partially ordered set contains a maximal totally ordered subset.*

This Theorem happens to be equivalent to the Axiom of Choice and to the Zorn Lemma. Hence, depending on one's preferences, it can be taken as a basic assumption (our choice) or derived from either the Axiom of Choice or the Zorn Lemma.

Theorem 1.5.6 (Hahn-Banach) *Suppose M is a subspace of a normed space X and that $f \in M^\star$. Then there exists $F \in X^\star$ such that $\|F\| = \|f\|$ and*

$$F(x) = f(x) \quad \textit{for all} \quad x \in M.$$

PROOF Let us suppose first that $\mathbb{K} = \mathbb{R}$ so that f is real valued. Let $\mathrm{c} = \|f\|$.

Choose any $x_0 \in X \backslash M$ and let

$$M_0 = \{\, x + \lambda x_0 \mid x \in M, \lambda \in \mathbb{R} \,\}.$$

Observe that for all $x, y \in M$,

$$f(x - y) \leq \mathrm{c}\|x - y\| \leq \mathrm{c}\|x - x_0\| + \mathrm{c}\|x_0 - y\|$$

$$f(x) - \mathrm{c}\|x - x_0\| \leq f(y) + \mathrm{c}\|y - x_0\|.$$

Hence one can find $a \in \mathbb{R}$ such that for all $x \in M$,

$$f(x) - c\|x - x_0\| \leq a \leq f(x) + c\|x - x_0\|$$

$$|f(x) - a| \leq c\|x - x_0\|.$$

Replacing x by $-x/\lambda$ gives

$$|f(x) + \lambda a| \leq c\|x + \lambda x_0\| \quad \text{for all} \quad x \in M,\ \lambda \in \mathbb{R}.$$

Hence, if $f_0 : M_0 \to \mathbb{R}$ is defined by

$$f_0(x + \lambda x_0) = f(x) + \lambda a,$$

then $f_0 \in M_0^\star$ and f_0 is an extension of f with $\|f_0\| = c$.

Let $\mathcal{P}$ be the collection of all pairs (N, g) where $N \supset M$ is a subspace of X and $g \in N^\star$ is an extension of f with $\|g\| = c$. Define partial ordering $\prec$ on $\mathcal{P}$ by

$$(N, g) \prec (N', g') \quad \text{if and only if} \quad N \subset N' \text{ and } g' \text{ extends } g.$$

The Hausdorff Maximality Theorem implies the existence of a maximal totally ordered subcollection $\mathcal{T}$ of $\mathcal{P}$.

Define S to be the union of all N where $(N, g) \in \mathcal{T}$. The fact that $\mathcal{T}$ is totally ordered implies that S is a subspace of X and allows us to define $F : S \to \mathbb{R}$ as follows: if $x \in S$, then x belongs to some N with $(N, g) \in \mathcal{T}$, let $F(x) = g(x)$. Note that $(S, F) \in \mathcal{P}$, and since $m \prec (S, F)$ for all $m \in \mathcal{T}$, the maximality of $\mathcal{T}$ implies $(S, F) \in \mathcal{T}$. If S would be a proper subspace of X, the first part of the proof would give us a further extension of F and this would contradict maximality of $\mathcal{T}$. Thus $S = X$, and the proof is complete for the case of real scalars.

Suppose now that $\mathbb{K} = \mathbb{C}$. Note that M, X are also real vector spaces and that if $u(x) = \operatorname{Re} f(x)$ for $x \in M$, then $u \in \mathfrak{B}(M, \mathbb{R})$ and, clearly, $\|u\| \leq \|f\|$. Let $U \in \mathfrak{B}(X, \mathbb{R})$ be the extension of u as obtained in the case of real scalars. Define

$$F(x) = U(x) - iU(ix) \quad \text{for all} \quad x \in X.$$

One can easily see that F is an extension of f and $F \in X^\star$. If $x \in X$ and if $\alpha \in \mathbb{C}$ is such that $\alpha F(x) = |F(x)|$, then $|F(x)| = F(\alpha x) = U(\alpha x) \leq \|U\|\|x\|$. Hence $\|F\| \leq \|U\| = \|u\| \leq \|f\|$ and, therefore, $\|F\| = \|f\|$. □

Theorem 1.5.7 *Let M be a subspace of a normed space X and assume that $x \in X$ does not belong to the closure of M. Then there exists $f \in X^\star$ such that*

$$\|f\| = 1, \quad f(x) = \operatorname{dist}(x, M) > 0, \quad f(y) = 0 \ \textit{for all}\ y \in M.$$

PROOF Define a subspace $N = \{y+\lambda x | y \in M, \lambda \in \mathbb{K}\}$ and let $g(y+\lambda x) = \lambda d$ for all $y \in M, \lambda \in \mathbb{K}$, where $d = \text{dist}(x, M)$. If $y \in M, \lambda \in \mathbb{K}\backslash\{0\}$, then

$$\|x - (-1/\lambda)y\| \geq d = g(x + (1/\lambda)y).$$

Hence $g \in N^\star$, $\|g\| \leq 1$. If $\{y_n\}$ is a sequence in M such that $\|x - y_n\| \to d$, then $|g(y_n - x)|/\|y_n - x\| = d/\|y_n - x\| \to 1$, and hence $\|g\| = 1$. Let f be the extension of g as provided by the Hahn-Banach Theorem 1.5.6. □

If X is a normed space and $x \in X$, then $f \in X^\star$ which is such that

$$f(x) = \|x\|^2 = \|f\|^2$$

is called a **normalized tangent functional** to x. Taking $M = \{0\}$ in the above Theorem 1.5.7 and rescaling proves their existence:

Corollary 1.5.8 *If X is a normed space and $x \in X$, then there exists $f \in X^\star$ such that $f(x) = \|x\|^2 = \|f\|^2$.*

EXAMPLE 1.5.9 Let $\|\cdot\|$ be the sup norm on $X = C[0,1]$. Pick $u \in X$ and let $x_0 \in [0,1]$ be such that $|u(x_0)| = \|u\|$. Define $f \in X^\star$ by $f(v) = \overline{u(x_0)}v(x_0)$. f is a normalized tangent functional to u. Note that u can have many normalized tangent functionals.

The above Corollary 1.5.8 implies that the dual of a normed space X separates points in X:

Corollary 1.5.10 *If X is a normed space and if $x, y \in X$ are such that $f(x) = f(y)$ for all $f \in X^\star$, then $x = y$.*

Definition 1.5.11 Let X be a normed space. The dual space $X^{\star\star}$ of $X^\star$ is called the **second dual space** of X and is again a Banach space. Note that to each $x \in X$ we can associate $F_x \in X^{\star\star}$ by $F_x(f) = f(x)$ for $f \in X^\star$. Since $|F_x(f)| \leq \|f\|\|x\|$, we have that $\|F_x\| \leq \|x\|$; by taking f to be a normalized tangent functional to x gives that $\|F_x\| = \|x\|$. Thus, the mapping $F : x \to F_x$ is a linear isometry of X onto a subspace of $X^{\star\star}$. Frequently, $F(X)$ is identified with X and, in this case, X is considered as a subspace of $X^{\star\star}$. X is said to be **reflexive** if $F(X) = X^{\star\star}$, i.e., X is reflexive iff for every $g \in X^{\star\star}$ there corresponds $x \in X$ such that $F_x = g$. For example, all L^p spaces, including ℓ^p, with $1 < p < \infty$ are reflexive. ℓ^1 is an example of a space that is not reflexive.

Definition 1.5.12 Let X and Y be normed spaces and let T be a *densely* defined linear operator from X to Y. Define a linear operator from $Y^\star$ to $X^\star$, called the

adjoint of T and denoted by $T^\star$, as follows: $g \in \mathcal{D}(T^\star)$ if and only if $g \in Y^\star$ and there exists $c < \infty$ such that

$$|g(Tx)| \leq c\|x\| \quad \text{for all} \quad x \in \mathcal{D}(T).$$

If $g \in \mathcal{D}(T^\star)$, then the fact that $\mathcal{D}(T)$ is dense in X implies (see Theorem 1.4.3) that there exists a unique $f \in X^\star$ such that $f(x) = g(Tx)$ for all $x \in \mathcal{D}(T)$. Define $T^\star g = f$.

Clearly, $T^\star$ is a linear operator and it is easy to show that $T^\star$ is always closed.

EXAMPLE 1.5.13 Suppose $p \in [1, \infty)$ and let $T \in \mathfrak{B}(L^p(A), L^p(B))$ be given by

$$(Tf)(x) = \int_A t(x,y)f(y)dy \quad \text{for} \quad f \in L^p(A),\ x \in B$$

as in Example 1.4.1. In view of identification $L^{p\star} = L^q$, where $1/q + 1/p = 1$ (see Example 1.5.2) the Fubini Theorem implies that

$$(T^\star g)(y) = \int_B t(x,y)g(x)dx \quad \text{for} \quad g \in L^q(B),\ y \in A.$$

Reapplying Example 1.4.1 gives $\|T^\star\| \leq M_1^{1/q} M_2^{1/p}$.

Theorem 1.5.14 *Let X and Y be normed spaces. If $T \in \mathfrak{B}(X,Y)$, then $T^\star \in \mathfrak{B}(Y^\star, X^\star)$ and $\|T\| = \|T^\star\|$.*

PROOF Since for all $g \in Y^\star$ and all $x \in X$,

$$|(T^\star g)(x)| = |g(Tx)| \leq \|g\|\|T\|\|x\|,$$

we have that $\|T^\star\| \leq \|T\|$. Corollary 1.5.8 implies that for every $x \in X$ there exists $g_x \in Y^\star$ such that $\|g_x\| \leq 1$ and $g_x(Tx) = \|Tx\|$. Hence

$$\|T^\star\| \geq \|T^\star g_x\| \geq g_x(Tx) = \|Tx\| \quad \text{for all} \quad x \in X,\ \|x\| \leq 1$$

implies that $\|T^\star\| \geq \|T\|$. □

Theorem 1.5.15 *If X and Y are normed spaces, Y is reflexive and T is a closed, densely defined, linear operator from X to Y, then $\mathcal{D}(T^\star)$ is dense in $Y^\star$.*

PROOF If $\mathcal{D}(T^\star)$ is not dense in $Y^\star$, then Theorem 1.5.7 and the fact that Y is reflexive imply that there exists $y \in Y$ such that $y \neq 0$ and $\ell(y) = 0$ for all $\ell \in \mathcal{D}(T^\star)$. Note that $(0, y) \notin M \equiv \{(x, Tx) \mid x \in \mathcal{D}(T)\}$, and since M is closed in $X \times Y$, Theorem 1.5.7 (see Exercise 10) implies the existence of $\ell_1 \in X^\star$ and $\ell_2 \in Y^\star$ such that $\ell_2(y) \neq 0$ and $\ell_1(x) + \ell_2(Tx) = 0$ for all $x \in \mathcal{D}(T)$. This implies that $\ell_2 \in \mathcal{D}(T^\star)$ and hence the contradiction $\ell_2(y) = 0$. □

Theorem 1.5.16 *Let X and Y be Banach spaces and suppose that T is a closed, densely defined, linear operator from X to Y. Then the following two statements are equivalent:*

(a) T is one-to-one, onto and $T^{-1} \in \mathfrak{B}(Y, X)$.

(b) $T^\star$ is one-to-one, onto and $T^{\star -1} \in \mathfrak{B}(X^\star, Y^\star)$.

Moreover, if (a) holds, then $(T^{-1})^\star = T^{\star -1}$.

PROOF Assume (a) and let $R = T^{-1}$. Note that for every $y \in Y$ and $g \in \mathcal{D}(T^\star)$,

$$g(y) = g(TRy) = (T^\star g)(Ry) = (R^\star T^\star g)(y).$$

Hence $g = R^\star T^\star g$ and $T^\star$ is one-to-one. For $f \in X^\star$ and $x \in \mathcal{D}(T)$,

$$f(x) = f(RTx) = (R^\star f)(Tx) \quad \text{hence} \quad (R^\star f) \in \mathcal{D}(T^\star) \text{ and } T^\star R^\star f = f.$$

Therefore $T^\star$ is onto, $T^{\star -1} = R^\star$ and, by Theorem 1.5.14, we have proved (b) as well as the 'moreover' part.

Assume (b). Choose any $x \in \mathcal{D}(T)$. Corollary 1.5.8 implies there is an $f \in X^\star$ such that $\|f\| \leq 1$ and $f(x) = \|x\|$. Hence if $g = T^{\star -1} f$, then

$$\|x\| = f(x) = (T^\star g)(x) = g(Tx) = (T^{\star -1} f)(Tx) \leq \|T^{\star -1}\| \|Tx\|. \tag{1.19}$$

Thus T is one-to-one. If a sequence $\{Tx_n\}$ converges to some $y \in Y$, then (1.19) implies that $\{x_n\}$ converges to some $x \in X$. Closedness of T implies that $x \in \mathcal{D}(T)$, $Tx = y$ and therefore $\mathcal{R}(T)$ is closed. If there would exist $y \in Y \backslash \mathcal{R}(T)$, then Theorem 1.5.7 would imply existence of $g \in Y^\star$ such that $g(y) > 0$ and $g(Tx) = 0$ for all $x \in \mathcal{D}(T)$. However, this implies $g \in \mathcal{D}(T^\star)$, $T^\star g = 0$, and, since $T^\star$ is one-to-one we would have $g = 0$ which contradicts $g(y) > 0$. Therefore T is onto, and $T^{-1} \in \mathfrak{B}(Y, X)$, by the inequality (1.19). □

EXAMPLE 1.5.17 Suppose $T : \mathcal{D}(T) \subset L^p(0,1) \to L^p(0,1)$, $p \in [1, \infty)$, is given by

$$Tf = f' \quad \text{for} \quad f \in \mathcal{D}(T) = \{f \in AC[0,1] \mid f' \in L^p(0,1),\ f(0) = 0\}.$$

In Example 1.4.10 it is shown that T is closed and densely defined. It can be easily seen that T is one-to-one, onto and

$$(T^{-1} g)(x) = \int_0^x g(t)dt \quad \text{for} \quad g \in L^p(0,1).$$

Thus Theorem 1.5.16 implies $T^{\star -1} = (T^{-1})^\star$ and, in view of Example 1.5.13, we have

$$(T^{\star -1} h)(x) = \int_x^1 h(t)dt \quad \text{for} \quad h \in L^q(0,1),$$

$1/q + 1/p = 1$. Therefore, $T^\star$ can be identified with the following linear operator in $L^q(0,1)$:

$$T^\star g = -g' \quad \text{for} \quad g \in \mathcal{D}(T^\star) = \{g \in AC[0,1] \mid g' \in L^q(0,1),\ g(1) = 0\}.$$

1.6 Spectrum

Definition 1.6.1 Let T be a linear operator in a normed space X with scalar field $\mathbb{K}$. The **resolvent set** of T, denoted by $\rho(T)$, is defined to be the set of all scalars $\lambda \in \mathbb{K}$ for which there exists $R(\lambda) \in \mathfrak{B}(X)$ such that

(1) for every $y \in X$ we have that $R(\lambda)y \in \mathcal{D}(T)$ and $(T-\lambda)R(\lambda)y = y$,

(2) $R(\lambda)(T-\lambda)x = x$ for all $x \in \mathcal{D}(T)$.

When $\lambda \in \rho(T)$, $R(\lambda)$ is called the **resolvent** of T at λ and will be usually denoted by $(T-\lambda)^{-1}$. $\sigma(T) = \mathbb{K}\backslash\rho(T)$ is called the **spectrum** of T. The set of $\lambda \in \mathbb{K}$ for which there exists $x \in \mathcal{D}(T)$, $x \neq 0$, such that $Tx = \lambda x$, is called the **point spectrum** of T and is denoted by $\sigma_p(T)$. The elements of $\sigma_p(T)$ are called the **eigenvalues** of T and the nonzero members of $\mathcal{N}(T-\lambda)$ are called the **eigenvectors** of T.

Note that if $\rho(T)$ is not empty, then T has to be closed. Observe that if (1) is satisfied and λ is not an eigenvalue of T, then (2) is also satisfied. Also, if $\lambda \in \rho(T)$, then $T-\lambda$ is one-to-one and onto. In particular $\sigma_p(T) \subset \sigma(T)$. The converse follows immediately from the Closed Graph Theorem:

Corollary 1.6.2 *If T is a closed linear operator in a Banach space, then $\lambda \in \rho(T)$ if and only if*

$$T-\lambda \text{ is one-to-one and onto.}$$

Suppose $\lambda, \zeta \in \rho(T)$. If $x \in X$, then $y = (T-\lambda)^{-1}x - (T-\zeta)^{-1}x \in \mathcal{D}(T)$ and

$$(T-\lambda)y = x - (T-\zeta+\zeta-\lambda)(T-\zeta)^{-1}x = (\lambda-\zeta)(T-\zeta)^{-1}x.$$

Applying $(T-\lambda)^{-1}$ on both sides gives the **resolvent identity**

$$(T-\lambda)^{-1} - (T-\zeta)^{-1} = (\lambda-\zeta)(T-\lambda)^{-1}(T-\zeta)^{-1} \quad \text{for} \quad \lambda, \zeta \in \rho(T). \tag{1.20}$$

Observe that this identity implies that $(T-\lambda)^{-1}$, $(T-\zeta)^{-1}$ commute for $\lambda, \zeta \in \rho(T)$.

EXAMPLE 1.6.3 Suppose $T_0 : \mathcal{D}(T_0) \subset L^p(0,1) \to L^p(0,1)$, $p \in [1,\infty]$, is given by

$$T_0 f = f' \quad \text{for} \quad f \in \mathcal{D}(T_0) = \{f \in C^1[0,1] \mid f(0) = 0\}.$$

Since $T_0 f - \lambda f \in C[0,1]$, $T_0 - \lambda$ is never onto and hence $\sigma(T_0) = \mathbb{K}$.

EXAMPLE 1.6.4 Suppose $T_1 : \mathcal{D}(T_1) \subset L^p(0,1) \to L^p(0,1)$, $p \in [1,\infty]$, is given by

$$T_1 f = f' \quad \text{for} \quad f \in \mathcal{D}(T_1) = \{f \in AC[0,1] \mid f' \in L^p(0,1),\ f(0) = f(1)\}$$

and let $\mathbb{K} = \mathbb{C}$.

It can be easily seen that λ is an eigenvalue of T_1 iff $e^{\lambda} = 1$ iff $\lambda = 2in\pi$, $n = 0, \pm 1, \ldots$ The corresponding eigenvectors are $e^{\lambda x}$. If $\lambda \notin \sigma_p(T_1)$, then solving $f' - \lambda f = g$, $f(0) = f(1)$, suggests that we define

$$(R(\lambda)g)(x) = \int_0^1 t(x,y)g(y)dy \quad \text{where} \quad t(x,y) = \frac{e^{\lambda(x-y)}}{1-e^{\lambda}} \times \begin{cases} e^{\lambda} & \text{if } x \le y \\ 1 & \text{if } x > y \end{cases}.$$

It is easy to verify that $R(\lambda) \in \mathfrak{B}(L^p(0,1))$ and that it satisfies the requirements stated in Definition 1.6.1. Thus $\sigma(T_1) = \sigma_p(T_1)$.

EXAMPLE 1.6.5 The operator $T : \mathcal{D}(T) \subset L^p(0,l) \to L^p(0,l)$, given by

$$Tf = f' \quad \text{for} \quad f \in \mathcal{D}(T) = \{f \in AC[0,l] \mid f' \in L^p(0,l),\ f(0) = 0\},$$

has an empty spectrum whenever $l \in (0,\infty)$, $p \in [1,\infty]$ and $\mathbb{K} = \mathbb{C}$. Its resolvent is given by

$$((T-\lambda)^{-1}g)(x) = \int_0^x e^{\lambda(x-s)}g(s)ds \quad \text{for} \quad \lambda \in \mathbb{C},\ g \in L^p(0,l).$$

EXAMPLE 1.6.6 Define $T_2 : \mathcal{D}(T_2) \subset L^p(\mathbb{R}) \to L^p(\mathbb{R})$, $p \in [1,\infty]$, by $T_2 f = f'$ with domain

$$\mathcal{D}(T_2) = \{f \in L^p(\mathbb{R}) \mid f \in AC[-a,a] \text{ for all } a \in (0,\infty),\ f' \in L^p(\mathbb{R})\},$$

and let $\mathbb{C}$ be the field of scalars.

Assume $\lambda \in \mathbb{C}$ and $\operatorname{Re}\lambda > 0$. Solving $f' - \lambda f = g$ suggests the following definition of the resolvent:

$$(R(\lambda)g)(x) = -\int_x^{\infty} e^{\lambda(x-s)}g(s)ds \quad \text{for} \quad g \in L^p(\mathbb{R}),\ x \in \mathbb{R}.$$

Applying Example 1.4.1, we see that $R(\lambda) \in \mathfrak{B}(L^p(\mathbb{R}))$ and $\|R(\lambda)\| \le 1/\operatorname{Re}\lambda$. It is easy to verify (1) in Definition 1.6.1 and, since λ cannot be an eigenvalue of T_2, we have that $\lambda \in \rho(T_2)$ and $R(\lambda) = (T_2 - \lambda)^{-1}$.

If $\lambda \in \mathbb{C}$ and $\operatorname{Re}\lambda < 0$, then one can similarly show that $\lambda \in \rho(T_2)$ and

$$((T_2-\lambda)^{-1}g)(x) = \int_{-\infty}^x e^{\lambda(x-s)}g(s)ds \quad \text{for} \quad g \in L^p(\mathbb{R}),\ x \in \mathbb{R}.$$

Thus, every $\lambda \in \mathbb{C}$, with $\operatorname{Re}\lambda \neq 0$, belongs to the resolvent set of T_2 and

$$\|(T_2-\lambda)^{-1}\| \le \frac{1}{|\operatorname{Re}\lambda|}. \tag{1.21}$$

If $p = \infty$, then $\sigma_p(T_2) = \sigma(T_2) = i\mathbb{R}$.

If $p < \infty$, then T_2 has no eigenvalues. However, if $\lambda \in \mathbb{C}$, with $\operatorname{Re}\lambda = 0$, then $T_2 - \lambda$ is not onto and hence $\lambda \in \sigma(T_2)$. To see this, try to find $f \in \mathcal{D}(T_2)$ such that $T_2 f - \lambda f = g$ where $g(x) = e^{\lambda x}$ for $x \in [0,1]$ and $g(x) = 0$ for other $x \in \mathbb{R}$.

EXAMPLE 1.6.7 Define $S : \mathcal{D}(S) \subset L^p(\mathbb{R}) \to L^p(\mathbb{R})$, $p \in [1,\infty]$, by $Sf = -f''$ with domain

$$\mathcal{D}(S) = \{f \in C^1(\mathbb{R}) \mid f' \in AC[-a,a] \text{ for all } a \in (0,\infty);\ f, f', f'' \in L^p(\mathbb{R})\}$$

and let $\mathbb{C}$ be the field of scalars.

If $\lambda \in \mathbb{C}$ and $\lambda \notin [0,\infty)$, then it can be easily seen that

$$(S-\lambda)^{-1} = -(T_2 - \zeta)^{-1}(T_2 + \zeta)^{-1},$$

where T_2 is as in Example 1.6.6 and $\zeta = (-\lambda)^{1/2}$ with $\mathrm{Re}\,\zeta > 0$. Similarly, as in Example 1.6.6, one can show that $\sigma(S) = [0,\infty)$.

Theorem 1.6.8 *Suppose X is a Banach space, $T \in \mathfrak{B}(X)$ and λ is a scalar such that $|\lambda| > \|T^n\|^{1/n}$ for some integer $n \geq 1$. Then $\lambda \in \rho(T)$ and*

$$(T-\lambda)^{-1}x = -\sum_{k=0}^{\infty} \lambda^{-k-1} T^k x \quad \textit{for all} \quad x \in X. \tag{1.22}$$

PROOF Choose $x \in X$ and define $R_x y = -\lambda^{-1}x + \lambda^{-1}Ty$ for $y \in X$. By induction,

$$R_x^k y = -\lambda^{-1}x - \lambda^{-2}Tx - \cdots - \lambda^{-k}T^{k-1}x + \lambda^{-k}T^k y \quad \text{for} \quad k \geq 1,\ y \in X.$$

Thus $\|R_x^n y - R_x^n z\| \leq |\lambda|^{-n}\|T^n\|\|y - z\|$ and hence Theorem 1.1.3 implies existence of a unique $z_x \in X$ such that $R_x z_x = z_x$, i.e. $(T-\lambda)z_x = x$. Thus, $T - \lambda$ is one-to-one and onto, hence, Corollary 1.6.2 implies that $\lambda \in \rho(T)$ and $z_x = (T-\lambda)^{-1}x$. Theorem 1.1.3 implies also that $\lim_{k\to\infty} R_x^k 0 = z_x$ which implies (1.22). □

EXAMPLE 1.6.9 *Initial value problem for a system of n linear ordinary differential equations.* Suppose $-\infty < a < b < \infty$ and let A be an $n \times n$ matrix with entries $A_{ij} \in L^1(a,b)$. Assume also that $c \in \mathbb{C}^n$ and $f = (f_1, \ldots, f_n)^\top$ where $f_i \in L^1(a,b)$. We will show that there exists a unique $g = (g_1, \ldots, g_n)^\top : [a,b] \to \mathbb{C}^n$ such that each $g_i \in AC[a,b]$, $g(a) = \mathrm{c}$ and

$$g'(x) = A(x)g(x) + f(x) \quad \text{for almost all} \quad x \in [a,b].$$

For $u = (u_1, \ldots, u_n)^\top \in X \equiv C([a,b], \mathbb{C}^n)$, define

$$\|u\| = \max_{a \leq x \leq b} \max_{1 \leq i \leq n} |u_i(x)|.$$

X is a Banach space (Example 1.3.2). For $u \in X$, define $Ku \in X$ by $(Ku)(x) = \int_a^x A(y)u(y)dy$ and let $h \in X$ be given by $h(x) = \mathrm{c} + \int_a^x f(s)ds$. Observe that it is enough to show that there exists a unique $g \in X$ such that $g = Kg + h$. This will be accomplished by showing that $1 \in \rho(K)$.

Choose any $u \in X$. Define $u^{(0)} = u$, $u^{(k+1)} = Ku^{(k)}$ for $k = 0, 1, 2, \ldots$ and observe that $K^k u = u^{(k)}$. Define also $v_k(x) = \max_{1 \le i \le n} |u_i^{(k)}(x)|$ and note

$$v_{k+1}(x) \le \int_a^x G(s) v_k(s)\, ds \quad \text{where} \quad G(x) = \max_{1 \le i \le n} \sum_{j=1}^{n} |A_{ij}(x)|.$$

By induction,

$$v_k(x) \le \frac{1}{(k-1)!} \int_a^x \left(\int_y^x G(s) ds \right)^{k-1} G(y) v_0(y)\, dy \quad \text{for all} \quad k \ge 1,\ x \in [a, b].$$

Thus $v_k(x) \le M^k \|u\|/k!$, where $M = \int_a^b G(s) ds$. Hence, if $\lambda \in \mathbb{C}$, $\lambda \ne 0$, then

$$|\lambda|^{-k} \|K^k u\| = |\lambda|^{-k} \|u^{(k)}\| \le \|u\| (M/|\lambda|)^k / k! \overset{k \to \infty}{\longrightarrow} 0.$$

Hence $|\lambda| > \|K^k\|^{1/k}$ for k large and, by Theorem 1.6.8, $\lambda \in \rho(K)$. Observe also that if $u = (u_1, \ldots, u_n)^\top \in \mathcal{R}(K)$, then each u_i is absolutely continuous and hence $\mathcal{R}(K) \ne X$ and $\sigma(K) = \{0\}$.

EXAMPLE 1.6.10 *The Volterra integral equation of the second kind* is represented by

$$f(x) - \int_a^x g(x, y) f(y)\, dy = h(x) \quad \text{for almost all} \quad x \in (a, b). \tag{1.23}$$

We shall show that, for given $h \in L^p(a, b)$, there exists a unique $f \in L^p(a, b)$ for which equation (1.23) is satisfied. It is assumed that $-\infty \le a < b \le \infty$, $1 < p < \infty$, $g : (a, b) \times (a, b) \to \mathbb{C}$ is Lebesgue measurable and

$$\int_a^b \left(\int_a^x |g(x, y)|^{\frac{p}{p-1}} dy \right)^{p-1} dx < \infty.$$

When $u \in L^p(a, b)$, Hölder's inequality implies that

$$\left(\int_a^x |g(x, y) u(y)| dy \right)^p \le G(x) \int_a^x |u(y)|^p dy \le G(x) \|u\|_p^p, \tag{1.24}$$

where

$$G(x) = \left(\int_a^x |g(x, y)|^{\frac{p}{p-1}} dy \right)^{p-1}.$$

Since $M \equiv \int_a^b G(x) dx < \infty$, one can define $K \in \mathfrak{B}(L^p(a, b))$ by

$$(Ku)(x) = \int_a^x g(x, y) u(y)\, dy.$$

Fix any $u \in L^p(a, b)$ and define $v_n(x) = \int_a^x |(K^n u)(y)|^p dy$ for $n = 0, 1, 2, \ldots$. Inequality (1.24) implies that

$$v_{n+1}(x) \le \int_a^x G(s) v_n(s)\, ds$$

and, as in Example 1.6.9, we have that $v_n(x) \leq M^n v_0(b)/n!$ for $n \geq 1$. Therefore $\|K^n u\|_p = v_n(b)^{1/p} \leq (M^n/n!)^{1/p}\|u\|_p$. Theorem 1.6.8 implies that every $\lambda \in \mathbb{C}$, $\lambda \neq 0$, belongs to $\rho(K)$. Thus the solution of the Volterra equation (1.23) is $f = (1-K)^{-1}h$ and equation (1.22) implies that

$$\lim_{n\to\infty} \left\| f - \sum_{k=0}^{n} K^k h \right\|_p = 0.$$

Theorem 1.6.11 *Suppose that T is a linear operator defined in a Banach space X. Then $\rho(T)$ is an open set and $\sigma(T)$ is a closed set. Moreover, if $\lambda \in \rho(T)$, $\zeta \in \mathbb{K}$, are such that $|\lambda - \zeta|\|(T-\lambda)^{-1}\| < 1$ then, $\zeta \in \rho(T)$. Furthermore, the resolvent of T is differentiable, $\mathfrak{B}(X)$ valued, function and its derivative is*

$$\frac{d}{d\lambda}(T-\lambda)^{-1} = ((T-\lambda)^{-1})^2 \quad \textit{for all} \quad \lambda \in \rho(T).$$

PROOF Observe that the 'moreover' part implies that $\rho(T)$ is an open set and that $\sigma(T) = \rho(T)^c$ is a closed set. To prove the 'moreover' part, use Theorem 1.6.8, to define

$$R = (T-\lambda)^{-1}(1 - (\zeta-\lambda)(T-\lambda)^{-1})^{-1} \in \mathfrak{B}(X).$$

It is easy to see that $(T-\zeta)Rx = x$ for $x \in X$ and $R(T-\zeta)y = y$ for $y \in \mathcal{D}(T)$. Hence $\zeta \in \rho(T)$, $(T-\zeta)^{-1} = R$. (1.22) gives the bound on $\|R\|$,

$$\|(T-\zeta)^{-1}\| \leq \frac{\|(T-\lambda)^{-1}\|}{1 - |\zeta-\lambda|\|(T-\lambda)^{-1}\|}.$$

If $\zeta \neq \lambda$, the resolvent identity (1.20) implies

$$\frac{1}{\zeta-\lambda}((T-\zeta)^{-1} - (T-\lambda)^{-1}) - ((T-\lambda)^{-1})^2 = (T-\lambda)^{-1}((T-\zeta)^{-1} - (T-\lambda)^{-1})$$

$$= (\zeta-\lambda)(T-\zeta)^{-1}((T-\lambda)^{-1})^2.$$

Using the bound on $\|R\|$ gives

$$\left\|\frac{1}{\zeta-\lambda}((T-\zeta)^{-1} - (T-\lambda)^{-1}) - ((T-\lambda)^{-1})^2\right\| \leq \frac{|\lambda-\zeta|\|(T-\lambda)^{-1}\|^3}{1 - |\zeta-\lambda|\|(T-\lambda)^{-1}\|}$$

and this proves the differentiability of $(T-\lambda)^{-1}$. □

A function $f : D \to X$, where D is a nonempty open set in $\mathbb{C}$ and X is a complex Banach space, is said to be **analytic** in D if it is differentiable in D, i.e., there exists $f' : D \to X$ such that

$$\lim_{w\to z} \left\| \frac{f(w) - f(z)}{w - z} - f'(z) \right\| = 0 \quad \text{for all} \quad z \in D.$$

Theorem 1.6.11 implies that the resolvent of an operator is analytic when its resolvent set is not empty and the scalar field is $\mathbb{C}$.

EXAMPLE 1.6.12 Define $T_2 : \mathcal{D}(T_2) \subset L^p(\mathbb{R}) \to L^p(\mathbb{R})$, $p \in [1, \infty]$, by $T_2 f = f'$ with domain

$$\mathcal{D}(T_2) = \{f \in L^p(\mathbb{R}) \mid f \in AC[-a, a] \text{ for all } a \in (0, \infty),\ f' \in L^p(\mathbb{R})\}$$

and let $\mathbb{C}$ be the field of scalars. In Example 1.6.6, it is shown that $\sigma(T_2) = i\mathbb{R}$. Thus, if $\lambda \in \mathbb{C}$, $\operatorname{Re}\lambda \neq 0$, then taking $\zeta = i\operatorname{Im}\lambda$ in Theorem 1.6.11 gives $|\operatorname{Re}\lambda| \|(T_2 - \lambda)^{-1}\| \geq 1$. Hence (1.21) implies

$$\|(T_2 - \lambda)^{-1}\| = \frac{1}{|\operatorname{Re}\lambda|} \quad \text{for all} \quad \lambda \in \mathbb{C} \text{ with } \operatorname{Re}\lambda \neq 0. \tag{1.25}$$

Theorem 1.6.13 *Let T be a closed and densely defined linear operator in a Banach space X. Then $\sigma(T) = \sigma(T^\star)$ and*

$$((T - \lambda)^{-1})^\star = (T^\star - \lambda)^{-1} \quad \textit{for all} \quad \lambda \in \rho(T).$$

PROOF It is easy to see that $(T-\lambda)^\star = T^\star - \lambda$ for all scalars λ. Theorem 1.5.16 implies the rest. □

Lemma 1.6.14 *If A, B are linear operators in a Banach space such that A is an extension of B and $\rho(A) \cap \rho(B)$ is not empty, then $A = B$.*

PROOF If $x \in \mathcal{D}(A)$ and $\lambda \in \rho(A) \cap \rho(B)$ then $(A - \lambda)x = (B - \lambda)y$ for some $y \in \mathcal{D}(B)$. Hence $(A - \lambda)(x - y) = 0$ and, therefore, $x = y$, implying $A = B$. □

Definition 1.6.15 Let T be a linear operator in a Banach space X. A set of scalars $f(Tx)$ - where x ranges over all $x \in \mathcal{D}(T)$ with $\|x\| = 1$, and f is a normalized tangent functional corresponding to this x - is called a **numerical range** of T. Since elements of a Banach space can have several different normalized linear functionals, T could, in principle, have several different numerical ranges. T is said to be **accretive** if it has a numerical range that is contained in $\{z \in \mathbb{K} | \operatorname{Re} z \geq 0\}$, i.e., for each $x \in \mathcal{D}(T)$ with $\|x\| = 1$, it is required that there exists $f \in X^\star$ such that $f(x) = \|f\| = 1$ and $\operatorname{Re} f(Tx) \geq 0$. T is said to be **dissipative** if $-T$ is accretive.

Note that $\sigma_p(T)$ is contained in any numerical range of T: if $\lambda \in \sigma_p(T)$, then $Tx = \lambda x$ for some x with norm 1, hence, if f is a normalized tangent functional to x then $f(Tx) = \lambda$. A partial converse

Theorem 1.6.16 *Suppose that T is a linear operator in a Banach space X and that its numerical range is contained in a nonempty closed set Γ. Let $\mathcal{O}$ be a connected subset of Γ^c. If $\mathcal{R}(T - \lambda) = X$ for some $\lambda \in \mathcal{O}$, then $\mathcal{O} \subset \rho(T)$ and for every $\zeta \in \mathcal{O}$,*

$$\|(T - \zeta)^{-1}\| \leq 1/\operatorname{dist}(\zeta, \Gamma). \tag{1.26}$$

PROOF If $x \in \mathcal{D}(T)$, $x \neq 0$, then for every scalar μ,

$$\|Tx - \mu x\| \geq |f(Tx - \mu x)| = \|x\||f(Tx)/\|x\| - \mu| \geq \|x\|\text{dist}(\mu, \Gamma)$$

where $f \in X^\star$ is the normalized tangent functional to $\|x\|^{-1}x$. This implies (1.26) for all ζ in the open set $\mathcal{O}_1 \equiv \rho(T) \cap \Gamma^c$ and that $\lambda \in \mathcal{O}_1 \cap \mathcal{O}$.

Let $\mathcal{O}_2 = \sigma(T) \cap \Gamma^c$. Take any $x \in \mathcal{O}_2$ and $\zeta \in B(x, r)$, where $r = \text{dist}(x, \Gamma)/2$. Note that $\text{dist}(\zeta, \Gamma) > r$. If $\zeta \in \rho(T)$, then $\zeta \in \mathcal{O}_1$ and (1.26) implies

$$|x - \zeta|\|(T - \zeta)^{-1}\| \leq |x - \zeta|/\text{dist}(\zeta, \Gamma) < 1.$$

Hence Theorem 1.6.11 implies $x \in \rho(T)$, which is not true, and therefore $\zeta \in \sigma(T)$. Hence $\zeta \in \mathcal{O}_2$, proving that $\mathcal{O}_2$ is open.

Since $\mathcal{O}_1$, $\mathcal{O}_2$ are disjoint and $\mathcal{O} \subset \mathcal{O}_1 \cup \mathcal{O}_2$, the connectedness of $\mathcal{O}$ implies that $\mathcal{O}_2 \cap \mathcal{O}$ is empty and therefore $\mathcal{O} \subset \mathcal{O}_1$, which completes the proof. □

Corollary 1.6.17 *If T is an accretive linear operator in a Banach space X such that $\mathcal{R}(T - \lambda) = X$ for some scalar λ with $\text{Re}\,\lambda < 0$, then every scalar ζ with $\text{Re}\,\zeta < 0$ belongs to the resolvent set of T and $\|(T - \zeta)^{-1}\| \leq 1/|\text{Re}\,\zeta|$.*

EXAMPLE 1.6.18 Let $X = \{f \in C[0,1] \mid f(0) = f(1) = 0\}$ and let $\|\cdot\|$ be the sup norm on X. Assume $a \in C[0,1]$ and $a(x) \in (0, \infty)$ for $x \in [0,1]$. Define

$$Tf = -af'' \quad \text{for} \quad f \in \mathcal{D}(T) \equiv \{f \in C^2[0,1] \mid f, f'' \in X\}.$$

Let us show first that T is accretive. Choose any $f \in \mathcal{D}(T)$ and define $\Lambda \in X^\star$ by $\Lambda(g) = \overline{f(x_0)}g(x_0)$ where $x_0 \in (0,1)$ is chosen so that $h \equiv |f|^2$ attains maximum at x_0. Λ is a normalized tangent functional to f and

$$\text{Re}\,\Lambda(Tf) = a(x_0)(|f'(x_0)|^2 - h''(x_0)/2) \geq 0.$$

It is easy to show that $0 \in \rho(T)$ by showing that

$$(T^{-1}f)(x) = \int_0^x \frac{(1-x)y}{a(y)} f(y)dy + \int_x^1 \frac{(1-y)x}{a(y)} f(y)dy \quad \text{for} \quad f \in X,\ x \in [0,1].$$

Since the resolvent set is open (Theorem 1.6.11) we have that $\mathcal{R}(T - \lambda) = X$ when $|\lambda|$ is small. Hence, all assumptions of Corollary 1.6.17 are satisfied.

1.7 Compact Linear Operators

We shall now restrict our investigation of operators to a special class which have many of the characteristics of operators in finite dimensional spaces. Many integral operators and, hence, inverses of differential operators, will fall in this class. Let us begin by showing that all norms are equivalent on finite dimensional spaces.

Theorem 1.7.1 *A finite dimensional normed space is complete and, hence, is a Banach space. Moreover, if $\{u_1, \ldots, u_n\}$ is a basis for the normed space X and if $T : \mathbb{K}^n \to X$ is defined by*

$$T\alpha = \alpha_1 u_1 + \cdots + \alpha_n u_n \quad \textit{for all} \quad \alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n,$$

then T is one-to-one, onto, bounded and T^{-1} is bounded.

PROOF The definition of the basis gives that T is one-to-one and onto. $\|T\alpha\| \leq (\sum \|u_i\|^2)^{1/2}\|\alpha\|$ implies that T is bounded. If T^{-1} is not bounded, then there exists a bounded sequence $\{x_m\}$ in X such that $\|T^{-1}x_m\| \to \infty$ as $m \to \infty$. Let $\alpha^m = \|T^{-1}x_m\|^{-1}T^{-1}x_m$. Then $\|\alpha^m\| = 1$ and $T\alpha^m \to 0$ as $m \to \infty$. The Heine-Borel Theorem implies that $\{\alpha^m\}$ has a convergent subsequence and, hence, by renaming a subsequence, we may assume that $\{\alpha^m\}$ converges to some $\alpha \in \mathbb{K}^n$ with $\|\alpha\| = 1$. Since $T\alpha^m \to 0$ and $\|T\alpha - T\alpha^m\| \leq \|T\|\|\alpha - \alpha^m\|$, we have that $T\alpha = 0$, which contradicts the fact that T is one-to-one. Therefore, T^{-1} is bounded. If $\{y_k\}$ is a Cauchy sequence in X, then boundedness of T^{-1} implies that $\{T^{-1}y_k\}$ is a Cauchy sequence in $\mathbb{K}^n$, which is complete. Hence $\{T^{-1}y_k\}$ converges to some $\beta \in \mathbb{K}^n$. Boundedness of T implies that $\{y_k\}$ converges to $T\beta$. □

Recall that a set is compact in a euclidean space if and only if it is closed and bounded. In infinite dimensional spaces the situation is very different:

Theorem 1.7.2 *If a normed space X contains a nonempty open subset V such that $V \subset K$ for some compact set K, then X is a finite dimensional space.*

PROOF Choose $x_0 \in V$, $r > 0$ such that $B(x_0, r) \subset K$. Compactness of K implies that there exist $x_1, \ldots, x_m$ in K such that $K \subset B(x_1, r/2) \cup \cdots \cup B(x_m, r/2)$. Let Y denote the set of all linear combinations of $x_1 - x_0, \ldots, x_m - x_0$. So, Y is a finite dimensional subspace of X, $\dim Y \leq m$.

If $x \in B(0,1)$, then $x_0 + rx \in B(x_0, r)$. Hence $x_0 + rx \in B(x_i, r/2)$ for some i and, therefore, $x_0 + rx = x_i + ry$, i.e., $x = (1/r)(x_i - x_0) + y$ for some $y \in B(0, 1/2)$. Therefore,

$$B(0,1) \subset Y + B(0, 1/2)$$

and, since Y is a vector space, we have $B(0, 2^{-n}) \subset Y + B(0, 2^{-n-1})$. This implies that if $B(0,1) \subset Y + B(0, 2^{-n})$, then

$$B(0,1) \subset Y + B(0, 2^{-n}) \subset Y + Y + B(0, 2^{-n-1}) = Y + B(0, 2^{-n-1}).$$

Therefore, $B(0,1) \subset Y + B(0, 2^{-n})$ for all $n \geq 1$. Thus, if $x \in B(0,1)$, then there exists a sequence $\{y_n\}$ in Y such that $\|x - y_n\| < 2^{-n}$. Theorem 1.7.1

implies that Y is complete and, therefore, $x \in Y$. Rescaling gives that $X = Y$. □

Let X and Y be Banach spaces. $T \in \mathfrak{B}(X,Y)$ is said to be **compact** if for every bounded sequence $\{x_n\}_{n=1}^{\infty}$ in X there exist integers $n_1 < n_2 < \cdots$ such that the sequence $\{Tx_{n_k}\}_{k=1}^{\infty}$ converges to some element of Y.

EXAMPLE 1.7.3 Suppose $t \in C_0(\mathbb{R}^{n+m})$ and define $T \in \mathfrak{B}(L^2(\mathbb{R}^m), L^2(\mathbb{R}^n))$ by

$$(Tu)(x) = \int_{\mathbb{R}^m} t(x,y)u(y)\,dy \quad \text{for all} \quad u \in L^2(\mathbb{R}^m),\ x \in \mathbb{R}^n.$$

We will show that T is compact.

Since t has compact support, there exists $c < \infty$ such that $t(x,y) = 0$ if $|x| > c$ or $|y| > c$. Let $\{u_k\}$ be a sequence in $L^2(\mathbb{R}^m)$ with $M = \sup_k \|u_k\| < \infty$. Note

$$|(Tu_k)(x)| \le \left(\int_{|y|\le c} |t(x,y)|^2 dy\right)^{1/2} \|u_k\| \le \|t\|_\infty (2c)^{m/2} M.$$

Let any $\varepsilon > 0$ be given. Uniform continuity of t on the compact set $|x| \le c+1$, $|y| \le c$ implies that there exists $\delta \in (0,1)$ such that $|t(x,y) - t(x',y)| < \varepsilon$ for all y whenever $|x - x'| < \delta$. Hence, if $|x - x'| < \delta$, then

$$|(Tu_k)(x) - (Tu_k)(x')| \le \left(\int_{|y|\le c} |t(x,y) - t(x',y)|^2 dy\right)^{1/2} \|u_k\| \le \varepsilon (2c)^{m/2} M.$$

Therefore, $\{Tu_k\}$ is a uniformly bounded and equicontinuous family of functions that vanish outside B, the closed ball of radius c in $\mathbb{R}^n$. The Arzela-Ascoli Theorem 1.1.5 implies that there exist $k_1 < k_2 < \ldots$ such that the subsequence $\{Tu_{k_j}\}_{j=1}^{\infty}$ converges uniformly in B to some $v \in C(B)$. Define $v = 0$ outside B. Hence $\{Tu_{k_j}\}_{j=1}^{\infty}$ converges to v in $L^2(\mathbb{R}^n)$.

Theorem 1.7.4 *Let X and Y be Banach spaces and $T \in \mathfrak{B}(X,Y)$.*

(a) T is compact if and only if the closure of $\{Tx | x \in X, \|x\| \le 1\}$ is compact.

(b) If $\dim \mathcal{R}(T) < \infty$, then T is compact.

(c) If T is compact and $\mathcal{R}(T)$ is closed, then $\dim \mathcal{R}(T) < \infty$.

(d) If $T_n \in \mathfrak{B}(X,Y)$ are compact for $n \ge 1$ and $\lim_{n\to\infty} \|T_n - T\| = 0$, then T is compact. Compact operators form a closed subspace of $\mathfrak{B}(X,Y)$.

(e) If Z is a Banach space and $S \in \mathfrak{B}(Y,Z)$, then ST is compact if either T or S is compact.

PROOF If T is compact and $y_n \in \overline{A}$, $A = \{Tx | x \in X, \|x\| \leq 1\}$, then there are $x_n \in X$ with $\|x_n\| \leq 1$ such that $\|y_n - Tx_n\| < 1/n$ and, since a subsequence of $\{Tx_n\}$ converges, the same subsequence of $\{y_n\}$ converges. Therefore $\overline{A}$ is compact. The rest of (a) is obvious. (b) follows from Theorem 1.7.1.

To prove (c), suppose $y_n \in \mathcal{R}(T)$ and $\|y_n\| \leq 1$ for $n \geq 1$. Since T maps onto $\mathcal{R}(T)$, and since $\mathcal{R}(T)$ is complete because $\mathcal{R}(T)$ is closed and Y is complete, the Open Mapping Theorem 1.4.7 implies that $y_n = Tx_n$ for some bounded sequence $\{x_n\}$ in X. Thus $\{y_n\}$ has a convergent subsequence and the closed unit ball in $\mathcal{R}(T)$ is compact and $\dim \mathcal{R}(T) < \infty$ by Theorem 1.7.2.

To prove (d), let $\{x_n\}$ be a sequence in X with $M = \sup_n \|x_n\| < \infty$. Let A_1 denote an infinite set of integers such that the sequence $\{T_1 x_n\}_{n \in A_1}$ converges. For $k \geq 2$ let $A_k \subset A_{k-1}$ denote an infinite set of integers such that the sequence $\{T_k x_n\}_{n \in A_k}$ converges. Choose $n_1 \in A_1$ and $n_k \in A_k$, $n_k > n_{k-1}$ for $k \geq 2$. Choose $\varepsilon > 0$. Let k be such that $\|T - T_k\| M < \varepsilon/4$ and note that

$$\|Tx_{n_i} - Tx_{n_j}\| \leq \|(T-T_k)(x_{n_i} - x_{n_j})\| + \|T_k x_{n_i} - T_k x_{n_j}\| < \varepsilon/2 + \|T_k x_{n_i} - T_k x_{n_j}\|.$$

Since $\{T_k x_{n_i}\}_{i=1}^{\infty}$ converges, $\{Tx_{n_i}\}_{i=1}^{\infty}$ is a Cauchy sequence.

(e) is obvious. □

Corollary 1.7.5 *Suppose X is a Banach space and $T \in \mathfrak{B}(X)$ is compact.*

(i) If λ is a scalar, $\lambda \neq 0$, then $\dim \mathcal{N}(T - \lambda) < \infty$.

(ii) If $\dim X = \infty$, then $0 \in \sigma(T)$.

PROOF If $Y = \mathcal{N}(T - \lambda)$, then T restricted to Y has closed range Y. If $0 \in \rho(T)$, then $\mathcal{R}(T) = X$ is closed and, hence, has to be finite dimensional. □

EXAMPLE 1.7.6 Assume that $-\infty < a < b < \infty$ and $1 \leq p \leq \infty$. Define

$$(Tf)(x) = \int_a^x f(y)dy.$$

Note that $T \in \mathfrak{B}(L^1(a,b), L^p(a,b))$. We want to show that T is compact if $p < \infty$.

Fix $n \geq 1$, $\delta = (b-a)/n$ and define for $f \in L^1(a,b)$

$$(T_n f)(x) = \sum_{k=0}^{n-1} \chi_{[r_k, r_{k+1})}(x) \int_a^{r_k} f(s)ds,$$

where $r_k = a + k\delta$ and χ_A is used to denote the characteristic function of the set A. Note that T_n has an n-dimensional range and, hence, Theorem 1.7.4 implies that T_n is

a compact member of $\mathfrak{B}(L^1(a,b), L^p(a,b))$. If $r_k \le x < r_{k+1}$, then

$$\begin{aligned} |(Tf - T_n f)(x)| = \left| \int_{r_k}^{x} f(s)ds \right| &\le \int_{r_k}^{r_{k+1}} |f(s)|ds \le \|f\|_1 \\ \int_{r_k}^{r_{k+1}} |Tf - T_n f| &\le \delta \int_{r_k}^{r_{k+1}} |f| \\ \|Tf - T_n f\|_1 &\le \delta \|f\|_1 \\ \|Tf - T_n f\|_p &\le \delta^{1/p} \|f\|_1. \end{aligned}$$

Hence, (d) of Theorem 1.7.4 implies that T is compact if $p < \infty$.

Since multiplication by the L^∞ function is a bounded operator in $L^p(a,b)$, the operator

$$(Sf)(x) = k_1(x) \int_a^x k_2(y) f(y) dy,$$

with $k_1, k_2 \in L^\infty(a,b)$, belongs to $\mathfrak{B}(L^1(a,b), L^p(a,b))$ and is compact if $p < \infty$.

EXAMPLE 1.7.7 Suppose $M \subset \mathbb{R}^m$ and $N \subset \mathbb{R}^n$ are nonempty Lebesgue measurable sets and that $t : N \times M \to \mathbb{C}$ is measurable with

$$\int_N \int_M |t(x,y)|^2 dy\, dx < \infty.$$

If $u \in L^2(M)$, then

$$\left(\int_M |t(x,y)u(y)| dy \right)^2 \le \left(\int_M |t(x,y)|^2 dy \right) \int_M |u(y)|^2 dy.$$

Therefore, we can define $T \in \mathfrak{B}(L^2(M), L^2(N))$ by

$$(Tu)(x) = \int_M t(x,y)u(y)\, dy \quad \text{for all} \quad u \in L^2(M),\ x \in N \tag{1.27}$$

$$\|T\| \le \|t\|_{L^2(N \times M)} = \left(\int_N \int_M |t(x,y)|^2 dy\, dx \right)^{1/2}. \tag{1.28}$$

Let us show that T is compact.

Define $s = t$ in $N \times M$ and $s(x,y) = 0$ if $x \in \mathbb{R}^n \backslash N$ or $y \in \mathbb{R}^m \backslash M$. Using s in place of t in equation (1.27) defines $S \in \mathfrak{B}(L^2(\mathbb{R}^m), L^2(\mathbb{R}^n))$. Choose $t_k \in C_0(\mathbb{R}^{n+m})$ so that $t_k \to s$ in $L^2(\mathbb{R}^{n+m})$ and let $T_k \in \mathfrak{B}(L^2(\mathbb{R}^m), L^2(\mathbb{R}^n))$ be defined as T in (1.27) but with t_k in place of t. In Example 1.7.3, it is shown that T_k are compact. (1.28) implies $\|T_k - S\| \le \|t_k - s\|_{L^2(\mathbb{R}^{n+m})} \to 0$. Hence Theorem 1.7.4 implies that S is compact and so is T.

Theorem 1.7.8 *Suppose X and Y are Banach spaces and $T \in \mathfrak{B}(X,Y)$. Then T is compact if and only if $T^\star$ is compact.*

PROOF Suppose T is compact. Let $\{f_n\}$ be a bounded sequence in $Y^\star$ and let $M = \sup_n \|f_n\|$. Note that K, the closure of $\{Tx | x \in X, \|x\| \leq 1\}$, is compact and

$$|f_n(y)| \leq M\|T\| \quad \text{for all} \quad y \in K,\ n \geq 1,$$

$$|f_n(y) - f_n(z)| \leq M\|y - z\| \quad \text{for all} \quad y, z \in K,\ n \geq 1.$$

The Arzela-Ascoli Theorem 1.1.5 implies that a subsequence $\{f_{n_j}\}$ converges uniformly on K and, therefore,

$$\begin{aligned} \|T^\star f_{n_i} - T^\star f_{n_j}\| &= \sup_{\|x\| \leq 1} |f_{n_i}(Tx) - f_{n_j}(Tx)| \\ &\leq \sup_{y \in K} |f_{n_i}(y) - f_{n_j}(y)| \end{aligned}$$

implies that $\{T^\star f_{n_j}\}$ is a Cauchy sequence. Hence $T^\star$ is compact.

Assume $T^\star$ is compact. Let $F : X \to X^{\star\star}$ be the linear isometry given by $(Fx)(f) = f(x)$ for $x \in X$, $f \in X^\star$. If $\{x_n\}$ is a bounded sequence in X, then $\{Fx_n\}$ is a bounded sequence in $X^{\star\star}$ and, since $T^{\star\star}$ is compact, we have that a subsequence $\{T^{\star\star}Fx_{n_j}\}$ converges. Observe that for $x \in X$, $g \in Y^\star$, we have

$$g(Tx) = (T^\star g)(x) = (Fx)(T^\star g) = (T^{\star\star}Fx)(g)$$

$$\|T^{\star\star}Fx\| = \|Tx\|,$$

implying that $\{Tx_{n_j}\}$ converges. □

Lemma 1.7.9 *If N is a finite dimensional subspace of a normed space X, then there exists a closed subspace M of X such that*

$$X = M + N \quad \textit{and} \quad M \cap N = \{0\}.$$

PROOF Let $\{u_1, \ldots, u_n\}$ be a basis for N. Each $x \in N$ has a unique representation of the form

$$x = \lambda_1(x)u_1 + \cdots + \lambda_n(x)u_n, \quad \lambda_i(x) \in \mathbb{K}.$$

By Theorem 1.7.1, $\lambda_i \in N^\star$. The Hahn-Banach Theorem gives extensions $\Lambda_i \in X^\star$ of λ_i. Let $M = \cap_{i=1}^n \mathcal{N}(\Lambda_i)$. If $x \in X$ and $y = \Lambda_1(x)u_1 + \cdots + \Lambda_n(x)u_n \in N$, then $x - y \in M$ because $\Lambda_i(x) = \lambda_i(y) = \Lambda_i(y)$ for $1 \leq i \leq n$. □

Lemma 1.7.10 *If X is a Banach space and $T \in \mathfrak{B}(X)$ is compact, then $\mathcal{R}(T - \lambda)$ is closed for every scalar $\lambda \neq 0$.*

PROOF By Corollary 1.7.5, $N = \mathcal{N}(T-\lambda)$ is a finite dimensional subspace of X. Let M be a closed subspace of X as given by Lemma 1.7.9. Define $S \in \mathfrak{B}(M,X)$ by $Sx = Tx - \lambda x$. To show that $\mathcal{R}(T-\lambda) = \mathcal{R}(S)$ is closed, it is sufficient to show that there exists $r > 0$ such that

$$r\|x\| \leq \|Sx\| \quad \text{for all} \quad x \in M$$

because this implies that if $\{Sx_n\}$ is a Cauchy sequence in $\mathcal{R}(S)$, then $\{x_n\}$ converges to some $x \in M$. Hence $\lim_{n\to\infty} Sx_n = Sx$.

Now, if there would be no such r, then for each $n \geq 1$ there should exist $x_n \in M$ such that $\|x_n\| = 1$, $\|Sx_n\| < 1/n$. Thus, for some subsequence, $Tx_{n_j} \to y \in X$. $\lambda x_{n_j} = Tx_{n_j} - Sx_{n_j} \to y$ implies $y \in M$, $\|y\| = |\lambda| > 0$ and $Ty = \lambda y$, which contradicts the fact that S is one-to-one. □

Lemma 1.7.11 *If M is a subspace of a normed space X and $\overline{M} \neq X$, then for each $\varepsilon > 0$ there exists $x_\varepsilon \in X$ such that*

$$\|x_\varepsilon\| = 1 \text{ and } \operatorname{dist}(x_\varepsilon, M) \geq 1 - \varepsilon.$$

PROOF Assume $\varepsilon \in (0,1)$. Choose $y \in X \backslash \overline{M}$ and note that $\alpha = \operatorname{dist}(y,M) > 0$. Choose $w \in M$ such that $\|y-w\| \leq \alpha/(1-\varepsilon)$. If $x_\varepsilon = \|y-w\|^{-1}(y-w)$, then

$$\|x_\varepsilon - z\| = \frac{1}{\|y-w\|}\| y - w - \|y-w\|z \| \geq \frac{\alpha}{\|y-w\|} \geq 1-\varepsilon$$

for $z \in M$. □

Theorem 1.7.12 *If X is a Banach space, $T \in \mathfrak{B}(X)$ and T is compact, then*

$$\sigma(T)\backslash\{0\} = \sigma_p(T)\backslash\{0\}.$$

PROOF Clearly, $\sigma_p(T)\backslash\{0\} \subset \sigma(T)\backslash\{0\}$. Assume $\lambda \in \sigma(T)\backslash\{0\}$ and $\lambda \notin \sigma_p(T)$.

Let $S = T - \lambda$, $X_0 = X$, $X_{n+1} = SX_n$ for $n \geq 0$. Since $X_n = S^n X$, we have that $X_{n+1} \subset X_n$. $\lambda \in \sigma(T)$ implies that $X_1 \neq X_0$. This, and the fact that S is one-to-one, implies that X_{n+1} is a proper subspace of X_n for $n \geq 0$. By Lemma 1.7.10, X_n are closed.

Lemma 1.7.11 implies that, for every $n \geq 0$, there exists $x_n \in X_n$ such that $\|x_n\| = 1$ and $\operatorname{dist}(x_n, X_{n+1}) \geq 1/2$. If $m > n \geq 0$, then

$$x_m - (1/\lambda)(Sx_n - Sx_m) \in X_{n+1}$$

$$\|Tx_n - Tx_m\| = |\lambda|\|x_n - x_m + (1/\lambda)(Sx_n - Sx_m)\| \geq |\lambda|/2.$$

This contradicts the fact that $\{Tx_n\}$ should have a convergent subsequence. □

Theorem 1.7.13 *Suppose X is a Banach space and $T \in \mathfrak{B}(X)$ is compact. Then,*

$$\dim \mathcal{N}(T-\lambda) = \dim \mathcal{N}(T^\star - \lambda) < \infty$$

for all scalars $\lambda \neq 0$.

PROOF Let $m = \dim \mathcal{N}(T-\lambda)$ and $n = \dim \mathcal{N}(T^\star - \lambda)$. It will be shown that $n \leq m$. This implies $\dim \mathcal{N}(T^{\star\star} - \lambda) \leq n$ and, since $\mathcal{N}(T-\lambda)$ can be identified with a subspace of $\mathcal{N}(T^{\star\star} - \lambda)$ (see Definition 1.5.11) we have $n = m$.

To get a contradiction, assume $n > m$. $n \geq 1$ implies $\lambda \in \sigma(T^\star) = \sigma(T)$ (see Theorem 1.6.13). Hence, Theorem 1.7.12 implies $\lambda \in \sigma_p(T)$ and, therefore, $m \geq 1$. Let $x_1, \ldots, x_m$ be a basis of $\mathcal{N}(T-\lambda)$. Each $x \in \mathcal{N}(T-\lambda)$ has a unique representation $x = \ell_1(x)x_1 + \cdots + \ell_m(x)x_m$. Let $f_i \in X^\star$ denote extensions of $\ell_i \in \mathcal{N}(T-\lambda)^\star$ (Theorem 1.7.1) as given by the Hahn-Banach Theorem.

Let $\psi_1, \ldots, \psi_n$ be a basis of $\mathcal{N}(T^\star - \lambda)$. Define $g_1 = \psi_1$ and choose $y_1 \in X$ such that $g_1(y_1) = 1$. If $y_1, \ldots, y_k$ in X and $g_1, \ldots, g_k$ in $X^\star$ are such that

$$g_i(y_i) = 1, \quad g_i(y_j) = 0 \text{ if } j \neq i \tag{1.29}$$

$$g_i \in \operatorname{span}\{\psi_1, \ldots, \psi_i\}$$

for $i = 1, \ldots, k$, for some $1 \leq k < n$, define

$$g_{k+1} = \psi_{k+1} - \sum_{i=1}^{k} \psi_{k+1}(y_i) g_i.$$

$g_{k+1}(y) = 1$ for some $y \in X$ because $\psi_1, \ldots, \psi_{k+1}$ are linearly independent. Define

$$y_{k+1} = y - \sum_{i=1}^{k} g_i(y) y_i$$

and verify that (1.29) holds for $i, j \leq k+1$. Thus, we may assume that $k = n$.

Define a compact $S \in \mathfrak{B}(X)$ by

$$Sx = Tx - \sum_{i=1}^{m} f_i(x) y_i.$$

If $Sx = \lambda x$, then for $1 \leq j \leq m$ we have

$$0 = g_j(\lambda x - Sx) = \lambda g_j(x) - g_j(Tx) + f_j(x) = \lambda g_j(x) - (T^\star g_j)(x) + f_j(x) = f_j(x).$$

Therefore $Sx = Tx = \lambda x$ and, also, $\ell_j(x) = 0$. Thus, $x = 0$ and $S - \lambda$ is one-to-one. Theorem 1.7.12 implies that $Sx - \lambda x = y_{m+1}$ for some x, giving the contradiction that

$$1 = g_{m+1}(y_{m+1}) = g_{m+1}(Tx) - \lambda g_{m+1}(x) = (T^\star g_{m+1})(x) - \lambda g_{m+1}(x) = 0.$$

□

The following Theorem and Theorem 1.7.12 are perhaps the most often quoted properties of compact operators.

Theorem 1.7.14 *If X is a Banach space, $T \in \mathfrak{B}(X)$ and T is compact, then, for each $r > 0$, there exist at most finitely many $\lambda \in \sigma_p(T)$ for which $|\lambda| > r$.*

PROOF Suppose that for some $r > 0$ there exists a sequence of distinct eigenvalues $\{\lambda_n\}$ of T such that $|\lambda_n| > r$ for all $n \geq 1$.

Choose $x_n \neq 0$ such that $Tx_n = \lambda_n x_n$, and let $X_n = \text{span}\{x_1, \cdots, x_n\}$. Since $\{x_1, \cdots, x_n\}$ are linearly independent, we have that X_{n-1} is a proper closed subspace of X_n. By Lemma 1.7.11, for each $n > 1$ there exists $y_n \in X_n$ such that $\|y_n\| = 1$ and $\text{dist}(y_n, X_{n-1}) \geq 1/2$. If $m > n > 1$, then

$$(T - \lambda_m)y_m - Ty_n \in X_{m-1}$$

$$\|Ty_m - Ty_n\| = |\lambda_m| \| y_m + \lambda_m^{-1}((T - \lambda_m)y_m - Ty_n) \| \geq |\lambda_m|/2 \geq r/2.$$

This contradicts the fact that $\{Ty_n\}$ should have a convergent subsequence. □

A linear operator T in a Banach space X is said to have **compact resolvent** if there exists λ_0 in the resolvent set of T such that $(T - \lambda_0)^{-1}$ is compact.

EXAMPLE 1.7.15 $T_1 : \mathcal{D}(T_1) \subset L^p(0,1) \to L^p(0,1)$, $p \in [1, \infty]$, given by

$$T_1 f = f' \quad \text{for} \quad f \in \mathcal{D}(T_1) = \{f \in AC[0,1] \mid f' \in L^p(0,1),\ f(0) = f(1)\}$$

has resolvent given in Example 1.6.4. It is easy to see that T_1 has compact resolvent (when $p = 1$, see Example 1.7.6).

Consider $T_2 : \mathcal{D}(T_2) \subset L^p(\mathbb{R}) \to L^p(\mathbb{R})$, $p \in [1, \infty]$, given by $T_2 f = f'$,

$$\mathcal{D}(T_2) = \{f \in L^p(\mathbb{R}) \mid f \in AC[-a, a] \text{ for all } a \in (0, \infty),\ f' \in L^p(\mathbb{R})\}.$$

In Example 1.6.6, it is shown that $\sigma(T_2) = i\mathbb{R}$. Hence (a) of Theorem 1.7.16 implies that it cannot have compact resolvent.

We shall see that differential operators typically have compact resolvents when defined in bounded domains; when defined in unbounded domains, they usually do not have compact resolvents. For an exception to this rule, see Hamiltonian associated with quadratic potential in Section 2.9.

Among the results of this section, the following Theorem will be used most often in the rest of this book.

Theorem 1.7.16 *Suppose that T is a linear operator in a Banach space X and that there exists λ_0 in the resolvent set of T such that $(T - \lambda_0)^{-1}$ is compact. Then*

(a) $\{\lambda \in \sigma(T) \mid |\lambda| < r\}$ *is a finite set for every* $r < \infty$

(b) $\sigma(T) = \sigma_p(T) = \{\lambda_0 + 1/\mu \mid \mu \neq 0,\ \mu \in \sigma_p((T-\lambda_0)^{-1})\}$

(c) $T - \lambda$ *has closed range for all scalars* λ

(d) $\dim \mathcal{N}(T-\lambda) < \infty$ *for all scalars* λ

(e) $(T-\lambda)^{-1}$ *is compact for each* $\lambda \in \rho(T)$.

Moreover, if T *is in addition densely defined, then for each scalar* λ *we have*

$$\dim \mathcal{N}(T-\lambda) = \dim \mathcal{N}(T^\star - \lambda).$$

PROOF Let $K = (T-\lambda_0)^{-1}$. The resolvent identity (1.20) can be written as

$$(T-\lambda)^{-1} = K(1 + (\lambda - \lambda_0)(T-\lambda)^{-1}) \quad \text{for} \quad \lambda \in \rho(T).$$

Hence Theorem 1.7.4 implies (e). It can be easily verified that

$$\mathcal{N}(T-\lambda) = \mathcal{N}((\lambda-\lambda_0)K - 1) \quad \text{for each scalar } \lambda$$

and, hence, Corollary 1.7.5 implies (d). If T is densely defined, then Theorem 1.6.13 implies

$$\mathcal{N}(T^\star - \lambda) = \mathcal{N}(((\lambda-\lambda_0)K)^\star - 1) \quad \text{for each scalar } \lambda.$$

Hence, Theorem 1.7.13 implies the 'moreover' part.

If $(T-\lambda)x_n \to y$, then applying K gives $(1-(\lambda-\lambda_0)K)x_n \to Ky$ and, by Lemma 1.7.10, $(1-(\lambda-\lambda_0)K)x = Ky$ for some x. Hence $(T-\lambda)x = y$ and (c) follows.

If $\lambda \notin S \equiv \{\lambda_0 + 1/\mu \mid \mu \neq 0,\ \mu \in \sigma_p(K)\}$, then $1 \notin \sigma_p((\lambda-\lambda_0)K)$ and, by Theorem 1.7.12, $1 \notin \sigma((\lambda-\lambda_0)K)$. Hence $K(1-(\lambda-\lambda_0)K)^{-1}$ belongs to $\mathfrak{B}(X)$ and it is easy to see that it is equal to $(T-\lambda)^{-1}$. Therefore, $\lambda \notin \sigma(T)$, implying that $\sigma(T) \subset S$. Clearly, $S \subset \sigma_p(T) \subset \sigma(T)$ and, hence, $S = \sigma_p(T) = \sigma(T)$.

If $\lambda \in \sigma(T)$, $|\lambda| < r$, then by (b), $\lambda = \lambda_0 + 1/\mu$ for some $\mu \in \sigma_p(K)$ with $|1/\mu| = |\lambda - \lambda_0| < r + |\lambda_0|$. By Theorem 1.7.14, there exist at most finitely many such μ. This proves (a). □

The following **Fredholm alternative** applies to operators with closed range (see Lemma 1.7.10 and Theorem 1.7.16).

Theorem 1.7.17 *Suppose* T *is a densely defined linear operator in a Banach space* X *and that the range of* T *is closed. Then, for each* $y \in X$, *the following two statements are equivalent:*

(a) There exists $x \in \mathcal{D}(T)$ such that $Tx = y$

(b) $f(y) = 0$ whenever $f \in \mathcal{N}(T^\star)$.

PROOF If (a) is not true, then, by Theorem 1.5.7, there exists $f \in X^\star$ such that $f(y) > 0$ and $f(z) = 0$ for all $z \in \mathcal{R}(T)$. $f(Tx) = 0$ for all $x \in \mathcal{D}(T)$ implies $f \in \mathcal{N}(T^\star)$ (see Definition 1.5.12) and, hence, (b) is not true.

If (a) holds and $f \in \mathcal{N}(T^\star)$, then $f(y) = f(Tx) = (T^\star f)(x) = 0$. □

1.8 Boundary Value Problems for Linear ODEs

Let $A(t)$ be $n \times n$ matrices for t in a bounded interval $[a, b]$ with entries $A_{ij} \in L^1(a, b)$. Matrices $A^*(t)$ are defined by $A^*{}_{ij} = \overline{A_{ji}}$. Using Example 1.6.9, we can construct $n \times n$ matrices X, Y of absolutely continuous functions on $[a, b]$ so that

$$X'(t) + A(t)X(t) = 0 \quad \text{for} \quad t \in [a, b] \text{ a.e.}, \quad X(a) = I$$

$$Y'(t) - A^*(t)Y(t) = 0 \quad \text{for} \quad t \in [a, b] \text{ a.e.}, \quad Y(a) = I.$$

Since $(X^*Y)' = 0$, we have that $X^*(t)Y(t) = X^*(a)Y(a) = I$ for all $t \in [a, b]$ which implies also that $X^{-1} = Y^*$. Note also if $f \in (L^1(a, b))^n$, $\alpha \in \mathbb{C}^n$ and

$$x(t) = X(t)\alpha + X(t) \int_a^t Y^*(s) f(s)\, ds, \tag{1.30}$$

then x is the unique solution of $x' + Ax = f$ a.e. on $[a, b]$, $x(a) = \alpha$.

Let us consider now a boundary value problem,

$$x'(t) + A(t)x(t) = 0 \quad \text{for} \quad t \in [a, b] \text{ a.e.}, \qquad M_1 x(a) + M_2 x(b) = 0, \tag{1.31}$$

where M_1, M_2 are $n \times n$ matrices.

Since $x = X\alpha$, with $\alpha \in \mathbb{C}^n$, is a solution of (1.31) iff $(M_1 + M_2 X(b))\alpha = 0$, we have that the number of linearly independent solutions of (1.31) is $\dim \mathcal{N}(M_1 + M_2 X(b))$. On the other hand, x, given by Equation (1.30), will satisfy the nonhomogeneous problem

$$x'(t) + A(t)x(t) = f(t) \quad \text{for} \quad t \in [a, b] \text{ a.e.}, \qquad M_1 x(a) + M_2 x(b) = \gamma \tag{1.32}$$

provided we can find α such that

$$(M_1 + M_2 X(b))\alpha + M_2 X(b) \int_a^b Y^*(s) f(s)\, ds = \gamma. \tag{1.33}$$

Thus, we proved that uniqueness implies existence for (1.32). In other words,

Theorem 1.8.1 *If problem (1.31) has only the trivial solution* 0, *then problem (1.32) has a unique solution for any* $f \in (L^1(a,b))^n$ *and any* $\gamma \in \mathbb{C}^n$.

Theorem 1.8.2 *Assume that problem (1.31) has only the trivial solution and that* $1 \le p \le \infty$. *The mapping of* $(f,\gamma) \in (L^1(a,b))^n \times \mathbb{C}^n$ *into the solution of problem (1.32),* $x \in (L^p(a,b))^n$ *is bounded, linear and, when* $p < \infty$, *is also compact.*

PROOF Using (1.30) and (1.33), we see that each component of x consists of a finite sum of terms of the form

$$k_1(t)\gamma_i, \quad k_1(t)\int_a^b k_2(s)f_i(s)ds, \quad k_1(t)\int_a^t k_2(s)f_i(s)ds \tag{1.34}$$

for some $k_1, k_2 \in C[a,b]$ that do not depend on f, γ. It is easy to see that the map is bounded and linear. If $\{(f^{(j)}, \gamma^{(j)})\}$ is a bounded sequence in $(L^1(a,b))^n \times \mathbb{C}^n$, then the terms of x_j that are like the first two terms of (1.34) have obviously uniformly convergent subsequences. It is shown in Example 1.7.6 that the terms of x_j, that are like the third term of (1.34), have convergent subsequences in $L^p(a,b)$ if $p < \infty$. □

To study the **adjoint problem** of problem (1.31),

$$y'(t) - A^*(t)y(t) = 0 \quad \text{for} \quad t \in [a,b] \text{ a.e.}, \qquad N_1 y(a) + N_2 y(b) = 0, \tag{1.35}$$

we assume in the rest of this Section that the dimension of the null space of the $n \times 2n$ matrix $[M_1, M_2]$ is n and that $n \times n$ matrices N_1, N_2 are constructed as follows. Choose $u^j, v^j \in \mathbb{C}^n$ such that $M_1 u^j + M_2 v^j = 0$ for $j = 1, \ldots, n$, and that the vectors

$$\begin{bmatrix} u^j \\ v^j \end{bmatrix} \quad j = 1, \ldots, n \qquad \text{are linearly independent.}$$

Define

$$N_1 = \begin{bmatrix} \overline{u_1^1} & \overline{u_2^1} & \cdots & \overline{u_n^1} \\ \cdots & \cdots & \cdots & \cdots \\ \overline{u_1^n} & \overline{u_2^n} & \cdots & \overline{u_n^n} \end{bmatrix}, \qquad N_2 = \begin{bmatrix} -\overline{v_1^1} & -\overline{v_2^1} & \cdots & -\overline{v_n^1} \\ \cdots & \cdots & \cdots & \cdots \\ -\overline{v_1^n} & -\overline{v_2^n} & \cdots & -\overline{v_n^n} \end{bmatrix}.$$

Using notation $(x,y) = x_1\overline{y_1} + \cdots + x_n\overline{y_n}$ for $x, y \in \mathbb{C}^n$, observe that

$$(N_1 y)_k = (y, u^k), \quad (N_2 y)_k = -(y, v^k) \quad \text{for} \quad y \in \mathbb{C}^n, k = 1, \ldots, n,$$

which implies that, for all $\alpha, y_1, y_2 \in \mathbb{C}^n$, we have

$$(\alpha, N_1 y_1 + N_2 y_2) = (x_1, y_1) - (x_2, y_2),$$

where $x_1 = \sum_{k=1}^n \alpha_k u^k$, $x_2 = \sum_{k=1}^n \alpha_k v^k$. Therefore

(a) If $M_1x_1 + M_2x_2 = 0$ and $N_1y_1 + N_2y_2 = 0$, then $(x_1, y_1) = (x_2, y_2)$.

(b) If $(x_1, y_1) = (x_2, y_2)$ whenever $M_1x_1 + M_2x_2 = 0$, then $N_1y_1 + N_2y_2 = 0$.

Theorem 1.8.3 *For a given $f \in (L^1(a,b))^n$, the problem*

$$x'(t) + A(t)x(t) = f(t) \quad \text{for} \quad t \in [a,b] \text{ a.e.}, \qquad M_1x(a) + M_2x(b) = 0 \tag{1.36}$$

has a solution if and only if

$$\int_a^b (f(s), y(s))\, ds = 0 \quad \text{for every solution } y \text{ of problem (1.35).} \tag{1.37}$$

PROOF Assume first (1.37). Equations (1.30) and (1.33) imply that problem (1.36) has a solution if for every $\beta \in \mathcal{N}(M_1{}^* + X^*(b)M_2{}^*)$ we have that

$$(\, M_2X(b) \int_a^b Y^*(s)f(s)\, ds,\, \beta\,) = 0.$$

Choose any $\beta \in \mathcal{N}(M_1{}^* + X^*(b)M_2{}^*)$. Since

$$(\, M_2X(b) \int_a^b Y^*(s)f(s)\, ds,\, \beta\,) = \int_a^b (\, f(s),\, Y(s)X^*(b)M_2{}^*\beta\,)\, ds,$$

it is enough to show that $y(t) = Y(t)X^*(b)M_2{}^*\beta$ is a solution of (1.35). So we need to show that $N_1y(a) + N_2y(b) = 0$. This can be done by using property (b) above: if $M_1x_1 + M_2x_2 = 0$ then

$$\begin{aligned}(x_1, y(a)) - (x_2, y(b)) &= (x_1, X^*(b)M_2{}^*\beta) - (x_2, M_2{}^*\beta) \\ &= -(x_1, M_1{}^*\beta) - (x_2, M_2{}^*\beta) = 0.\end{aligned}$$

To show the converse, let x be a solution of (1.36) and let y be a solution of (1.35). Then $(x,y)' = (f,y)$ and, by property (a) above,

$$0 = (x(b), y(b)) - (x(a), y(a)) = \int_a^b (f(s), y(s))\, ds.$$

□

Theorem 1.8.4 *Problem (1.31) has the same number of linearly independent solutions as its adjoint problem (1.35).*

PROOF Using the definition of N_1, N_2, one can easily show that

$$(N_1 + N_2Y(b))^*\beta = \sum_{i=1}^{n} \beta_i(u^i - Y^*(b)v^i) \quad \text{for} \quad \beta \in \mathbb{C}^n.$$

Now, for each $x \in \mathcal{N}(M_1 + M_2X(b))$, there exists a unique $\alpha \in \mathbb{C}^n$ such that

$$x = \sum_{i=1}^{n} \alpha_i u^i \quad \text{and} \quad X(b)x = \sum_{i=1}^{n} \alpha_i v^i.$$

Hence

$$0 = \sum_{i=1}^{n} \alpha_i(u^i - X(b)^{-1}v^i) = (N_1 + N_2Y(b))^*\alpha.$$

Therefore we can define a map $T : \mathcal{N}(M_1 + M_2X(b)) \to \mathcal{N}((N_1 + N_2Y(b))^*)$ by $Tx = \alpha$. One can easily check that T is linear, one-to-one and onto. Hence

$$\dim \mathcal{N}(M_1 + M_2X(b)) = \dim \mathcal{N}((N_1 + N_2Y(b))^*) = \dim \mathcal{N}(N_1 + N_2Y(b))$$

which proves the Theorem. □

EXAMPLE 1.8.5 Assume $-\infty < a < b < \infty$ and that $q, r^{-1}, f \in L^1(a,b)$. Consider the problem of finding $x \in AC[a,b]$ such that $rx' \in AC[a,b]$ and

$$\left.\begin{array}{c} (rx')' + qx = f \quad \text{a.e. on } [a,b] \\ \alpha_1\, x(a) + \alpha_2\, (rx')(a) = 0 \\ \beta_1\, x(b) + \beta_2\, (rx')(b) = 0 \end{array}\right\} \tag{1.38}$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{C}$ are such that $|\alpha_1| + |\alpha_2| \neq 0$ and $|\beta_1| + |\beta_2| \neq 0$.

If the homogeneous version (when $f = 0$) of problem (1.38) has only the trivial solution 0, then problem (1.38) has a unique solution by Theorem 1.8.1. In this case, Theorem 1.8.2 implies that if $\{f_k\}$ is a bounded sequence in $L^1(a,b)$, then the corresponding solutions of (1.38) are uniformly bounded and have a convergent subsequence in $L^p(a,b)$ if $p < \infty$.

It is an easy exercise to apply Theorem 1.8.3 giving that problem (1.38) has a solution iff

$$\int_a^b f(t)x(t)\,dt = 0,$$

for all solutions x of the homogeneous version of problem (1.38). This fact is also known as the Fredholm alternative.

1.9 Exercises

1. Let $M = (0, 1]$. For $x, y \in M$, define

$$d(x, y) = \min\{|x - y|, 1 - |x - y|\}.$$

Is d a metric on M? If so, is M complete with this metric?

2. Show that there exists a unique $u \in C([0, 1], \mathbb{R})$ such that

$$e^{-10x}u(x) + \int_0^x (x + y)^2 \sin u(y)\, dy = 1 \quad \text{for all} \quad x \in [0, 1].$$

3. Let Ω be a nonempty open set in $\mathbb{R}^n$. For $i \geq 1$, define

$$K_i = \{x \in \overline{B(0, i)} \mid |x - y| \geq 1/i \text{ for every } y \in \Omega^c\}.$$

Show that K_i is compact, $K_i \subset K_{i+1}$ for $i \geq 1$ and $\Omega = \cup_{i=1}^{\infty} K_i$. Choose $i_0 \in \mathbb{N}$ such that K_{i_0} is not empty. Define

$$p_i(f) = \sup_{x \in K_i} |f(x)| \quad \text{for} \quad f \in C(\Omega),\ i \geq i_0.$$

Show that p_i are seminorms on $C(\Omega)$ and that

$$d(f, g) = \sum_{i=i_0}^{\infty} \frac{2^{-i} p_i(f - g)}{1 + p_i(f - g)} \quad \text{for} \quad f, g \in C(\Omega)$$

defines a metric on $C(\Omega)$ which makes $C(\Omega)$ a complete metric space.

4. Assuming Hölder's inequality (1.9) with $r = 1$, show (1.9) and (1.10).

5. Show that the mapping $\sum_1^{\infty} a_n \to \sum_1^{\infty} \lambda_n a_n$, defined by the sequence $\{\lambda_n\}$, maps convergent series into convergent series if and only if $\sum_{n=1}^{\infty} |\lambda_{n+1} - \lambda_n| < \infty$.

6. Suppose $z \in \mathbb{C}$, $|z| < 1$, $p \in [1, \infty]$ $(Tx)_n = \sum_{k=1}^{n} z^{n-k} x_k$ for $n \geq 1$ and $x = \{x_k\} \in \ell^p$. Show that T is a bounded linear operator in ℓ^p.

7. Let T be a bounded linear operator on a Banach space X such that the sequence $x, Tx, T^2x, \ldots$ is bounded for every $x \in X$. Show that if $\lambda \in \sigma(T)$, then $|\lambda| \leq 1$.

8. Find a linear operator that is not closable.

9. Show that $\ell^q = \ell^{p\star}$ when $p, q \in (1, \infty)$, $1/p + 1/q = 1$.

10. Let X_1, X_2 be normed spaces and $\Lambda \in (X_1 \times X_2)^\star$. Show that there exist $\ell_1 \in X_1{}^\star$, $\ell_2 \in X_2{}^\star$ such that $\Lambda((x, y)) = \ell_1(x) + \ell_2(y)$ for all $x \in X_1$, $y \in X_2$.

11. Assume that x belongs to a normed space X and that for some $c \in [0, \infty)$ we have that $|\ell(x)| \leq c\|\ell\|$ for all $\ell \in X^\star$. Show that $\|x\| \leq c$.

12. Suppose that X is a normed space and $S \subset X$ is such that

$$\sup_{x \in S} |\ell(x)| < \infty \quad \text{for all} \quad \ell \in X^\star.$$

Such sets S are called **weakly bounded**. Show that S is bounded in X. (*Hint:* use problem 11 above and the uniform boundedness principle.)

13. Find a bounded sequence in $L^1(0,1)$ which has no weakly convergent subsequence.

14. Find a bounded sequence in $L^\infty(0,1)$ which has no weakly convergent subsequence.

15. Suppose that A and B are closed operators in a Banach space X, $\mathcal{D}(A) \supset \mathcal{D}(B)$ and that B is one-to-one and onto. Show that $AB^{-1} \in \mathfrak{B}(X)$.

16. Find a bounded linear operator in a complex Banach space whose spectrum is empty.

17. Prove that if T is a linear operator and $a \neq 0$, b are scalars, then

$$\sigma(aT + b) = a\sigma(T) + b.$$

18. Show that

$$\|(T - \lambda)^{-1}\| \geq \frac{1}{\operatorname{dist}(\lambda, \sigma(T))} \quad \text{for} \quad \lambda \in \rho(T)$$

for every linear operator T with a nonempty spectrum.

19. Let $Tu = u'$ for $u \in C_u^1(\mathbb{R})$. Show that both T and $-T$ are accretive in $C_u(\mathbb{R})$. Find the spectrum of T.

20. Suppose $g \in L^p(0,\ell)$, $1 \leq p \leq \infty$, $\ell \in (0,\infty)$, $u' \in AC[0,\ell]$, $u(0) = u(\ell) = 0$ and

$$u(x) - \lambda u''(x) = g(x) \quad \text{for} \quad x \in (0,\ell) \text{ a.e.}$$

where $\lambda \in (0,\infty)$. Show that $\|u\|_p \leq \|g\|_p$.

21. Let X be an infinite dimensional Banach space. Is it possible to find compact sets K_1, K_2, ... such that $X = \cup_{n=1}^\infty K_n$?

22. For $f \in L^2(0,\infty)$, define

$$(Tf)(x) = \frac{1}{x} \int_0^x f(s)ds \quad \text{for} \quad x \in (0,\infty).$$

Show that $T \in \mathfrak{B}(L^2(0,\infty))$ and that T is not compact.

23. Assume that $p \in [1, \infty]$ and $S : \mathcal{D}(S) \subset L^p(\mathbb{R}) \to L^p(\mathbb{R})$ is defined by $Sf = -f''$ with the domain

$$\mathcal{D}(S) = \{f \in C^1(\mathbb{R}) \mid f' \in AC[-a, a] \text{ for all } a \in (0, \infty);\ f, f', f'' \in L^p(\mathbb{R})\}.$$

Show that S does not have compact resolvent.

24. Fix $l \in (0, \infty)$ and $p \in [1, \infty]$. For $f \in L^p(0, l)$, define

$$(Sf)(x) = \frac{1}{\sqrt{\pi}} \int_0^x \frac{f(y)}{\sqrt{x - y}} dy \quad \text{for} \quad x \in [0, l].$$

Show that $S \in \mathfrak{B}(L^p(0, l))$ and $S^2 = T^{-1}$, where T is a linear operator in $L^p(0, l)$ defined by

$$Tf = f' \quad \text{for} \quad f \in \mathcal{D}(T) \equiv \{f \in AC[0, l] \mid f' \in L^p(0, l), f(0) = 0\}.$$

Show that, if $f \in \mathcal{D}(T)$ then $Sf \in \mathcal{D}(T)$ and $TSf = STf$. (*Hint:* $ST^{-1} = T^{-1}S$.) Show that if $f \in \mathcal{D}(T)$, then there exists a unique $g \in L^p(0, l)$ which solves **Abel's integral equation**:

$$f(x) = \frac{1}{\sqrt{\pi}} \int_0^x \frac{g(y)}{\sqrt{x - y}} dy \quad \text{for} \quad x \in [0, l],$$

moreover, $g = TSf = STf$. Find the spectrum of S.

25. Find a linear operator in a Banach space X which has compact resolvent and whose domain is not dense in X.

26. For what $f \in L^1(0, \pi)$ does there exist $u \in AC[0, \pi]$ such that $u' \in AC[0, \pi]$ and

$$u'' + 4u = f \quad \text{a.e. on } [0, \pi],$$

$$u(0) = u(\pi), \quad u'(0) = u'(\pi)?$$

27. Show that $\sigma(S) = [0, \infty)$ where S is as in Example 1.6.7.

Chapter 2

Linear Operators in Hilbert Spaces

2.1 Orthonormal Sets

Let H be a vector space with a scalar field $\mathbb{K}$ - which is either $\mathbb{R}$ or $\mathbb{C}$. H is said to be an **inner product space** if for every pair of x and y in H there corresponds $(x, y) \in \mathbb{K}$, called the **inner product** or **scalar product** of x and y, such that

(i) $(x, y) = \overline{(y, x)}$ for all $x, y \in H$

(ii) $(x + y, z) = (x, z) + (y, z)$ for all $x, y, z \in H$

(iii) $(\lambda x, y) = \lambda(x, y)$ for all $x, y \in H$, $\lambda \in \mathbb{K}$

(iv) $(x, x) \geq 0$ for all $x \in H$

(v) $(x, x) = 0$ if and only if $x = 0$.

If $(x, y) = 0$, then x is said to be **orthogonal** to y. For $M \subset H$, one defines $M^{\perp}$ to be the set of all $y \in H$ such that $(x, y) = 0$ for all $x \in M$. For $x \in H$ define also

$$\|x\| = (x, x)^{1/2}. \tag{2.1}$$

x is said to be **normalized** if $\|x\| = 1$. If $x \neq 0$, then 'normalized x' means $\|x\|^{-1}x$.

Theorem 2.1.1 *If H is an inner product space, then for all x and y in H*

(a) $|(x, y)| \leq \|x\|\|y\|$ *(Schwarz Inequality)*

(b) $\|x + y\| \leq \|x\| + \|y\|$.

PROOF (a) is obvious if $x = 0$; otherwise it follows by taking $\delta = -\overline{(x,y)}/\|x\|^2$ in

$$0 \le \|\delta x + y\|^2 = |\delta|^2\|x\|^2 + 2\mathrm{Re}\,(\delta x, y) + \|y\|^2.$$

This identity, with $\delta = 1$ and (a), implies (b). □

Thus, equation (2.1) defines a norm on an inner product space H. If this norm makes H complete, then H is said to be a **Hilbert space**.

EXAMPLE 2.1.2 The most typical Hilbert space is ℓ^2 - with the inner product defined by

$$(x,y) = \sum_{n=1}^{\infty} x_n\overline{y_n} \quad \text{for all} \quad x = \{x_n\}_{n=1}^{\infty},\ y = \{y_n\}_{n=1}^{\infty} \in \ell^2.$$

EXAMPLE 2.1.3 Let Ω be a nonempty (Lebesgue) measurable set in $\mathbb{R}^n$. $L^2(\Omega)$ is a Hilbert space with inner product

$$(f,g) = \int_\Omega f(x)\overline{g(x)}dx, \qquad f,g \in L^2(\Omega).$$

A subset E of a Hilbert space is said to be **orthonormal** if each element of E has norm 1 and if every pair of distinct elements in E is orthogonal. An orthonormal set is said to be **complete** if no non-zero element is orthogonal to every element in the set. A complete orthonormal set is often called an **orthonormal basis**.

If $\{y_1, \ldots, y_n\}$ are linearly independent in a Hilbert space, then one can obtain an orthonormal set $\{\psi_1, \ldots, \psi_n\}$ such that

$$\mathrm{span}\{\psi_1, \ldots, \psi_k\} = \mathrm{span}\{y_1, \ldots, y_k\} \quad \text{for } k = 1, \ldots, n$$

by defining $\psi_1 = y_1/\|y_1\|$ and

$$\psi_k = \frac{y_k - \sum_{j=1}^{k-1}(y_k, \psi_j)\psi_j}{\left\|y_k - \sum_{j=1}^{k-1}(y_k, \psi_j)\psi_j\right\|} \qquad \text{for } k = 2, \ldots, n.$$

This is known as the **Gram-Schmidt orthogonalization** procedure. If, for example, the Gram-Schmidt orthogonalization is applied in the Hilbert space $L^2(-1,1)$ to the polynomials $y_n(x) = x^n$, $n \ge 0$, then the corresponding orthonormal ψ_n are proportional to the **Legendre polynomials**.

Using the Hausdorff Maximality Theorem 1.5.5 it can be easily seen that any orthonormal set is contained in a complete orthonormal set (Exercise 1). We shall not need this fact. However, along the way we shall show completeness of many given orthonormal sets - which is quite a different story.

Lemma 2.1.4 *Let $\{\phi_i\}_{i=1}^{\infty}$ be an orthonormal set and let $\{\alpha_i\}_{i=1}^{\infty}$ be a sequence of scalars such that $\sum_{i=1}^{\infty}|\alpha_i|^2 < \infty$. Then $\sum_{i=1}^{\infty}\alpha_i\phi_i$ converges and does not depend on the order of summation. Moreover,*

$$\|\sum_{i=1}^{\infty}\alpha_i\phi_i\|^2 = \sum_{i=1}^{\infty}|\alpha_i|^2 \quad \textit{and} \quad (\sum_{i=1}^{\infty}\alpha_i\phi_i, \phi_j) = \alpha_j \textit{ for } j \geq 1.$$

PROOF If $x_n = \sum_{i=1}^{n}\alpha_i\phi_i$ and $m > n$, then

$$\|x_m - x_n\|^2 = \sum_{i=n+1}^{m}|\alpha_i|^2$$

and hence $\{x_i\}$ converges to some x. By definition, $\sum_{i=1}^{\infty}\alpha_i\phi_i = x$. Since $\|x_n\|^2 = \sum_{i=1}^{n}|\alpha_i|^2$, we have that $\|x\|^2 = \sum_{i=1}^{\infty}|\alpha_i|^2$. $(x, \phi_j) = \alpha_j$ follows from the observation $(x_n, \phi_j) = \alpha_j$ for $n \geq j \geq 1$.

Let k map the set of positive integers one-to-one and onto itself. If $z_n = \sum_{i=1}^{n}\alpha_{k(i)}\phi_{k(i)}$, then, as above, $\{z_i\}$ converges to some z, $(z_n, x) = \sum_{i=1}^{n}|\alpha_{k(i)}|^2$. Since

$$\|x - z\|^2 = \|x\|^2 - (x, z) - (z, x) + \|z\|^2$$

and each of the terms on the right hand side equals $\sum_{i=1}^{\infty}|\alpha_i|^2$, we have that $x = z$. □

Theorem 2.1.5 *Let E be an orthonormal subset of a Hilbert space H and $y \in H$. Then $(y, x) \neq 0$ for at most countably many $x \in E$. The series $\sum_{x\in E}(y, x)x$ converges and does not depend on the order of summation of non-zero terms,*

$$y - \sum_{x\in E}(y, x)x \in E^{\perp} \quad \textit{and} \quad \left\|y - \sum_{x\in E}(y, x)x\right\|^2 = \|y\|^2 - \sum_{x\in E}|(y, x)|^2. \tag{2.2}$$

PROOF If $x_1, \ldots, x_n$ are distinct elements of E, then

$$\left\|y - \sum_{i=1}^{n}(y, x_i)x_i\right\|^2 = \|y\|^2 - \sum_{i=1}^{n}|(y, x_i)|^2 \tag{2.3}$$

and, in particular,

$$\sum_{i=1}^{n}|(y, x_i)|^2 \leq \|y\|^2. \tag{2.4}$$

So, if x_i were chosen so that $|(y, x_i)| > \|y\|/m$ for some integer m, then (2.4) implies $n < m^2$ and therefore

$$E_y = \{x \in E \mid (y, x) \neq 0\} = \cup_{m\geq 1}\{x \in E \mid |(y, x)| > \|y\|/m\}$$

is at most a countable set. Thus, if $x_1, x_2, \ldots$ is an enumeration of E_y, then (2.4) and Lemma 2.1.4 imply that

$$z = \sum_{x \in E} (y, x)x = \sum_i (y, x_i)x_i$$

converges and is independent of the order of summation. Lemma 2.1.4 also implies $(z, w) = (y, w)$ for $w \in E$ - add w to the orthonormal set $x_1, x_2, \ldots$ Using convergence of the sum in equation (2.3) gives (2.2). □

Theorem 2.1.6 *If E is an orthonormal subset of a Hilbert space H, then the following statements are equivalent:*

(i) E is complete

(ii) $z = \sum_{x \in E}(z, x)x$ for all z in a dense subset of H

(iii) $y = \sum_{x \in E}(y, x)x$ for all $y \in H$

(iv) $\|y\|^2 = \sum_{x \in E} |(y, x)|^2$ for all $y \in H$ (Parseval equality).

PROOF Equivalence of (iii) and (iv) follows from equation (2.2). If $y \in H$, then, by Theorem 2.1.5, $y - \sum_{x \in E}(y, x)x$ is orthogonal to all $w \in E$, hence (i) implies (iii). (iii) obviously implies (i) and (ii). To see that (ii) implies (iii) define a linear map P on H by $Py = y - \sum_{x \in E}(y, x)x$; (2.2) implies $\|P\| \leq 1$ and hence for every $y \in H$ we have that $\|Py\| = \|P(y - z) + Pz\| \leq \|y - z\|$ for all z in the dense set and therefore $Py = 0$. □

EXAMPLE 2.1.7 Let

$$\phi_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx} \quad \text{for} \quad x \in \mathbb{R},\ n = 0, \pm 1, \pm 2, \ldots.$$

It is easy to check that ϕ_n form an orthonormal set in the Hilbert space $H = L^2(-\pi, \pi)$ with the usual inner product. We shall show now that the set is also complete. A different proof of completeness is given in Example 2.6.9.

Let us note first that (2.2) implies that for each $f \in L^2(-\pi, \pi)$ we have that

$$\sum_{n=-\infty}^{\infty} |(f, \phi_n)|^2 \leq \|f\|^2 \tag{2.5}$$

and this implies

$$\lim_{n \to \infty} \int_{-\pi}^{\pi} f(x) \cos nx \, dx = \lim_{n \to \infty} \int_{-\pi}^{\pi} f(x) \sin nx \, dx = 0. \tag{2.6}$$

Since $C_0(-\pi,\pi)$ is dense in $L^1(-\pi,\pi)$, see Example 1.3.4, (2.6) is true also for all $f \in L^1(-\pi,\pi)$.

If $f : \mathbb{R} \to \mathbb{C}$ is periodic with period 2π, $f \in L^1(-\pi,\pi)$, $x \in \mathbb{R}$ and there exist $f(x+), f(x-) \in \mathbb{C}$ and $\sigma, \varepsilon, M \in (0,\infty)$ such that

$$|f(x+h) - f(x+)| \le Mh^\sigma \quad \textit{and} \quad |f(x-h) - f(x-)| \le Mh^\sigma \quad \textit{for} \quad h \in (0,\varepsilon),$$

then

$$\lim_{n\to\infty} \sum_{k=-n}^{n} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \mathrm{e}^{-iks} ds \right) \mathrm{e}^{ikx} = \frac{f(x+) + f(x-)}{2}. \tag{2.7}$$

To prove (2.7) define

$$D_n(s) = \frac{1}{2\pi} \sum_{k=-n}^{n} \mathrm{e}^{iks} = \frac{\sin(2n+1)\frac{s}{2}}{2\pi \sin\frac{s}{2}}, \qquad s/2\pi \notin \mathbb{Z}$$

$$P_n(x) = \sum_{k=-n}^{n} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \mathrm{e}^{-iks} ds \right) \mathrm{e}^{ikx} = \int_{-\pi}^{\pi} f(s) D_n(x-s) ds.$$

Periodicity of f and D_n and the fact that $\int_0^\pi D_n(s) ds = 1/2$ imply

$$\begin{aligned} P_n(x) &= \int_{x-\pi}^{x+\pi} f(s) D_n(x-s) ds \\ &= \int_0^\pi f(x+s) D_n(s) ds + \int_0^\pi f(x-s) D_n(s) ds \\ &= \frac{f(x+) + f(x-)}{2} + I_n^+(x) + I_n^-(x) \end{aligned}$$

where

$$I_n^\pm(x) = \int_0^\pi (f(x \pm s) - f(x\pm)) D_n(s) ds = \int_0^\pi \frac{f(x \pm s) - f(x\pm)}{2\pi \sin\frac{s}{2}} \sin\tfrac{(2n+1)s}{2} ds.$$

(2.6) implies $\lim_{n\to\infty} I_n^\pm(x) = 0$ and hence (2.7) follows.

If $f \in C_0^1(-\pi,\pi)$, then (2.7) implies

$$\lim_{n\to\infty} \sum_{k=-n}^{n} (f, \phi_k)\phi_k(x) = \lim_{n\to\infty} P_n(x) = f(x) \quad \text{for all} \quad x \in (-\pi,\pi)$$

and since, by Theorem 2.1.5, P_n converge in H, we have that

$$\sum_{k=-\infty}^{\infty} (f, \phi_k)\phi_k = f. \tag{2.8}$$

Since $C_0^1(-\pi,\pi)$ is dense in $L^2(-\pi,\pi)$, Theorem 2.1.6 implies that (2.8) is true for all $f \in L^2(-\pi,\pi)$. This proves completeness of the set $\{\phi_k\}_{k=-\infty}^{\infty}$ in $L^2(-\pi,\pi)$.

The series $\sum_{k=-\infty}^{\infty}(f,\phi_k)\phi_k$ is called the **Fourier series** of f. Its partial sums P_n can be written also as

$$P_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos kx + b_k \sin kx \tag{2.9}$$

where

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx\,dx, \qquad b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kx\,dx. \tag{2.10}$$

Using these a_k, b_k, the Parseval equality can be written as

$$\frac{|a_0|^2}{2} + \sum_{k=1}^{\infty}|a_k|^2 + |b_k|^2 = \frac{1}{\pi}\int_{-\pi}^{\pi}|f(x)|^2 dx \quad \text{for} \quad f \in L^2(-\pi,\pi). \tag{2.11}$$

EXAMPLE 2.1.8 If $f \in L^2(0,\pi)$, then

$$f(x) = \sum_{k=1}^{\infty} b_k \sin kx \qquad \text{(series converges in } L^2(0,\pi)) \tag{2.12}$$

where

$$b_k = \frac{2}{\pi}\int_0^{\pi} f(x)\sin kx\,dx, \qquad k = 1,2,\ldots. \tag{2.13}$$

To see this define $f(x) = -f(-x)$ for $x \in (-\pi,0)$, hence $a_k = 0$ in Example 2.1.7 and therefore (2.9) and (2.8) imply (2.12). The series in (2.12) is called a **Fourier sine series**. Note that (2.12) and Theorem 2.1.6 imply that

$$\sqrt{\frac{2}{\pi}}\sin kx, \qquad k = 1,2,\ldots$$

form a complete orthonormal set in $L^2(0,\pi)$.

If you make an even extension of $f \in L^2(0,\pi)$, then Example 2.1.7 implies

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos kx \qquad \text{(series converges in } L^2(0,\pi)) \tag{2.14}$$

where

$$a_k = \frac{2}{\pi}\int_0^{\pi} f(x)\cos kx\,dx, \qquad k = 0,1,2,\ldots. \tag{2.15}$$

The series in (2.14) is called a **Fourier cosine series**.

EXAMPLE 2.1.9 Čebyšev polynomials T_n are defined as follows:

$$T_0(x) = 1,\ T_1(x) = x,$$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \quad \text{for} \quad n \geq 1.$$

By induction, one can easily show that

$$T_n(\cos y) = \cos(ny) \quad \text{for} \quad n \geq 0,\ -\infty < y < \infty$$

and, in particular, $|T_n(x)| \leq 1$ for $n \geq 0$, $-1 \leq x \leq 1$.

Suppose $g \in C^1[-1,1]$. Let $f(y) = g(\cos(y))$ and note that $f \in C^1(\mathbb{R})$ and f has period 2π. Since f is even, its Fourier series becomes a Fourier cosine series, hence (2.7) and (2.9) imply that

$$g(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n T_n(x) \quad \text{for} \quad -1 \leq x \leq 1 \tag{2.16}$$

where

$$a_n = \frac{2}{\pi} \int_{-1}^{1} \frac{g(x)T_n(x)}{\sqrt{1-x^2}} dx = \frac{2}{\pi} \int_0^{\pi} f(y) \cos(ny)\, dy \quad \text{for} \quad n \geq 0.$$

Integration by parts implies that

$$-na_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f'(y) \sin ny\, dy.$$

Hence (2.11) implies

$$\sum_{n=1}^{\infty} |na_n|^2 = \frac{1}{\pi} \int_{-\pi}^{\pi} |f'(y)|^2 dy$$

which implies $\sum_{n=0}^{\infty} |a_n| < \infty$ and therefore the series in (2.16) converges uniformly.

One handy consequence of the uniform convergence in (2.16) is the following. *If $h \in C^1(0,1)$ is bounded and $h(x) \neq 0$ for $0 < x < 1$, then the set of finite linear combinations of*

$$h(x), \quad h(x)x, \quad h(x)x^2, \quad \cdots$$

is dense in $L^p(0,1)$ for every $p \in [1,\infty)$. To see this, pick $u \in L^p(0,1)$ and $\varepsilon > 0$. Pick $v \in C_0^1(0,1)$ such that $\|u - v\|_p < \varepsilon/(1 + \|h\|_\infty)$ (see Example 1.3.4). Define $g \in C^1[-1,1]$ by $g = v/h$ in $[0,1]$ and $g = 0$ in $[-1,0]$. Uniform convergence in (2.16) implies the existence of a polynomial Q such that $\|g - Q\|_\infty < \varepsilon/(1 + \|h\|_\infty)$, hence,

$$\|u - hQ\|_p = \|u - v + h(v/h - Q)\|_p \leq \|u - v\|_p + \|h\|_\infty \|v/h - Q\|_\infty < \varepsilon.$$

Theorem 2.1.10 *If I is an interval in $\mathbb{R}$ and $\{\phi_1, \phi_2, \ldots\}$ is a complete orthonormal set in $L^2(I)$, then $\phi_{k_1}(x_1)\phi_{k_2}(x_2)\cdots\phi_{k_n}(x_n)$ with $k_i \geq 1$ form a complete orthonormal set in $L^2(I^n)$ for any $n \geq 1$.*

PROOF True if $n = 1$, let us assume that the Theorem is true for some value of $n \geq 1$ and suppose that $f \in L^2(I^{n+1})$ is such that

$$\int_I \int_{I^n} f(x,t)u(x)\phi_k(t)\, dx\, dt = 0$$

for all $k \geq 1$ and all u in the form $u(x) = u(x_1, \ldots, x_n) = \phi_{k_1}(x_1)\cdots\phi_{k_n}(x_n)$ with $k_i \geq 1$. Note that $f(\cdot,t) \in L^2(I^n)$ for almost all $t \in I$; hence $g_u(t) \equiv \int f(x,t)u(x)dx$ exists for these t and $g_u \in L^2(I)$. Completeness of $\{\phi_k\}$ implies $g_u = 0$ a.e. in I and the exceptional set may be chosen to be independent of u

since there are only countable many of them. The induction assumption implies that $f(\cdot,t)=0$ in $L^2(I^n)$ for almost all t. Therefore, $f=0$ a.e. in I^{n+1} and the Theorem is true for $n+1$ and hence for all n. □

EXAMPLE 2.1.11 In view of Example 2.1.7 and Theorem 2.1.10, we have that

$$e^{2\pi i k\cdot x}=e^{2\pi i(k_1x_1+\cdots+k_nx_n)}\quad(k\in\mathbb{Z}^n)$$

form a complete orthonormal set in $L^2((0,1)^n)$ for any $n\geq 1$.

2.2 Adjoints

Theorem 2.2.1 *Every nonempty closed convex subset of a Hilbert space contains a unique element of minimal norm.*

PROOF Choose x_n in the subset M so that $\|x_n\|\to d\equiv\inf\{\|x\|\,|\,x\in M\}$. Since $(1/2)(x_n+x_m)\in M$, we have $\|x_n+x_m\|^2\geq 4d^2$. Hence, the parallelogram law

$$\|x+y\|^2+\|x-y\|^2=2\|x\|^2+2\|y\|^2 \tag{2.17}$$

implies

$$\|x_n-x_m\|^2\leq 2(\|x_n\|^2-d^2)+2(\|x_m\|^2-d^2) \tag{2.18}$$

and therefore $\{x_n\}$ is a Cauchy sequence in the complete space. Since M is closed, $\{x_n\}$ converges to some $x\in M$ and $\|x\|=d$. If $y\in M$ and $\|y\|=d$, then the parallelogram law implies, as in (2.18), that $x=y$. □

Theorem 2.2.2 *If M is a closed subspace of a Hilbert space H, then $(M^\perp)^\perp=M$ and for each $x\in H$ there exist unique $y\in M$, $z\in M^\perp$ such that $x=y+z$.*

PROOF Let $E=\{x-y|y\in M\}$. It is easy to see that E is convex and closed. Theorem 2.2.1 implies that there exists $y\in M$ such that $\|x-y\|\leq\|x-w\|$ for all $w\in M$. Let $z=x-y$. Choose $w\in M$, $w\neq 0$, let $\alpha=(z,w)/\|w\|^2$ and note that

$$\|z\|^2\leq\|z-\alpha w\|^2=\|z\|^2-|(z,w)/\|w\||^2$$

implies $(z,w)=0$. Therefore $z\in M^\perp$. If $x=y'+z'$ for some $y'\in M$, $z'\in M^\perp$, then $y'-y=z-z'$. Hence $\|y-y'\|^2=(y'-y,z-z')=0$ implies uniqueness.

Obviously, $M\subset(M^\perp)^\perp$. If $x\in(M^\perp)^\perp$ and y and z are as above, then $x-y\in(M^\perp)^\perp$ and $x-y=z\in M^\perp$. Hence $x=y$ and $(M^\perp)^\perp=M$. □

The above Theorem implies that if M is not a dense subspace of a Hilbert space H, then $\overline{M}^\perp\neq\{0\}$, hence, there exists $x\in H$ such that $x\neq 0$ and $(y,x)=0$ for all $y\in M$. This argument is often used to show that a given subspace is dense.

Theorem 2.2.3 *Let M be a closed subspace of a Hilbert space H. Then there exist $P, Q \in \mathfrak{B}(H)$ such that for all $x \in H$,*

$$x = Px + Qx,\ Px \in M,\ Qx \in M^{\perp} \tag{2.19}$$

$$\|x\|^2 = \|Px\|^2 + \|Qx\|^2.$$

Moreover, $PH = M$, $QH = M^{\perp}$, $P^2 = P$, $Q^2 = Q$, $\|P\| \leq 1$, $\|Q\| \leq 1$.

PROOF Theorem 2.2.2 allows us to define P, Q with equation (2.19). Other conclusions follow straightforwardly. □

The operator P, as given in Theorem 2.2.3, is called the **orthogonal projection** of H onto M.

Lemma 2.2.4 (Riesz) *If H is a Hilbert space and $f \in H^{\star}$, then there exists a unique $y \in H$ such that*

$$f(x) = (x, y) \quad \textit{for all} \quad x \in H.$$

PROOF If $f = 0$ take $y = 0$. Assume $f \neq 0$. Hence, there exists w such that $f(w) \neq 0$ and, since $\mathcal{N}(f)$ is closed, Theorem 2.2.2 implies that $w = y_0 + y_1$ for some $y_0 \in \mathcal{N}(f)$, $y_1 \in \mathcal{N}(f)^{\perp}$ and $y_1 \neq 0$. If $x \in H$, then $f(w') = 0$ where $w' = f(x)y_1 - f(y_1)x$; hence $(w', y_1) = 0$ and this implies $f(x) = (x, y)$ where $y = (\overline{f(y_1)}/\|y_1\|^2)y_1$. If $f(x) = (x, y')$ for all $x \in H$, then $(x, y - y') = 0$ for all $x \in H$, taking $x = y - y'$ gives $y = y'$. □

This Lemma enables us to completely characterize all bounded linear functionals in a Hilbert space:

Theorem 2.2.5 *Suppose H is a Hilbert space. For each $y \in H$ define $\omega(y) : H \to \mathbb{K}$ by*

$$(\omega(y))(x) = (x, y) \quad \textit{for all} \quad x \in H.$$

Then

(a) $\omega(y) \in H^{\star}$ and $\|\omega(y)\| = \|y\|$ for all $y \in H$

(b) $\omega(y + z) = \omega(y) + \omega(z)$ for all $y, z \in H$

(c) $\omega(\lambda y) = \overline{\lambda}\omega(y)$ for all $y \in H$, $\lambda \in \mathbb{K}$

(d) ω is one-to-one and it maps onto $H^{\star}$

(e) for every $y \in H$, $\omega(y)$ is the unique normalized tangent functional to y

(f) H is reflexive.

PROOF It is easy to verify (a), (b) and (c). (d) follows directly from the Riesz Lemma 2.2.4. If f is a normalized tangent functional to $y \in H$, then

$$(y, \omega^{-1}(f)) = f(y) = \|y\|^2 = \|f\|^2 = \|\omega^{-1}(f)\|^2.$$

Hence $\|\omega^{-1}(f) - y\|^2 = 0$, which proves (e).

To prove (f), define an inner product on $H^\star$ by

$$(f, g) = (\omega^{-1}(g), \omega^{-1}(f)) \quad \text{for all} \quad f, g \in H^\star.$$

$(f, f) = \|\omega^{-1}(f)\|^2 = \|f\|^2$ implies that $H^\star$ is a Hilbert space. Hence if $F \in H^{\star\star}$, then, by Lemma 2.2.4, there exists $f \in H^\star$ such that for all $g \in H^\star$

$$F(g) = (g, f) = (\omega^{-1}(f), \omega^{-1}(g)) = g(\omega^{-1}(f))$$

which proves (f). □

Let T be a densely defined linear operator in a Hilbert space H. Define the **Hilbert space adjoint** $T^* : \mathcal{D}(T^*) \to H$ as follows. $y \in \mathcal{D}(T^*)$ if and only if there exists $z \in H$ such that

$$(Tx, y) = (x, z) \quad \text{for all} \quad x \in \mathcal{D}(T). \tag{2.20}$$

If $y \in \mathcal{D}(T^*)$, then the fact that $\mathcal{D}(T)$ is dense in H implies that there is only one $z \in H$ such that equation (2.20) holds, define $T^*y = z$. Hence

$$(Tx, y) = (x, T^*y) \quad \text{for all} \quad x \in \mathcal{D}(T),\ y \in \mathcal{D}(T^*).$$

Theorem 2.2.5 implies that the (Banach space) adjoint $T^\star$, see Definition 1.5.12, is related to the Hilbert space adjoint T^* as follows:

$$x \in \mathcal{D}(T^*) \quad \text{if and only if} \quad \omega(x) \in \mathcal{D}(T^\star)$$

$$\text{if } x \in \mathcal{D}(T^*), \text{ then } \omega(T^*x) = T^\star\omega(x).$$

A long but straightforward verification (Exercise 5), using the properties of $T^\star$ listed in Theorems 1.5.14, 1.7.8, 1.6.13, 1.7.16, gives:

Theorem 2.2.6 *Let H be a Hilbert space.*

(1) If $T \in \mathfrak{B}(H)$, then $T^* \in \mathfrak{B}(H)$ and $\|T\| = \|T^*\|$.

(2) If $T \in \mathfrak{B}(H)$, then T is compact if and only if T^* is compact.

(3) If T is a densely defined linear operator in H, then

(a) T^* is closed

(b) $\sigma(T^*) = \{ \lambda \mid \overline{\lambda} \in \sigma(T) \}$ if T is closed

(c) $(T^* - \overline{\lambda})^{-1} = ((T-\lambda)^{-1})^*$ for $\lambda \in \rho(T)$

(d) $\dim \mathcal{N}(T) = \dim \mathcal{N}(T^*)$ if T has compact resolvent.

Theorem 2.2.7 *If T is a closed and densely defined linear operator in a Hilbert space H, then T^* is densely defined and $(T^*)^* = T$.*

PROOF Note that $H \times H$ is a Hilbert space with an inner product defined by

$$[\{u,v\},\{x,y\}] = (u,x) + (v,y) \quad \text{for} \quad \{u,v\},\{x,y\} \in H \times H.$$

Let subspaces M and N of $H \times H$ be given by

$$M = \{ \{x, Tx\} \mid x \in \mathcal{D}(T) \}, \quad N = \{ \{-T^*y, y\} \mid y \in \mathcal{D}(T^*) \}.$$

$[\{x,Tx\},\{-T^*y,y\}] = -(x,T^*y) + (Tx,y) = 0$ implies $N \subset M^\perp$.

If $\{x,y\} \in M^\perp$, then $(u,x) + (Tu,y) = 0$ for all $u \in \mathcal{D}(T)$. Hence $T^*y = -x$ and therefore $N = M^\perp$. Note that $N^\perp = M$ by Theorem 2.2.2.

If $(T^*y, x) = (y,z)$ for all $y \in \mathcal{D}(T^*)$, then $\{x,z\} \in N^\perp = M$. Therefore $z = Tx$. This implies that $\mathcal{D}(T^*)$ is dense (otherwise we could start with $x = 0$ and $z \neq 0$) and that $T^{**}x = z = Tx$. Therefore T is an extension of T^{**}. We are done since T^{**} is obviously an extension of T. □

A linear operator S in a Hilbert space H is said to be **symmetric** if

$$(Sx, y) = (x, Sy) \quad \text{for all} \quad x, y \in \mathcal{D}(S).$$

A linear operator T in H is said to be **self-adjoint** if T is densely defined and $T = T^*$. Hence, T is self-adjoint iff it is symmetric, densely defined and $\mathcal{D}(T^*) = \mathcal{D}(T)$. If $T \in \mathfrak{B}(H)$, then T is self-adjoint if and only if T is symmetric. Note that self-adjoint operators are closed by Theorem 2.2.6.

The Fredholm alternative, Theorem 1.7.17, can be restated as

Theorem 2.2.8 *If T is a densely defined linear operator in a Hilbert space, then $\mathcal{R}(T)$ is closed if and only if $\mathcal{R}(T) = \mathcal{N}(T^*)^\perp$.*

PROOF Since $\mathcal{N}(T^*)^\perp$ is closed, so is $\mathcal{R}(T)$ if they are equal.

If $x \in \mathcal{D}(T)$, $z \in \mathcal{N}(T^*)$, then $(Tx, z) = (x, T^*z) = 0$ hence $\mathcal{R}(T) \subset \mathcal{N}(T^*)^\perp$.

Suppose $y \in \mathcal{N}(T^*)^\perp$ and that $\mathcal{R}(T)$ is closed. Theorem 2.2.2 implies that $y = y_1 + y_2$ for some $y_1 \in \mathcal{R}(T) \subset \mathcal{N}(T^*)^\perp$, $y_2 \in \mathcal{R}(T)^\perp$. Hence $y_2 = y - y_1 \in \mathcal{N}(T^*)^\perp$. Since $(Tx, y_2) = 0$ for all $x \in \mathcal{D}(T)$, we have that $y_2 \in \mathcal{N}(T^*)$ and hence $y_2 = 0$. Therefore $y = y_1 \in \mathcal{R}(T)$. □

2.3 Accretive Operators

Theorem 2.2.5 implies that, in a Hilbert space H, a linear operator T is **accretive** if and only if

$$\mathrm{Re}\,(Tx,x) \geq 0 \quad \text{for all} \quad x \in \mathcal{D}(T).$$

EXAMPLE 2.3.1 Assume that $a, b \in \mathbb{C}$, $\mathrm{Re}\,a \geq 0$, $\mathrm{Re}\,b \geq 0$, $c, d \in \mathbb{R}$ and let

$$Tu = -au_{xx} - bu_{yy} + cu_x + du_y \quad \text{for} \quad u \in \mathcal{D}(T) = C_0^\infty(\mathbb{R}^2).$$

It is easy to see that $\mathrm{Re}\,(Tu,u) = \mathrm{Re} \int\int a|u_x|^2 + b|u_y|^2 \geq 0$ for $u \in \mathcal{D}(T)$, hence, T is accretive in $L^2(\mathbb{R}^2)$.

A linear operator T in a Hilbert space H is said to be **m-accretive** if it is accretive and $\mathcal{R}(T-\lambda) = H$ for some scalar λ with $\mathrm{Re}\,\lambda < 0$. Since bounded operators have bounded spectrum, we have that bounded accretive operators are also m-accretive. If T is an m-accretive linear operator and $y \in \mathcal{D}(T)^\perp$, then $y = Tx - \lambda x$ for some $x \in \mathcal{D}(T)$ and some scalar λ with $\mathrm{Re}\,\lambda < 0$ hence

$$0 = \mathrm{Re}\,(y,x) = \mathrm{Re}\,(Tx,x) - \mathrm{Re}\,\lambda\|x\|^2 \geq -\mathrm{Re}\,\lambda\|x\|^2$$

implying $x = 0$, $y = 0$ and therefore *m-accretive linear operators are densely defined.* The following Theorem implies that m-accretive operators are closed.

Theorem 2.3.2 *T is an m-accretive linear operator in a Hilbert space if and only if every scalar ζ with $\mathrm{Re}\,\zeta < 0$ belongs to the resolvent set of T and*

$$\|(T-\zeta)^{-1}\| \leq 1/|\mathrm{Re}\,\zeta|.$$

PROOF If T is m-accretive, then Corollary 1.6.17 implies the conclusion. Suppose now that $(-\infty, 0) \subset \rho(T)$ and that $\|(T+\alpha)^{-1}\| \leq 1/\alpha$ for every $\alpha > 0$. If $x \in \mathcal{D}(T)$, then $\|x\| = \|(T+\alpha)^{-1}(Tx+\alpha x)\| \leq \alpha^{-1}\|Tx+\alpha x\|$. Thus

$$\alpha^2\|x\|^2 \leq \|Tx\|^2 + 2\alpha\mathrm{Re}\,(Tx,x) + \alpha^2\|x\|^2.$$

Hence $2\mathrm{Re}\,(Tx,x) \geq -\|Tx\|^2/\alpha$ and letting $\alpha \to \infty$ implies $\mathrm{Re}\,(Tx,x) \geq 0$. □

Corollary 2.3.3 *If H is a Hilbert space, $A \in \mathfrak{B}(H)$ and $A - \lambda$ is accretive for some $\lambda > 0$, then $0 \in \rho(A)$ and $\|A^{-1}\| \leq 1/\lambda$.*

We shall now employ Newton's method of finding a zero of $A - x^2 = 0$, where A is a bounded accretive operator. This will give us a **square root** of the operator. A more detailed study of fractional powers of operators is presented in Section 6.1.

Theorem 2.3.4 *Suppose that H is a Hilbert space and that $A \in \mathfrak{B}(H)$ is accretive. Then there exists a unique accretive $B \in \mathfrak{B}(H)$ such that $B^2 = A$. Moreover, $BS = SB$ for every $S \in \mathfrak{B}(H)$ such that $AS = SA$. Furthermore, if A is symmetric, then B is symmetric.*

PROOF Define $A_0 = I$ and assume that for some $n \geq 0$ we have $A_n \in \mathfrak{B}(H)$ such that for all $x \in H$

(a) $\mathrm{Re}\,(A_n x, x) \geq 2^{-n}\|x\|^2$

(b) $\mathrm{Re}\,(AA_n^{-1}x, x) \geq 0$

(c) $\mathrm{Re}\,(A_n x, AA_n^{-1}x) \geq 0$

(d) $A_n S = SA_n$ for every $S \in \mathfrak{B}(H)$ such that $AS = SA$.

Define

$$A_{n+1} = \tfrac{1}{2}(A_n + AA_n^{-1}) \tag{2.21}$$

and observe that

$$\mathrm{Re}\,(A_{n+1}x, x) = \tfrac{1}{2}\mathrm{Re}\,(A_n x, x) + \tfrac{1}{2}\mathrm{Re}\,(AA_n^{-1}x, x) \geq 2^{-n-1}\|x\|^2,$$

proving (a) for $n+1$. If $y = A_n A_{n+1}^{-1}x$ and $z = AA_n^{-2}y$, then

$$\begin{aligned} 2(AA_{n+1}^{-1}x, x) &= 2(AA_n^{-1}y, A_{n+1}A_n^{-1}y) \\ &= (AA_n^{-1}y, (I + AA_n^{-2})y) \\ &= (AA_n^{-1}y, y) + (A_n z, z), \end{aligned}$$

which implies (b) for $n+1$. If $y = A_{n+1}^{-1}x$, then

$$\begin{aligned} 4(A_{n+1}x, AA_{n+1}^{-1}x) &= 4(A_{n+1}^2 y, Ay) \\ &= ((A_n^2 + 2A + A^2A_n^{-2})y, Ay) \\ &= (A_n^2 y, Ay) + 2\|Ay\|^2 + (A^2A_n^{-2}y, Ay) \\ &= (A_n z, AA_n^{-1}z) + 2\|Ay\|^2 + (AA_n^{-1}u, A_n u), \end{aligned}$$

where $z = A_n y$ and $u = AA_n^{-1}y$, proving (c) for $n+1$. If $S \in \mathfrak{B}(H)$ is such that $AS = SA$, then also $A_n^{-1}S = SA_n^{-1}$. Hence (d) clearly holds also for $n+1$.

We therefore have $A_n \in \mathfrak{B}(H)$ for $n \geq 0$ which satisfy (a)-(d) and (2.21).

If $n \geq 0$ and $y = A_{n+1}^{-1}x$, then by (c),

$$\begin{aligned} \|(A_n - AA_n^{-1})A_{n+1}^{-1}x\|^2 &= \|A_n y\|^2 - 2\mathrm{Re}\,(A_n y, AA_n^{-1}y) + \|AA_n^{-1}y\|^2 \\ &\leq \|A_n y\|^2 + 2\mathrm{Re}\,(A_n y, AA_n^{-1}y) + \|AA_n^{-1}y\|^2 \\ &= \|2x\|^2. \end{aligned}$$

Hence $\|(A_n - AA_n^{-1})A_{n+1}^{-1}\| \leq 2$ and therefore

$$\begin{aligned} 2(A_{n+2} - A_{n+1}) &= A_{n+1} - A_n + A(A_{n+1}^{-1} - A_n^{-1}) \\ &= (A_{n+1} - A_n)(A_{n+1} - AA_n^{-1})A_{n+1}^{-1} \\ &= (A_{n+1} - A_n)\tfrac{1}{2}(A_n - AA_n^{-1})A_{n+1}^{-1} \\ \|A_{n+2} - A_{n+1}\| &\leq \tfrac{1}{2}\|A_{n+1} - A_n\| \\ \|A_{n+1} - A_n\| &\leq 2^{-n}\|A_1 - A_0\|. \end{aligned}$$

If $m > n$, then

$$\|A_m - A_n\| = \|A_m - A_{m-1} + A_{m-1} - A_{m-2} + \cdots + A_{n+1} - A_n\| \leq 2^{1-n}\|A_1 - A_0\|.$$

Hence $\{A_n\}$ is a Cauchy sequence in $\mathfrak{B}(H)$ and Theorem 1.4.4 implies that A_n converge to some $B \in \mathfrak{B}(H)$. Letting $n \to \infty$ in (a) implies that B is accretive and in

$$2A_{n+1}A_n = A_n^2 + A \tag{2.22}$$

implies that $B^2 = A$.

Suppose now that $S \in \mathfrak{B}(H)$ is accretive and $S^2 = A$. Completing the square in (2.22) gives

$$A_{n+1}^2 - S^2 = (A_{n+1} - A_n)^2.$$

Since $S^3 = SA = AS$, we have that $SA_{n+1} = A_{n+1}S$ and therefore

$$(A_{n+1} - S)(A_{n+1} + S) = (A_{n+1} - A_n)^2.$$

Since $\operatorname{Re}((A_{n+1} + S)x, x) \geq 2^{-n-1}\|x\|^2$, we have that $\|(A_{n+1} + S)^{-1}\| \leq 2^{n+1}$ and hence

$$\begin{aligned} \|A_{n+1} - S\| &= \|(A_{n+1} - A_n)^2(A_{n+1} + S)^{-1}\| \\ &\leq 2^{n+1}\|A_{n+1} - A_n\|^2 \\ &\leq 2^{-n+1}\|A_1 - A_0\|^2, \end{aligned}$$

which implies that $S = B$.

When A is symmetric we can at the beginning add an induction hypothesis

$$(A_n x, y) = (x, A_n y) \quad \text{for all} \quad x, y \in H \tag{2.23}$$

which then implies

$$\begin{aligned} (A_{n+1}x, y) &= \tfrac{1}{2}(A_n x, y) + \tfrac{1}{2}(AA_n^{-1}x, A_n A_n^{-1}y) \\ &= \tfrac{1}{2}(x, A_n y) + \tfrac{1}{2}(A_n AA_n^{-1}x, A_n^{-1}y) \\ &= \tfrac{1}{2}(x, A_n y) + \tfrac{1}{2}(x, AA_n^{-1}y). \end{aligned}$$

Hence A_{n+1} is symmetric and therefore all A_n are symmetric. Letting $n \to \infty$ in (2.23) shows that B is symmetric. □

Theorem 2.3.5 *Suppose that T is an m-accretive linear operator in a Hilbert space and that $0 \in \rho(T)$. Then there exists a unique m-accretive linear operator S such that $S^2 = T$. Moreover,*

(a) $0 \in \rho(S)$

(b) $RS^{-1} = S^{-1}R$ *for every* $R \in \mathfrak{B}(H)$ *such that* $RT^{-1} = T^{-1}R$

(c) for every $\varepsilon > 0$, $x \in \mathcal{D}(S)$ *there exists* $y \in \mathcal{D}(T)$ *such that* $\|Sx - Sy\| < \varepsilon$

(d) if T is symmetric, then S is symmetric.

PROOF Since $\mathrm{Re}\,(T^{-1}x, x) = \mathrm{Re}\,(T^{-1}x, TT^{-1}x) \geq 0$ Theorem 2.3.4, implies the existence of a unique accretive $B \in \mathfrak{B}(H)$ such that $B^2 = T^{-1}$. B is clearly one-to-one; hence we can define a linear operator $S : \mathcal{R}(B) \to H$ by $S = B^{-1}$. Since B is accretive, S is accretive. By construction, 0 belongs to an open set $\rho(S)$; hence S is m-accretive. $(S^{-1})^2 = T^{-1}$ implies $S^2 = T$.

Suppose that C is an m-accretive operator such that $C^2 = T$. Since T is one-to-one and onto, the same is true for C, and since C is closed, we have that $0 \in \rho(C)$. It is easy to see that C^{-1} is accretive and $(C^{-1})^2 = T^{-1}$, hence, Theorem 2.3.4 implies $C^{-1} = B$ and $C = S$.

To show (c), define $[x, y] = (Sx, Sy)$ for $x, y \in \mathcal{D}(S)$. Since $0 \in \rho(S)$, we have that $[\cdot, \cdot]$ makes $\mathcal{D}(S)$ into a Hilbert space. Define a linear operator R in $\mathcal{D}(S)$ to be the restriction of T to $\mathcal{D}(R)$ where $x \in \mathcal{D}(R)$ iff $x \in \mathcal{D}(T)$ and $Tx \in \mathcal{D}(S)$. It is easy to see that R is m-accretive in $\mathcal{D}(S)$ and hence densely defined.

If T is symmetric, then T^{-1} is symmetric. Hence B is symmetric by Theorem 2.3.4, and therefore $S = B^{-1}$ is symmetric. □

Theorem 2.3.6 *Every densely defined accretive linear operator in a Hilbert space has an m-accretive extension.*

PROOF Let S be a densely defined accretive linear operator. Note that $\|x\|^2 \leq \mathrm{Re}\,((S+1)x, x) \leq \|(S+1)x\|\|x\|$, which implies

$$\|x\| \leq \|(S+1)x\| \quad \text{for} \quad x \in \mathcal{D}(S). \tag{2.24}$$

Pick w in the Hilbert space H. By Theorem 2.2.2,

$$w = 2y + z \quad \text{for some } z \in \overline{\mathcal{R}(1+S)},\ y \in \overline{\mathcal{R}(1+S)}^{\perp}.$$

There exists $\{x_n\}$ such that $\lim_{n\to\infty}(S+1)x_n = z$. (2.24) implies that $\{x_n\}$ converges to some $x \in H$ which does not depend on the particular choice of

$\{x_n\}$. Define $u = y + x$ and let $\mathcal{D}(T)$ denote the set of all such u. Since the mapping $w \to u$ is linear, we can define a linear operator T by

$$Tu = w - u = y + z - x \tag{2.25}$$

provided that $u = 0$ implies $w = 0$ - which will be shown next.

Take any $v \in \mathcal{D}(S)$ and any scalar α and note

$$(u + \alpha v, w - u + \alpha Sv)$$

$$= \|y\|^2 - (y, x + \alpha v) + (x + \alpha v, y) + \lim_{n\to\infty} (x_n + \alpha v, S(x_n + \alpha v)).$$

Hence

$$\mathrm{Re}\,(u + \alpha v, w - u + \alpha Sv) \geq 0. \tag{2.26}$$

If, in particular, $u = 0$, then $\mathrm{Re}\,\alpha(v, w + \alpha Sv) \geq 0$. Thus if $\alpha = -(w, v)\varepsilon$ with $\varepsilon > 0$, $\varepsilon \to 0$, then $-|(v, w)|^2 \geq 0$; hence $(v, w) = 0$ and, because $\mathcal{D}(S)$ is dense, we have $w = 0$. Therefore T is a well defined by (2.25), it is easy to see that T is an extension of S and (2.26) implies that T is accretive. By construction, $\mathcal{R}(T + 1) = H$ and hence T is m-accretive. □

2.4 Weak Solutions

When solving PDEs one is usually faced with the following problem: for given f find u such that

$$Lu = f \tag{2.27}$$

where L is a differential operator. For a 'nice' ϕ it is usually a trivial matter to evaluate $L\phi$. However, in order to solve the problem for a large enough class of f, one has to be able to evaluate $L\phi$ for 'not so nice' ϕ. With a proper definition of L, including its domain, it can be very easy to solve (2.27) as well as many other problems involving L. Consider the following.

EXAMPLE 2.4.1 It was shown that the operator T in Example 2.3.1 is accretive. We shall see later, Theorem 3.1.7, that T is densely defined. Therefore Theorem 2.3.6 implies that T has an m-accretive extension S and, in particular, for every $f \in L^2(\mathbb{R}^2)$ and $\lambda > 0$ there exists $u \in \mathcal{D}(S)$ such that

$$Su + \lambda u = f.$$

It will be shown (Corollary 4.3.11) that, with this extension, we can also solve the evolution problem $w_t + Sw = 0$.

There are other ways to make extensions. The one based on the variational formulation, described later in the Sectorial Forms section, is perhaps the most often used to obtain extensions of elliptic operators. The one based on Theorem 2.3.6 is more general and applies also to hyperbolic problems.

Another way to solve (2.27) is the following. Find L', called a **formal adjoint** of L, such that $(Lu, v) = (u, L'v)$ for $v \in \mathcal{D}(L')$ (the space of test functions) and then try to find u, called a **weak solution** of (2.27), such that

$$(u, L'v) = (f, v) \quad \text{for all} \quad v \in \mathcal{D}(L'). \tag{2.28}$$

L' is called a formal adjoint because we usually do not know ahead of time what is the correct domain of L and hence what exactly is the adjoint of L. In view of Theorem 2.2.7, solutions obtained via various extensions can usually be interpreted also as weak solutions.

Theorem 2.4.2 *If L' is a linear operator in a Hilbert space H such that for some $c > 0$ we have*

$$c\|v\| \leq \|L'v\| \quad \textit{for all} \quad v \in \mathcal{D}(L'), \tag{2.29}$$

then for each $f \in H$ there exists a unique $u \in \overline{\mathcal{R}(L')}$ such that (2.28) holds.

PROOF L'^{-1} is a bounded map from $\mathcal{R}(L')$ to H, hence, Theorem 1.4.3 gives an extension B of L'^{-1} such that $B \in \mathfrak{B}(M, H)$, $M \equiv \overline{\mathcal{R}(L')}$. Let $f \in H$ be given. The Riesz Lemma 2.2.4 implies existence of $u \in M$ such that $(Bw, f) = (w, u)$ for all $w \in M$ and hence (2.28) follows by choosing $w = L'v$. If (2.28) holds for some $\hat{u} \in M$, then $(u - \hat{u}, w) = 0$ for all $w \in \mathcal{R}(L')$. Hence $u = \hat{u}$. □

Observe that if $L' - c$ is accretive for some $c > 0$, then (2.29) is satisfied.

Theorem 2.4.3 shows how weak solutions can be approximated numerically. This method, however, has not much in common with the widely used finite element method based on Theorem 2.8.11.

Theorem 2.4.3 *Suppose that $\mathcal{V}_1 \subset \mathcal{V}_2 \subset \cdots$ are finite dimensional subspaces of a Hilbert space H. Let $\mathcal{V} = \cup_{n=1}^{\infty} \mathcal{V}_n$ and suppose that $L' : \mathcal{V} \to H$ is linear ($\mathcal{D}(L') = \mathcal{V}$) and such that the inequality (2.29) is satisfied with some $c > 0$.*

Pick any $f \in H$ and let $u \in \overline{\mathcal{R}(L')}$ be such that (2.28) holds. Then, for each $n \geq 1$, there exists a unique $z_n \in \mathcal{V}_n$ such that

$$(L'z_n, L'v) = (f, v) \quad \textit{for all} \quad v \in \mathcal{V}_n.$$

Moreover, $\lim_{n\to\infty} \|L'z_n - u\| = 0$.

PROOF Let $[u,v] = (L'u, L'v)$ for $u, v \in \mathcal{V}_n$. (2.29) implies that $[\cdot,\cdot]$ is an inner product on $\mathcal{V}_n$ and, since $\mathcal{V}_n$ is finite dimensional, it is a Hilbert space (see Theorem 1.7.1). (2.29) and the Riesz Lemma 2.2.4 imply existence of a unique $z_n \in \mathcal{V}_n$ such that $(v,f) = [v,z_n]$ for all $v \in \mathcal{V}_n$.

Take any $v \in \mathcal{V}_n$ and note that

$$\begin{aligned} \|L'z_n - u\|^2 &= (L'z_n - u, L'v - u) + (L'z_n - u, L'(z_n - v)) \\ &= (L'z_n - u, L'v - u) \\ &\leq \|L'z_n - u\|\|L'v - u\| \\ \|L'z_n - u\| &\leq \|L'v - u\| \\ \|L'z_n - u\| &= \inf_{v\in\mathcal{V}_n} \|L'v - u\| \equiv d_n. \end{aligned}$$

Since $\mathcal{V}_n \subset \mathcal{V}_{n+1}$, we have that $d_{n+1} \leq d_n$ and, since $u \in \overline{\mathcal{R}(L')}$, we have that $\lim_{n\to\infty} d_n = 0$. □

A key tool for obtaining a formal adjoint of a differential operator is

Lemma 2.4.4 *If Ω is any nonempty open set in $\mathbb{R}^n$, $n \geq 1$, $u \in C^m(\Omega)$, $m \geq 1$, $\varphi \in C_0^m(\Omega)$ and α is a multi-index with $|\alpha| \leq m$, then*

$$\int_\Omega u D^\alpha \varphi = (-1)^{|\alpha|} \int_\Omega (D^\alpha u)\varphi.$$

PROOF Suppose first that $u \in C^1(\Omega)$, $\varphi \in C_0^1(\Omega)$. Define $f = \varphi u$ in Ω, $f = 0$ in $\mathbb{R}^n \backslash \Omega$. Note that $f \in C_0^1(\mathbb{R}^n)$, hence $\text{supp}(f) \subset C \cap \Omega$ for some open finite cube C and therefore

$$\int_\Omega u\frac{\partial\varphi}{\partial x_i} + \varphi\frac{\partial u}{\partial x_i} = \int_{\text{supp}(f)} \frac{\partial f}{\partial x_i} = \int_C \frac{\partial f}{\partial x_i} = 0$$

$$\int_\Omega u\frac{\partial\varphi}{\partial x_i} = -\int_\Omega \varphi\frac{\partial u}{\partial x_i}.$$

Repeated use of the last identity completes the proof. □

After proving existence of a weak solution of a PDE, one usually attempts to prove its regularity which then implies that the weak solution is also the classical solution. Regularity of weak solutions is studied in Section 3.8, but for now let us prove only the key tool in one dimension.

Lemma 2.4.5 *If* $-\infty \le a < c < b \le \infty$, $f \in L^1_{loc}(a,b)$, $g \in L^1_{loc}(a,b)$ *and*

$$\int_a^b f\psi' = -\int_a^b g\psi \quad \text{for all} \quad \psi \in C_0^1(a,b),$$

then there exists a constant k *such that*

$$f(x) = k + \int_c^x g \quad \text{for almost all } x \in (a,b).$$

PROOF Define $u(x) = f(x) - \int_c^x g$. For $a < \alpha < \beta < \gamma < \delta < b$ let

$$\varphi(x) = \begin{cases} 0 & \text{if} \quad x < \alpha \\ \frac{x-\alpha}{\beta-\alpha} & \text{if} \quad \alpha \le x < \beta \\ 1 & \text{if} \quad \beta \le x < \gamma \\ \frac{\delta-x}{\delta-\gamma} & \text{if} \quad \gamma \le x < \delta \\ 0 & \text{if} \quad \delta \le x. \end{cases}$$

Let $\psi_\varepsilon(x) = (1/\varepsilon)\int_x^{x+\varepsilon} \varphi$. $\psi_\varepsilon \in C_0^1(a,b)$ for small positive ε, hence

$$0 = \int_a^b u\psi_\varepsilon' = \int_a^b u(x)\frac{\varphi(x+\varepsilon)-\varphi(x)}{\varepsilon}dx \xrightarrow{\varepsilon\to 0} \int_a^b u\varphi'$$

$$= \frac{1}{\beta-\alpha}\int_\alpha^\beta u - \frac{1}{\delta-\gamma}\int_\gamma^\delta u.$$

Thus $\int_\alpha^\beta u = k(\beta-\alpha)$ where $k = \int_\gamma^\delta u/(\delta-\gamma)$ and therefore $u = k$ on (a,γ). □

EXAMPLE 2.4.6 Let us examine weak solutions of $Lu = f$ in the Hilbert space $L^2(0,1)$ with $Lu = u'$. Let us choose for the space of test functions $C_0^1(0,1) \equiv \mathcal{D}(L')$ and $L'\varphi = -\varphi'$. Thus the weak formulation is: for given $f \in L^2(0,1)$, find $u \in L^2(0,1)$ such that

$$-\int_0^1 u(x)\varphi'(x)dx = \int_0^1 f(x)\varphi(x)dx \quad \text{for all} \quad \varphi \in C_0^1(0,1). \tag{2.30}$$

If $\varphi \in \mathcal{D}(L')$, then $\varphi(x) = \int_0^x \varphi'(s)ds$; hence $|\varphi(x)| \le \|\varphi'\|_2$ and (2.29) follows. Thus Theorem 2.4.2 proves existence of $u \in L^2(0,1)$ such that (2.30) holds. Lemma 2.4.5 now implies that for some constant k we have to have that $u(x) = k + \int_0^x f(s)ds$ a.e. and indeed all such u satisfy (2.30) as one can easily verify directly. The unique u given by Theorem 2.4.2 lies in the closure of the range of L'. Thus, $u = \lim_{n\to\infty} \varphi_n'$ hence $\int_0^1 u = 0$ which determines the constant.

EXAMPLE 2.4.7 We shall now show existence of a weak solution of

$$u_t - Au_{xx} - Bu_x - Cu = f, \qquad x \in \mathbb{R}, 0 \le t \le T \tag{2.31}$$

$$u(x,0) = 0, \qquad x \in \mathbb{R}, \tag{2.32}$$

assuming that $T \in (0,\infty)$ is arbitrary, $\Omega = \mathbb{R} \times (0,T)$, $A \in C_B^2(\Omega)$, $A \geq 0$, $B \in C_B^1(\Omega)$, $C \in C_B(\Omega)$, $f \in L^2(\Omega)$. In this example, all functions are assumed to be real valued, hence, the scalar field of the Hilbert space $L^2(\Omega)$ is now $\mathbb{R}$. (2.31) is a typical parabolic equation when $A > 0$.

Multiplying (2.31) by φ and formally integrating by parts, using (2.32), we conclude that an appropriate formal adjoint is

$$L'\varphi = -\varphi_t - (A\varphi)_{xx} + (B\varphi)_x - C\varphi$$

for $\varphi \in \mathcal{D}(L')$ which are required to satisfy:

(a) $\varphi \in C_u^2(\Omega)$, $\varphi(x,T) = 0$ for $x \in \mathbb{R}$

(b) there exists L_φ such that $\varphi(x,t) = 0$ for $|x| > L_\varphi$, $0 \leq t \leq T$.

Choose $\varphi \in \mathcal{D}(L')$ and integrate over x at fixed $t \in (0,T)$ in

$$\begin{aligned}
\int_{\mathbb{R}} \varphi L'\varphi &= -\int_{\mathbb{R}} \varphi_t\varphi - \int_{\mathbb{R}} (A\varphi)_{xx}\varphi + \int_{\mathbb{R}} (B\varphi)_x\varphi - \int_{\mathbb{R}} C\varphi^2 \\
&= -\int_{\mathbb{R}} \varphi_t\varphi + \int_{\mathbb{R}} (A\varphi)_x\varphi_x - \int_{\mathbb{R}} B\varphi\varphi_x - \int_{\mathbb{R}} C\varphi^2 \\
&= -\int_{\mathbb{R}} \varphi_t\varphi + \int_{\mathbb{R}} A\varphi_x^2 + \frac{1}{2}\int_{\mathbb{R}} (A_x - B)(\varphi^2)_x - \int_{\mathbb{R}} C\varphi^2 \\
&= -\int_{\mathbb{R}} \varphi_t\varphi + \int_{\mathbb{R}} A\varphi_x^2 - \frac{1}{2}\int_{\mathbb{R}} (A_{xx} - B_x + 2C)\varphi^2 \\
&\geq -\frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}} \varphi^2 - \frac{k}{2}\int_{\mathbb{R}} \varphi^2
\end{aligned}$$

where $k \geq 0$ is such that $k \geq A_{xx} - B + 2C$, hence,

$$\begin{aligned}
-\frac{d}{dt}\left(e^{kt}\int_{\mathbb{R}} \varphi^2\right) &\leq 2e^{kt}\int_{\mathbb{R}} \varphi L'\varphi \\
\mathrm{e}^{kt}\int_{\mathbb{R}} \varphi^2 &\leq 2\int_t^T ds\, \mathrm{e}^{ks}\int_{\mathbb{R}} dx\, \varphi L'\varphi \\
&\leq 2e^{kT}\int_0^T ds \int_{\mathbb{R}} dx\, |\varphi L'\varphi| \leq 2e^{kT}\|\varphi\|\|L'\varphi\| \\
\|\varphi\| &\leq 2Te^{kT}\|L'\varphi\|
\end{aligned}$$

and therefore Theorem 2.4.2 applies, proving existence of a weak solution.

EXAMPLE 2.4.8 A more elaborate version of Example 2.4.7 will now be presented, namely, proof of existence of a weak solution of a **symmetric hyperbolic system**,

$$A\frac{\partial u}{\partial t} + \sum_{k=1}^{n} A^k \frac{\partial u}{\partial x_k} + Bu = f, \qquad x \in \mathbb{R}^n, 0 \leq t \leq T \tag{2.33}$$

$$u(x,0) = 0, \qquad x \in \mathbb{R}^n, \tag{2.34}$$

assuming $T \in (0,\infty)$ is arbitrary, $\Omega = \mathbb{R}^n \times (0,T)$ and $u = u(x,t) = u(x_1,\dots,x_n,t)$ is a column vector with N components. All functions are real valued in this example and f belongs to the real Hilbert space $(L^2(\Omega))^N$. $A, A^1, \dots, A^n, B$ are $N \times N$ matrices with $A_{ij} \in C_B^1(\Omega)$, $A_{ij} = A_{ji}$, $A_{ij}^k \in C_B^1(\Omega)$, $A_{ij}^k = A_{ji}^k$, $B_{ij} \in C_B(\Omega)$ and A is uniformly positive definite, i.e. there exists $\sigma > 0$ such that

$$y^\top A(x,t)y \geq \sigma \|y\|^2 \quad \text{for} \quad y \in \mathbb{R}^N,\ x \in \mathbb{R}^n,\ 0 \leq t \leq T. \tag{2.35}$$

Multiplying (2.33) by $\varphi^\top$ and formally integrating by parts, using (2.34), we conclude that an appropriate formal adjoint is

$$L'\varphi = -\frac{\partial}{\partial t}(A\varphi) - \sum_{k=1}^n \frac{\partial}{\partial x_k}(A^k\varphi) + B^*\varphi$$

for $\varphi \in \mathcal{D}(L')$ which are required to satisfy:

(a) $\varphi \in (C_u^1(\Omega))^N$, $\varphi(x,T) = 0$ for $x \in \mathbb{R}^n$

(b) there exists L_φ such that $\varphi(x,t) = 0$ for $|x| > L_\varphi$, $0 \leq t \leq T$.

Choose $\varphi \in \mathcal{D}(L')$. Integration over $x \in \mathbb{R}^n$ at fixed $t \in (0,T)$, using repeatedly the fact that $2\varphi^\top D(A\varphi) = D(\varphi^\top A\varphi) + \varphi^\top (DA)\varphi$ for any symmetric A, gives

$$\begin{aligned}
\int \varphi^\top L'\varphi &= -\int \varphi^\top \frac{\partial}{\partial t}(A\varphi) - \sum_{k=1}^n \int \varphi^\top \frac{\partial}{\partial x_k}(A^k\varphi) + \int \varphi^\top B^*\varphi \\
2\int \varphi^\top L'\varphi &= -\frac{d}{dt}\int \varphi^\top A\varphi - \int \varphi^\top \frac{\partial A}{\partial t}\varphi - \sum_{k=1}^n \int \varphi^\top \frac{\partial A^k}{\partial x_k}\varphi + 2\int \varphi^\top B^*\varphi \\
&= -\frac{d}{dt}\int \varphi^\top A\varphi - \int \varphi^\top \left(\frac{\partial A}{\partial t} + \sum_{k=1}^n \frac{\partial A^k}{\partial x_k} - 2B^*\right)\varphi \\
&\geq -\frac{d}{dt}\int \varphi^\top A\varphi - \mathrm{c}\int \varphi^\top A\varphi,
\end{aligned}$$

where $\mathrm{c} \geq \|\partial A/\partial t + \sum_{k=1}^n \partial A^k/\partial x_k - 2B^*\|/\sigma$ and (2.35) was used, hence

$$\begin{aligned}
-\frac{d}{dt}\left(e^{ct}\int_{\mathbb{R}^n} \varphi^\top A\varphi\right) &\leq 2e^{ct}\int_{\mathbb{R}^n} \varphi^\top L'\varphi \\
e^{ct}\int_{\mathbb{R}^n} \varphi^\top A\varphi &\leq 2\int_t^T ds\, e^{cs}\int_{\mathbb{R}^n} dx\, \varphi^\top L'\varphi \\
\sigma\int_{\mathbb{R}^n} \varphi^\top\varphi &\leq 2e^{cT}\int_0^T ds\int_{\mathbb{R}^n} dx\, |\varphi^\top L'\varphi| \leq 2e^{cT}\|\varphi\|\|L'\varphi\| \\
\sigma\|\varphi\| &\leq 2Te^{cT}\|L'\varphi\|
\end{aligned}$$

and therefore Theorem 2.4.2 implies existence of $u \in (L^2(\Omega))^N$ such that

$$\int_0^T \int_{\mathbb{R}^n} u^\top L'\varphi = \int_0^T \int_{\mathbb{R}^n} f^\top\varphi \quad \text{for all} \quad \varphi \in \mathcal{D}(L').$$

Under some additional assumptions it can be proven that this u satisfies also (2.33) and (2.34).

Many other (hyperbolic) equations can be written as (2.33). For example, consider the following **wave equation**,

$$\frac{\partial^2 v}{\partial t^2} = \sum_{i,j=1}^{n} a_{ij}(x,t)\frac{\partial^2 v}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x,t)\frac{\partial v}{\partial x_i} + c(x,t)\frac{\partial v}{\partial t} + d(x,t)v + g,$$

where a_{ij} form a uniformly positive definite symmetric matrix. Let u be a vector of $N = n + 2$ components,

$$u_1 = \frac{\partial v}{\partial x_1}, \quad \cdots, \quad u_n = \frac{\partial v}{\partial x_n}, \quad u_{n+1} = \frac{\partial v}{\partial t}, \quad u_{n+2} = v.$$

This gives N equations in the form of (2.33):

$$\sum_{j=1}^{n} a_{ij}\frac{\partial u_j}{\partial t} - \sum_{j=1}^{n} a_{ij}\frac{\partial u_{n+1}}{\partial x_j} = 0 \qquad \text{for} \quad i = 1, \ldots, n$$

$$\frac{\partial u_{n+1}}{\partial t} - \sum_{i,j=1}^{n} a_{ij}\frac{\partial u_i}{\partial x_j} - \sum_{i=1}^{n} b_i u_i - c u_{n+1} - d u_{n+2} = g$$

$$\frac{\partial u_{n+2}}{\partial t} - u_{n+1} = 0.$$

2.5 Example: Constant Coefficient PDEs

Let Ω be any nonempty open set in $\mathbb{R}^n$, $n \geq 1$, and let a constant coefficient differential operator in the Hilbert space $L^2(\Omega)$ be given by

$$L = \sum_{|\alpha| \leq m} c_\alpha D^\alpha,$$

where c_α are arbitrary complex constants and m is any nonnegative integer called the order of the differential operator L. We shall show that for every $f \in L^2(\Omega)$ there exists a weak solution of $Lu = f$ provided that Ω is bounded and that at least one of c_α is nonzero.

Using Lemma 2.4.4 we see that $(L\varphi, \psi) = (\varphi, L'\psi)$ for all $\varphi, \psi \in C_0^\infty(\Omega)$ where

$$L' = \sum_{|\alpha| \leq m} (-1)^{|\alpha|}\, \overline{c_\alpha}\, D^\alpha.$$

Note $LL'\varphi = L'L\varphi$. Hence $(L\varphi, L\varphi) = (\varphi, L'L\varphi) = (\varphi, LL'\varphi) = (L'\varphi, L'\varphi)$, i.e.

$$\|L\varphi\| = \|L'\varphi\| \quad \text{for all} \quad \varphi \in C_0^\infty(\Omega). \tag{2.36}$$

Define a polynomial of n-variables $\zeta_1, \dots, \zeta_n$ by

$$P(\zeta) = \sum_{|\alpha| \le m} c_\alpha \zeta^\alpha = \sum_{|\alpha| \le m} c_\alpha \zeta_1^{\alpha_1} \cdots \zeta_n^{\alpha_n}$$

and let

$$P_k = \frac{\partial P}{\partial \zeta_k}, \quad P_{kk} = \frac{\partial^2 P}{\partial \zeta_k^2} \quad \text{for} \quad k = 1, \dots, n$$

$$P^{(\alpha)} = \left(\frac{\partial}{\partial \zeta_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial \zeta_n}\right)^{\alpha_n} P \quad \text{for} \quad \alpha = (\alpha_1, \dots, \alpha_n).$$

Note $L = P(D)$ and, for example,

$$P_1(D) = \sum_{|\alpha| \le m,\, \alpha_1 > 0} c_\alpha \alpha_1 D_1^{\alpha_1 - 1} D_2^{\alpha_2} \cdots D_n^{\alpha_n},$$

$P^{(\alpha)} = \alpha! c_\alpha$ when $|\alpha| = m$.

Lemma 2.5.1 *If for some a_k, d_k we have that $|x_k - a_k| \le d_k/2$ all $x \in \Omega$, then*

$$\|P_k(D)\varphi\| \le m d_k \|P(D)\varphi\| \quad \textit{for all} \quad \varphi \in C_0^\infty(\Omega).$$

PROOF Let us fix Ω and do induction on m. When $m = 0$ the assertion is obvious. Assume $m \ge 1$ and that the assertion is true for all constant coefficient differential operators of order $m - 1$, hence, we may in particular assume that

$$\|P_{kk}(D)\varphi\| \le (m-1) d_k \|P_k(D)\varphi\| \quad \text{for all} \quad \varphi \in C_0^\infty(\Omega).$$

Let $v(x) = x_k - a_k$ and note that (1.7) implies

$$P(D)(v\varphi) = vP(D)\varphi + P_k(D)\varphi \quad \text{for all} \quad \varphi \in C_0^\infty(\Omega).$$

Hence

$$\begin{aligned}
\|P_k(D)\varphi\|^2 &= (P(D)(v\varphi) - vP(D)\varphi, P_k(D)\varphi) \\
&= (P(D)(v\varphi), P_k(D)\varphi) - (vP(D)\varphi, P_k(D)\varphi) \\
&= (\varphi, vP_k(D)\psi) - (vP(D)\varphi, P_k(D)\varphi) \quad \text{where } \psi = P(D)'\varphi \\
&= (\varphi, P_k(D)(v\psi) - P_{kk}(D)\psi) - (vP(D)\varphi, P_k(D)\varphi) \\
&= (P_k(D)'\varphi, v\psi) - (P_{kk}(D)'\varphi, \psi) - (vP(D)\varphi, P_k(D)\varphi) \\
&\le \|P_k(D)'\varphi\| \|v\|_\infty \|\psi\| + \|P_{kk}(D)'\varphi\| \|\psi\| + \\
&\qquad \|v\|_\infty \|P(D)\varphi\| \|P_k(D)\varphi\| \\
&= 2\|v\|_\infty \|P_k(D)\varphi\| \|P(D)\varphi\| + \|P_{kk}(D)\varphi\| \|P(D)\varphi\| \\
&\le d_k \|P(D)\varphi\| \|P_k(D)\varphi\| + \|P_{kk}(D)\varphi\| \|P(D)\varphi\| \\
&\le m d_k \|P(D)\varphi\| \|P_k(D)\varphi\|.
\end{aligned}$$

□

Repeated applications of the above Lemma give

Theorem 2.5.2 (Hörmander) *If Ω is contained in a finite box $|x_k - a_k| \leq d_k/2$, $1 \leq k \leq n$, then for every multi-index $\beta = (\beta_1, \ldots, \beta_n)$ and every $\varphi \in C_0^\infty(\Omega)$,*

$$\|P^{(\beta)}(D)\varphi\| \leq m(m-1)\cdots(m-|\beta|+1)d_1^{\beta_1}\cdots d_n^{\beta_n}\|P(D)\varphi\|.$$

Theorem 2.5.3 *If Ω is bounded and if at least one of c_α is not zero, then for every $f \in L^2(\Omega)$ there exists $u \in L^2(\Omega)$ such that*

$$(u, L'\varphi) = (f, \varphi) \quad \textit{for all} \quad \varphi \in C_0^\infty(\Omega).$$

PROOF Note that $P^{(\beta)}$ is a nonzero constant for some β, hence, Theorem 2.5.2 and (2.36) imply (2.29) and therefore Theorem 2.4.2 applies. □

EXAMPLE 2.5.4 Δ is used to denote the **Laplace operator** (or the **Laplacian**)

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}.$$

Lemma 2.5.1 implies that if at least one coordinate of points in Ω is bounded, then for each $f \in L^2(\Omega)$ there exists a weak solution of the typical **elliptic equation** $\Delta u = f$. That is, there exists $u \in L^2(\Omega)$ such that

$$\int_\Omega u\Delta\varphi = \int_\Omega f\varphi \quad \text{for all} \quad \varphi \in C_0^\infty(\Omega).$$

Regularity of u is given in Example 3.8.4.

2.6 Self-adjoint Operators

Observe that if S is a symmetric linear operator in a Hilbert space H, then

(i) $(Sx, x) \in \mathbb{R}$ for all $x \in \mathcal{D}(S)$

(ii) eigenvalues of S are real

(iii) eigenvectors of S corresponding to different eigenvalues are orthogonal

(iv) if H is a complex Hilbert space and $\lambda = \lambda_r + i\lambda_i$, with $\lambda_r, \lambda_i \in \mathbb{R}$, then

$$\|(S-\lambda)x\|^2 = \|(S-\lambda_r)x\|^2 + \lambda_i^2\|x\|^2 \quad \text{for all} \quad x \in \mathcal{D}(S). \tag{2.37}$$

Note that the same holds for real symmetric matrices. In this section it will be shown that many other properties of real symmetric matrices also extend to operators that are self-adjoint - not just symmetric. Recall that a densely defined symmetric linear operator S is self-adjoint if and only if $\mathcal{D}(S^*) = \mathcal{D}(S)$. When S is unbounded, the task of showing that $\mathcal{D}(S^*) = \mathcal{D}(S)$ can be nontrivial (try Exercise 10). The following Lemma can sometimes make the job easier.

Lemma 2.6.1 *If T is a linear operator in a Hilbert space such that its resolvent set contains a real number λ, then the following statements are equivalent:*

(a) T is symmetric

(b) $(T-\lambda)^{-1}$ is symmetric

(c) T is self-adjoint.

PROOF Let H denote the Hilbert space and $K = (T-\lambda)^{-1}$. (a) implies

$$(Kx, y) = (Kx, (T-\lambda)Ky) = (x, Ky) \quad \text{for all} \quad x, y \in H$$

and this gives (b). Similarly, if (b) holds and $x, y \in \mathcal{D}(T)$, then

$$(Tx, y) - \lambda(x, y) = ((T-\lambda)x, K(T-\lambda)y) = (x, (T-\lambda)y) = (x, Ty) - \lambda(x, y)$$

and therefore T is symmetric.

Assume K is symmetric and that $y, z \in H$ are such that

$$(Tx, y) = (x, z) \quad \text{for all} \quad x \in \mathcal{D}(T).$$

If $\overline{\mathcal{D}(T)} \neq H$, then one could take $z \neq 0$, $y = 0$. Let $x = K(y - K(z - \lambda y))$ and note

$$\begin{aligned} \|y - K(z-\lambda y)\|^2 &= ((T-\lambda)x, y - K(z-\lambda y)) \\ &= (Tx, y) - \lambda(x, y) - ((T-\lambda)x, K(z-\lambda y)) \\ &= (x, z) - \lambda(x, y) - (x, z - \lambda y) = 0. \end{aligned}$$

Therefore $y = K(z - \lambda y) \in \mathcal{D}(T)$ and hence $Ty = z$. This shows that T is densely defined and that T is an extension of T^*. Since T is also symmetric it is self-adjoint. □

Theorem 2.6.2 *$\sigma(T) \subset \mathbb{R}$ for every self-adjoint operator T.*

PROOF Let H be a complex Hilbert space and $\lambda = \lambda_r + i\lambda_i$ with $\lambda_i \neq 0$.

(2.37) implies $\|(T-\lambda)x\| \geq |\lambda_i|\|x\|$ for $x \in \mathcal{D}(T)$, hence, $T - \lambda$ is one-to-one and since T is closed, we also have that $\mathcal{R}(T-\lambda)$ is closed. If $y \in \mathcal{R}(T-\lambda)^{\perp}$, then $(Tx, y) = (x, \overline{\lambda}y)$ for all $x \in \mathcal{D}(T)$, hence $T^*y = \overline{\lambda}y$ and, since symmetric T can have only real eigenvalues, we have that $y = 0$ and therefore $\mathcal{R}(T-\lambda) = H$ by Theorem 2.2.2. Corollary 1.6.2 implies that $\lambda \in \rho(T)$. □

If T is a symmetric operator in a complex Hilbert space, then $\pm iT$ are accretive. Theorem 2.6.2 implies that $\pm iT$ are m-accretive when T is self-adjoint. The converse also holds:

Theorem 2.6.3 *If T is a symmetric operator in a complex Hilbert space H, then T is self-adjoint if and only if $\mathcal{R}(T+i) = \mathcal{R}(T-i) = H$.*

PROOF The 'only if' part is given by Theorem 2.6.2. To prove the 'if' part, assume that $(Tx, y) = (x, z)$ for all $x \in \mathcal{D}(T)$. Choose $w \in \mathcal{D}(T)$ so that $(T-i)w = z - iy$. Hence $((T+i)x, y-w) = 0$ for all $x \in \mathcal{D}(T)$ and therefore $y = w$, $Ty = z$. If $\mathcal{D}(T)$ would not be dense in H, we could have chosen $y = 0$ and $z \neq 0$, but this is not possible since $Ty = z$. Thus, $\mathcal{D}(T)$ is dense in H, $\mathcal{D}(T) = \mathcal{D}(T^*)$ and therefore $T = T^*$. □

Our basic tool for obtaining various bounds involving self-adjoint operators is the following:

Theorem 2.6.4 *Suppose that $H \neq \{0\}$ is a Hilbert space and that $T \in \mathfrak{B}(H)$ is symmetric. Then $\|T\| = \sup_{\|x\|=1} |(Tx, x)| = |\lambda|$ for some $\lambda \in \sigma(T)$.*

PROOF Clearly $m \equiv \sup_{\|x\|=1} |(Tx, x)| \leq \|T\|$. To show $\|T\| \leq m$, observe that for all $x, y \in H$,

$$\begin{aligned} 2(Tx, y) + 2(Ty, x) &= (T(x+y), x+y) - (T(x-y), x-y) \\ &\leq m(\|x+y\|^2 + \|x-y\|^2) \\ &= 2m(\|x\|^2 + \|y\|^2). \end{aligned}$$

Hence if $Tx \neq 0$ and $y = (\|x\|/\|Tx\|)Tx$, then

$$2\|x\|\|Tx\| = (Tx, y) + (y, Tx) \leq m(\|x\|^2 + \|y\|^2) = 2m\|x\|^2$$

implies $\|Tx\| \leq m\|x\|$ - which is true also when $Tx = 0$. Thus, $\|T\| = m$.

Choose $x_n \in H$ such that $\|x_n\| = 1$ and $\|T\| = \lim_{n\to\infty} |(Tx_n, x_n)|$. By renaming a subsequence of $\{x_n\}$ we may assume that (Tx_n, x_n) converge to some real number λ with $|\lambda| = \|T\|$. Observe that

$$\begin{aligned} \|(T-\lambda)x_n\|^2 &= \|Tx_n\|^2 - 2\lambda(Tx_n, x_n) + \lambda^2\|x_n\|^2 \\ &\leq 2\lambda^2 - 2\lambda(Tx_n, x_n) \to 0. \end{aligned}$$

Hence the assumption $\lambda \notin \sigma(T)$ leads to the contradiction

$$1 = \|x_n\| = \|(T-\lambda)^{-1}(T-\lambda)x_n\| \leq \|(T-\lambda)^{-1}\|\,\|(T-\lambda)x_n\| \to 0.$$

□

A sharp version of Exercise 18 in Chapter 1:

Theorem 2.6.5 *If T is a self-adjoint operator in a Hilbert space $H \neq \{0\}$, then its spectrum is not empty and $\|(T-\lambda)^{-1}\| = \operatorname{dist}(\lambda, \sigma(T))^{-1}$ for all $\lambda \in \rho(T)$.*

PROOF Note that $\sigma(T) \subset \mathbb{R}$ by Theorem 2.6.2. Let $\lambda = \lambda_r + i\lambda_i \in \rho(T)$.

If $\lambda_r \in \sigma(T)$, then (2.37) implies that $\|(T-\lambda)^{-1}\| \leq 1/|\lambda_i| = \operatorname{dist}(\lambda, \sigma(T))^{-1}$.

Next let us consider the case when $\lambda_r \in \rho(T)$. Pick any $\delta \in (0, \infty)$ such that $(\lambda_r - \delta, \lambda_r + \delta) \subset \rho(T)$. If $\zeta \in \mathbb{R}$ and $|\zeta| > 1/\delta$, then $\zeta \in \rho((T-\lambda_r)^{-1})$ since

$$((T-\lambda_r)^{-1} - \zeta)^{-1} = \zeta^{-1}(T-\lambda_r)(\zeta^{-1} + \lambda_r - T)^{-1}.$$

Hence $\sigma((T-\lambda_r)^{-1}) \subset [-1/\delta, 1/\delta]$ and Theorem 2.6.4 implies that

$$\|(T-\lambda_r)^{-1}\| \leq 1/\delta. \tag{2.38}$$

Since $(T-\lambda_r)(T-\lambda_r)^{-1} = I$ and $H \neq \{0\}$, we have that $\|(T-\lambda_r)^{-1}\| \neq 0$ and hence not every $\delta \in (0, \infty)$ is such that $(\lambda_r - \delta, \lambda_r + \delta) \subset \rho(T)$. Let us take the largest possible δ, hence, either $\lambda_r + \delta$ or $\lambda_r - \delta$ belongs to $\sigma(T)$ and $\operatorname{dist}(\lambda, \sigma(T))^2 = \delta^2 + \lambda_i^2$. (2.38) implies that $\|(T-\lambda_r)x\| \geq \delta\|x\|$ for all $x \in \mathcal{D}(T)$ and hence (2.37) implies $\|(T-\lambda)x\|^2 \geq (\delta^2 + \lambda_i^2)\|x\|^2$. Therefore $\|(T-\lambda)^{-1}\| \leq \operatorname{dist}(\lambda, \sigma(T))^{-1}$ also in this case.

The reverse inequality follows from Theorem 1.6.11 (see Exercise 18 in Chapter 1). □

Theorem 2.6.6 *If T is a self-adjoint operator and $\lambda \in \mathbb{R}$, then $(-\infty, \lambda) \subset \rho(T)$ if and only if*

$$(Tx, x) \geq \lambda\|x\|^2 \quad \textit{for all} \quad x \in \mathcal{D}(T).$$

PROOF If $(-\infty, \lambda) \subset \rho(T)$, then Theorem 2.6.5 implies

$$\|(T-\lambda-\zeta)^{-1}\| \leq 1/|\operatorname{Re}\zeta| \quad \text{for} \quad \zeta \in \mathbb{K},\ \operatorname{Re}\zeta < 0$$

and hence Theorem 2.3.2 implies that $T-\lambda$ is m-accretive.

Suppose now that $T-\lambda$ is accretive and that $\zeta \in (-\infty, \lambda)$. Note that

$$(\lambda - \zeta)\|x\|^2 \leq ((T-\zeta)x, x) \leq \|(T-\zeta)x\|\|x\| \quad \text{for} \quad x \in \mathcal{D}(T)$$

and, since T is closed, we have that $\mathcal{R}(T-\zeta)$ is closed. If $y \in \mathcal{R}(T-\zeta)^{\perp}$, then $(Tx, y) = (x, \zeta y)$ for all $x \in \mathcal{D}(T)$, hence $Ty = \zeta y$ and, since $T-\zeta$ is one-to-one, we have that $y = 0$ and therefore $\mathcal{R}(T-\zeta) = H$ by Theorem 2.2.2. Corollary 1.6.2 implies that $\zeta \in \rho(T)$. □

Spectral representation for self-adjoint compact operators:

Theorem 2.6.7 (Hilbert-Schmidt) *Suppose H is a Hilbert space and $K \in \mathfrak{B}(H)$ is symmetric and compact.*

For each $\lambda \in \sigma_p(K)\backslash\{0\}$, define $n(\lambda) = \dim \mathcal{N}(K-\lambda)$ and choose an orthonormal set $\{\phi_{\lambda 1}, \ldots, \phi_{\lambda n(\lambda)}\}$ so that its span is equal to $\mathcal{N}(K-\lambda)$. (See Corollary 1.7.5, Theorem 1.7.14.)

Then

$$Kx = \sum_{\lambda \in \sigma_p(K)\backslash\{0\}} \sum_{i=1}^{n(\lambda)} \lambda(x, \phi_{\lambda i})\phi_{\lambda i} \quad \textit{for all} \quad x \in H.$$

PROOF Choose any $\varepsilon > 0$. Let $\Delta = \{\lambda \in \sigma_p(K) \,|\, |\lambda| > \varepsilon\}$ and

$$Rx = Kx - \sum_{\lambda \in \Delta} \sum_{i=1}^{n(\lambda)} \lambda(x, \phi_{\lambda i})\phi_{\lambda i} \quad \text{for all} \quad x \in H.$$

Clearly, $R \in \mathfrak{B}(H)$ and it can be easily seen that R is symmetric. We shall show that $\|R\| \leq \varepsilon$. Assume $R \neq 0$. Theorem 2.6.4 implies that $\|R\| = |\mu| \neq 0$ for some $\mu \in \sigma(R)$ and, since R is compact, see Theorem 1.7.4, we have, by Theorem 1.7.12, that μ is an eigenvalue of R. Thus, for some $z \in H$, $z \neq 0$,

$$\mu z = Kz - \sum_{\lambda \in \Delta} \sum_{i=1}^{n(\lambda)} \lambda(z, \phi_{\lambda i})\phi_{\lambda i}.$$

This and the fact that for all $\lambda \in \Delta$, $i = 1, \ldots, n(\lambda)$,

$$\mu(z, \phi_{\lambda i}) = (Kz, \phi_{\lambda i}) - \lambda(z, \phi_{\lambda i}) = (z, K\phi_{\lambda i}) - \lambda(z, \phi_{\lambda i}) = 0$$

imply that $\mu z = Kz$, and since $z \in \mathcal{N}(K-\mu) = \operatorname{span}\{\phi_{\mu 1}, \ldots, \phi_{\mu n(\mu)}\}$, we have to have that $\mu \notin \Delta$. Therefore $\|R\| = |\mu| \leq \varepsilon$. □

The following spectral representation for self-adjoint operators with compact resolvent will play a key role in many of the examples in the rest of the text.

Theorem 2.6.8 *Suppose that T is a symmetric linear operator with compact resolvent in a Hilbert space H.*

For each $\lambda \in \sigma_p(T)$, define $n(\lambda) = \dim \mathcal{N}(T-\lambda)$ and choose an orthonormal set $\{\phi_{\lambda 1}, \ldots, \phi_{\lambda n(\lambda)}\}$ so that its span is equal to $\mathcal{N}(T-\lambda)$. (See Theorem 1.7.16.)

Then

(a) $\{\phi_{\lambda i} \,|\, \lambda \in \sigma_p(T), i = 1, \ldots, n(\lambda)\}$ is a complete orthonormal set in H

(b) T is a self-adjoint operator

(c) $x \in \mathcal{D}(T)$ if and only if $\sum_{\lambda \in \sigma_p(T)} \sum_{i=1}^{n(\lambda)} |\lambda(x, \phi_{\lambda i})|^2 < \infty$

(d) $Tx = \sum_{\lambda\in\sigma_p(T)} \sum_{i=1}^{n(\lambda)} \lambda(x,\phi_{\lambda i})\phi_{\lambda i}$ *for all* $x \in \mathcal{D}(T)$.

PROOF Theorem 1.7.16 implies that we can pick $\lambda_0 \in \mathbb{R} \cap \rho(T)$ and that $K \equiv (T-\lambda_0)^{-1}$ is compact and, by Lemma 2.6.1, symmetric. Lemma 2.6.1 implies also (b). If $\mu \in \sigma_p(K)\backslash\{0\}$ and $\lambda = \lambda_0 + 1/\mu$, then

$$\mathcal{N}(K-\mu) = \mathcal{N}(T-\lambda) = \text{span}\{\phi_{\lambda 1}, \ldots, \phi_{\lambda n(\lambda)}\}$$

and therefore Theorem 2.6.7 implies

$$Kx = \sum_{\lambda\in\sigma_p(T)} \sum_{i=1}^{n(\lambda)} \frac{(x,\phi_{\lambda i})}{\lambda-\lambda_0}\phi_{\lambda i} \quad \text{for all} \quad x \in H. \tag{2.39}$$

Thus, if $(x,\phi_{\lambda i}) = 0$ for all $\lambda \in \sigma_p(T)$, $i = 1,\ldots,n(\lambda)$, then $Kx = 0$ and hence $x = (T-\lambda_0)Kx = 0$. This proves (a).

If $x \in \mathcal{D}(T)$, then $\lambda(x,\phi_{\lambda i}) = (x, T\phi_{\lambda i}) = (Tx,\phi_{\lambda i})$ and, by Theorem 2.1.6,

$$\sum_{\lambda\in\sigma_p(T)} \sum_{i=1}^{n(\lambda)} |\lambda(x,\phi_{\lambda i})|^2 = \sum_{\lambda\in\sigma_p(T)} \sum_{i=1}^{n(\lambda)} |(Tx,\phi_{\lambda i})|^2 = \|Tx\|^2.$$

If $\sum_{\lambda\in\sigma_p(T)} \sum_{i=1}^{n(\lambda)} |\lambda(x,\phi_{\lambda i})|^2 < \infty$, then let $y = \sum_{\lambda\in\sigma_p(T)} \sum_{i=1}^{n(\lambda)} \lambda(x,\phi_{\lambda i})\phi_{\lambda i}$. Since $(y-\lambda_0 x,\phi_{\lambda i}) = (\lambda-\lambda_0)(x,\phi_{\lambda i})$, equation (2.39) gives

$$K(y-\lambda_0 x) = \sum_{\lambda\in\sigma_p(T)} \sum_{i=1}^{n(\lambda)} (x,\phi_{\lambda i})\phi_{\lambda i} = x$$

and therefore $x \in \mathcal{D}(T)$ and $(T-\lambda_0)x = y - \lambda_0 x$. This proves both (c) and (d). □

EXAMPLE 2.6.9 Suppose $T : \mathcal{D}(T) \subset L^2(0,1) \to L^2(0,1)$ is given by

$$Tf = if' \quad \text{for} \quad f \in \mathcal{D}(T) \equiv \{f \in AC[0,1] \mid f' \in L^2(0,1),\ f(0) = f(1)\}.$$

One can easily check that T is symmetric and that

$$(T-\lambda)^{-1} = -iR(-i\lambda) \quad \text{for} \quad \lambda \notin \sigma(T) = \sigma_p(T) = \{0, \pm 2\pi, \pm 4\pi, \ldots\},$$

where R is the integral operator as in Example 1.6.4. Example 1.7.7 implies that $(T-\lambda)^{-1}$ is compact for $\lambda \in \rho(T)$ and therefore Theorem 2.6.8 implies that T is self-adjoint and that its eigenvectors

$$1\,,\ e^{2i\pi x}\,,\ e^{-2i\pi x}\,,\ e^{4i\pi x}\,,\ e^{-4i\pi x}\,,\ \ldots$$

form a complete orthonormal set in $L^2(0,1)$. This proof of completeness of the Fourier series is independent of the proof given in Example 2.1.7.

2.7 Example: Sturm-Liouville Problem

The problem of finding a scalar λ and a function $u \not\equiv 0$ such that

$$(ru')' + (q + \lambda p)u = 0 \quad \text{on} \quad [a, b]$$

$$\alpha_1 u(a) + \alpha_2 (ru')(a) = 0$$

$$\beta_1 u(b) + \beta_2 (ru')(b) = 0$$

is known as a **Sturm-Liouville problem**. It is assumed that

(i) $-\infty < a < b < \infty$ and $u, ru' \in AC[a, b]$

(ii) $p, q, r^{-1} \in L^1(a, b)$

(iii) $p(x), r(x) \in (0, \infty)$ for almost all $x \in [a, b]$, q is real valued

(iv) $\alpha_1, \alpha_2, \beta_1, \beta_2$ are real numbers with $|\alpha_1| + |\alpha_2| \neq 0$ and $|\beta_1| + |\beta_2| \neq 0$

(v) **one** of the following (va) or (vb) is true

(va) $\alpha_1 \alpha_2 \leq 0$, $\beta_1 \beta_2 \geq 0$, $q \leq C_0 p$ a.e. on $[a, b]$ for some $C_0 \in \mathbb{R}$

(vb) $r^{-1} \leq C_0 p$ a.e. on $[a, b]$ for some $C_0 < \infty$.

Let H denote the set of all equivalence classes of measurable functions f for which

$$\int_a^b p(x)|f(x)|^2 dx < \infty.$$

H is a Hilbert space with inner product

$$(f, g) = \int_a^b p(x) f(x) \overline{g(x)} dx.$$

Define a linear operator S in H as follows:

$$Su = -\frac{(ru')' + qu}{p} \qquad \text{for} \quad u \in \mathcal{D}(S),$$

where $\mathcal{D}(S)$ consists of all $u \in AC[a, b]$ such that $ru' \in AC[a, b]$ and

$$((ru')' + qu)/p \in H, \quad \alpha_1 u(a) + \alpha_2 (ru')(a) = \beta_1 u(b) + \beta_2 (ru')(b) = 0.$$

The Sturm-Liouville problem thus becomes the problem of finding eigenvalues and eigenvectors of the operator S.

Lemma 2.7.1 *S is symmetric.*

PROOF Choose $u, v \in \mathcal{D}(S)$ and let $w = u\overline{rv'} - ru'\overline{v} \in AC[a,b]$. Note

$$w' = u\overline{((rv')' + qv)} - ((ru')' + qu)\overline{v},$$

$$w(b) - w(a) = (Su, v) - (u, Sv).$$

Since the system

$$\alpha_1 u(a) + \alpha_2 (ru')(a) = 0$$
$$\alpha_1 \overline{v(a)} + \alpha_2 \overline{(rv')(a)} = 0$$

has a nontrivial solution α_1, α_2, we have $w(a) = 0$. Similarly, $w(b) = 0$. □

Lemma 2.7.2 *There exists $\theta \in \mathbb{R}$ such that $(Su, u) \geq \theta \|u\|^2$ for all $u \in \mathcal{D}(S)$.*

PROOF Note that for $u \in \mathcal{D}(S)$ we have

$(ru')(a)\overline{u(a)} = c_a |u(a)|^2$ where $c_a = -\alpha_1/\alpha_2$ if $\alpha_2 \neq 0$ and $c_a = 0$ otherwise

$(ru')(b)\overline{u(b)} = c_b |u(b)|^2$ where $c_b = -\beta_1/\beta_2$ if $\beta_2 \neq 0$ and $c_b = 0$ otherwise.

Therefore

$$\begin{aligned}(Su, u) &= [-ru'\overline{u}]_a^b + \int_a^b r|u'|^2 - q|u|^2 \\ &= c_a|u(a)|^2 - c_b|u(b)|^2 + \int_a^b r|u'|^2 - q|u|^2.\end{aligned}$$

Thus, if (va) holds, take $\theta = -C_0$. If (vb) holds let $c = |c_a| + |c_b| + \|q\|_1 + 1$ and note that for $u \in \mathcal{D}(S)$,

$$(Su, u) \geq -c\|u\|_\infty^2 + \int_a^b r|u'|^2.$$

However, $\int_a^b p|u|^2 = |u(y)|^2 \int_a^b p$ for some $y \in [a,b]$; hence, for all $x \in [a,b]$,

$$\begin{aligned}|u(x)|^2 &= |u(y)|^2 + 2\mathrm{Re} \int_y^x u'\overline{u} \\ &\leq \|u\|^2/\|p\|_1 + \int_a^b 2|u'||u| \\ &\leq \|u\|^2/\|p\|_1 + \int_a^b ((r/c)|u'|^2 + (c/r)|u|^2) \\ c\|u\|_\infty^2 &\leq c\|u\|^2/\|p\|_1 + \int_a^b r|u'|^2 + c^2 C_0 \int_a^b p|u|^2 \\ &\leq (c/\|p\|_1 + c^2 C_0)\|u\|^2 + \int_a^b r|u'|^2\end{aligned}$$

and therefore

$$(Su, u) \geq -(c/\|p\|_1 + c^2 C_0)\|u\|^2.$$

□

Lemma 2.7.3 *If $\lambda < \theta$, then $\lambda \in \rho(S)$ and $(S - \lambda)^{-1}$ is compact.*

PROOF If $f \in H$, then $pf \in L^1(a, b)$ and the statement $u \in \mathcal{D}(S)$ is such that $(S - \lambda)u = f$ is equivalent to the statement $u \in AC[a, b]$ is such that $ru' \in AC[a, b]$ and

$$\left.\begin{array}{c} (ru')' + (q + \lambda p)u = -pf \quad \text{a.e. on } [a, b] \\ \alpha_1 u(a) + \alpha_2 (ru')(a) = 0 \\ \beta_1 u(b) + \beta_2 (ru')(b) = 0 \end{array}\right\} \tag{2.40}$$

Thus, if u is a solution of the homogeneous version of (2.40), then $Su = \lambda u$ and hence $(Su, u) = \lambda\|u\|^2$, but by Lemma 2.7.2 $(Su, u) \geq \theta\|u\|^2$. Therefore u has to be the trivial solution 0. In view of Example 1.8.5 we have that $S - \lambda$ is one-to-one and onto.

Let $\{f_k\}$ be a bounded sequence in H and let $u_k = (S - \lambda)^{-1} f_k$. As in Example 1.8.5, $\{u_k\}$ is a uniformly bounded sequence and has a subsequence convergent in $L^1(a, b)$. Hence we can choose a subsequence $\{u_{k_i}\}$ which converges a.e. to some $u \in L^\infty(a, b)$. The Dominated Convergence Theorem (DCT) implies that $\{u_{k_i}\}$ converge in H to u. This implies that $(S - \lambda)^{-1}$ is bounded and compact and that $\lambda \in \rho(S)$. □

Theorem 2.7.4 *S is a self-adjoint operator,*

$$\sigma(S) = \sigma_p(S) = \{\lambda_1, \lambda_2, \cdots\},$$

for some $\lambda_1 < \lambda_2 < \cdots$ with $\lim_{n\to\infty} \lambda_n = \infty$ and $\dim \mathcal{N}(S - \lambda_n) = 1$ for all $n \geq 1$. Moreover, if $\phi_n \in \mathcal{N}(S - \lambda_n)$ are chosen so that $\|\phi_n\| = 1$ for $n \geq 1$, then $\{\phi_1, \phi_2, \cdots\}$ is a complete orthonormal set in H.

PROOF Lemmas 2.7.1, 2.7.3 and Theorem 2.6.8 imply that S is self-adjoint and that the normalized eigenvectors of S form a complete orthonormal system in H. If $\sigma_p(S)$ would be a finite set, then H would have to be spanned by finitely many vectors because, by Theorem 1.7.16, $n(\lambda) = \dim \mathcal{N}(S - \lambda) < \infty$ for each $\lambda \in \sigma_p(S)$. Therefore $\sigma_p(S)$ is an infinite set. By Theorem 1.7.16, $\sigma(S) = \sigma_p(S)$, and also Lemma 2.7.2, gives that $\sigma_p(S) = \{\lambda_1, \lambda_2, \cdots\}$ for some $\theta \leq \lambda_1 < \lambda_2 < \cdots$ with $\lim_{n\to\infty} \lambda_n = \infty$.

To show $\dim \mathcal{N}(S-\lambda_n)=1$, suppose $u, v \in \mathcal{N}(S-\lambda_n)$. The system

$$x_1 u(a) + x_2 (ru')(a) = 0$$

$$x_1 v(a) + x_2 (rv')(a) = 0$$

has a nontrivial solution $x_1 = \alpha_1$, $x_2 = \alpha_2$ hence there should exist c_1, c_2, not both zero, such that

$$c_1 u(a) + c_2 v(a) = 0$$

$$c_1 (ru')(a) + c_2 (rv')(a) = 0.$$

Thus, if $w = c_1 u + c_2 v$, then

$$(rw')' + (q + \lambda_n p) w = 0, \quad w(a) = (rw')(a) = 0$$

and hence the uniqueness for the initial value problem (Example 1.6.9) implies $w = 0$ and therefore $\mathcal{N}(S-\lambda_n)$ cannot contain two linearly independent vectors. □

EXAMPLE 2.7.5 In engineering applications one often encounters Sturm-Liouville problems similar to the following one:

$$u''(x) + \lambda u(x) = 0 \quad \text{for all} \quad x \in [0,1]$$

$$u(0) = 0, \quad u(1) + h u'(1) = 0$$

where $h \geq 0$. One can easily see that the eigenvectors u_n and the eigenvalues λ_n are given by

$$u_n(x) = \sin(\alpha_n x), \quad \lambda_n = \alpha_n^2, \quad n \geq 1$$

where α_n is the nth positive solution of equation

$$\tan \alpha = -h\alpha.$$

Theorem 2.7.4 says that normalized u_n, $n \geq 1$, form a complete orthonormal set in $L^2(0,1)$. Therefore, for all $f \in L^2(0,1)$,

$$\lim_{n\to\infty} \|f - \sum_{j=1}^{n} c_j u_j\|_2 = 0$$

where

$$c_j = \frac{2}{1 + h\cos^2 \alpha_j} \int_0^1 f u_j \quad \text{for all} \quad j \geq 1$$

and the Parseval equality, see Theorem 2.1.6, implies

$$2\int_0^1 |f|^2 = \sum_{j=1}^{\infty} |c_j|^2 (1 + h\cos^2 \alpha_j).$$

Selecting $h = 0$ gives an independent proof of (2.12).

2.8 Sectorial Forms

Hypotheses **H**1, **H**2, **H**3 will be in effect throughout this section.

(**H**1) $\mathcal{H}$ is a Hilbert space with an inner product denoted by $(\cdot,\cdot)$ and the corresponding norm by $\|\cdot\|$.

(**H**2) $\mathcal{V}$ is a dense subspace of $\mathcal{H}$. $\mathcal{V}$ is also a Hilbert space with an inner product $[\cdot,\cdot]$ and the corresponding norm $|\cdot|$. There exists $M_1 \in (0,\infty)$ such that

$$\|x\| \le M_1|x| \quad \text{for all} \quad x \in \mathcal{V}.$$

(**H**3) $\mathfrak{F} : \mathcal{V} \times \mathcal{V} \to \mathbb{K}$ is such that

$$\mathfrak{F}(\alpha x + \beta y, z) = \alpha\mathfrak{F}(x,z) + \beta\mathfrak{F}(y,z)$$

$$\mathfrak{F}(z, \alpha x + \beta y) = \overline{\alpha}\mathfrak{F}(z,x) + \overline{\beta}\mathfrak{F}(z,y)$$

for all vectors x, y, z in $\mathcal{V}$ and all scalar numbers α, β in $\mathbb{K}$. Moreover, there exist $a \in \mathbb{R}$, $M_2 \in (0,\infty)$, $M_3 \in (0,\infty)$ such that

$$|\mathfrak{F}(x,y)| \le M_2|x||y| \quad \text{for all} \quad x, y \in \mathcal{V}$$

$$\operatorname{Re}\mathfrak{F}(x,x) \ge M_3|x|^2 + a\|x\|^2 \quad \text{for all} \quad x \in \mathcal{V}.$$

Such $\mathfrak{F}$ are called **sectorial sesquilinear forms** and are obviously continuous. The adjective 'sectorial' is due to the fact that the numerical range of $\mathfrak{F}$ lies in a sector: if $x \in \mathcal{V}$ and $\|x\| = 1$, then $|x| \ge M_1^{-1}$ hence $\operatorname{Re}\mathfrak{F}(x,x) \ge a + M_3M_1^{-2}$ and

$$|\operatorname{Im}\mathfrak{F}(x,x)| \le M_2|x|^2 \le (M_2/M_3)(\operatorname{Re}\mathfrak{F}(x,x) - a),$$

therefore

$$\mathfrak{F}(x,x) \in \Gamma \equiv \{\, \zeta \in \mathbb{K} \mid \operatorname{Re}\zeta \ge a + M_3M_1^{-2} \text{ and } |\operatorname{Im}\zeta| \le M_2(\operatorname{Re}\zeta - a)/M_3 \,\}.$$

Define a linear operator A in $\mathcal{H}$ as follows. $x \in \mathcal{D}(A)$ if and only if $x \in \mathcal{V}$ and there exists $z \in \mathcal{H}$ such that

$$\mathfrak{F}(x,y) = (z,y) \quad \text{for all} \quad y \in \mathcal{V}. \tag{2.41}$$

Since $\mathcal{V}$ is dense in $\mathcal{H}$ we have, for each $x \in \mathcal{D}(A)$, only one $z \in \mathcal{H}$ such that (2.41) holds. Define $Ax = z$. Thus, $\mathfrak{F}$ has the following representation:

$$\mathfrak{F}(x,y) = (Ax,y) \quad \text{for all} \quad x \in \mathcal{D}(A),\ y \in \mathcal{V}.$$

A is called **the operator associated with** $\mathfrak{F}$. Note that if $x \in \mathcal{D}(A)$ and $\|x\| = 1$, then $(Ax,x) = \mathfrak{F}(x,x) \in \Gamma$. Hence, the numerical range of A is contained in Γ.

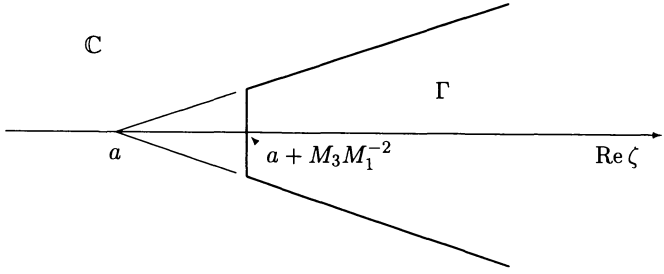


Figure 2.1: Γ containing numerical ranges of $\mathfrak{F}$ and A.

EXAMPLE 2.8.1 Let the Hilbert space $\mathcal{H}$ be $L^2(0,1)$ with the usual inner product and let $\mathcal{V}$ consist of $u \in AC[0,1]$ for which $u' \in L^2(0,1)$ and $u(0) = u(1) = 0$. Since $C_0^1(0,1)$ is dense in $\mathcal{H}$, (Example 1.3.4) so is $\mathcal{V}$. Define

$$[u,v] = \int_0^1 u'(x)\overline{v'(x)}dx, \quad |u| = [u,u]^{1/2} \quad \text{for} \quad u,v \in \mathcal{V}.$$

Since $u(x) = \int_0^x u'$ for $u \in \mathcal{V}$, we have that $\|u\| \le \|u\|_\infty \le |u|$; thus, $[\cdot,\cdot]$ is an inner product on $\mathcal{V}$ and it can be easily seen that $\mathcal{V}$ is complete and hence a Hilbert space.

Define $\mathfrak{F}(u,v) = [u,v]$ for $u,v \in \mathcal{V}$ and note that the assumptions **H1**, **H2**, **H3** are satisfied with $M_1 = M_2 = M_3 = 1$, $a = 0$. If $u \in \mathcal{D}(A)$, then

$$\int_0^1 u'\phi' = \int_0^1 (Au)\phi \quad \text{for all} \quad \phi \in C_0^1(0,1).$$

Hence Lemma 2.4.5 implies that $u' \in AC[0,1]$, $Au = -u''$ and therefore

$$\mathcal{D}(A) \subset \mathcal{V}_1 \equiv \{\, u \in AC[0,1] \mid u' \in AC[0,1],\ u'' \in L^2(0,1),\ u(0) = u(1) = 0 \,\}.$$

An integration by parts implies that $\mathcal{V}_1 \subset \mathcal{D}(A)$ and hence $\mathcal{D}(A) = \mathcal{V}_1$.

In Section 3.7 it is shown that strongly elliptic operators can be obtained in this way. Thus, $u_t + Au = 0$ becomes a generalized parabolic equation which is studied in Section 4.5 and where it is shown, in particular, that the initial value problem for $u_t + Au = 0$ always has a solution. With one more assumption on $\mathfrak{F}$, the initial value problem for the generalized wave equation $u_{tt} + Au = 0$ is solved in Section 4.4. Some ODE applications will be presented in the following sections.

The following Theorem presents the basic properties of A; its converse is given by Theorem 2.12.1.

Theorem 2.8.2 *The linear operator A is closed, densely defined in $\mathcal{H}$ and it has the following properties:*

(1) $\sigma(A) \subset \Gamma$ and $\|(A-\lambda)^{-1}\| \leq 1/\mathrm{dist}(\lambda, \Gamma)$ for $\lambda \in \mathbb{K}\backslash\Gamma$

(2) if the identity map $I : \mathcal{V} \to \mathcal{H}$ is compact, then A has compact resolvent

(3) $\mathcal{D}(A)$ is also dense in $\mathcal{V}$ (in $|\cdot|$ norm).

PROOF Choose $x \in \mathcal{V}$ and define $f_x(y) = \overline{\mathfrak{F}(x,y) - a(x,y)}$ for $y \in \mathcal{V}$. It can be easily verified that f_x is a bounded linear functional on $\mathcal{V}$ and hence, by Lemma 2.2.4, there exists a unique $Bx \in \mathcal{V}$ such that $f_x(y) = [y, Bx]$ for all $y \in \mathcal{V}$. Therefore

$$\mathfrak{F}(x,y) - a(x,y) = [Bx, y] \quad \text{for all} \quad x, y \in \mathcal{V}. \tag{2.42}$$

The linearity of $\mathfrak{F}(\cdot, y)$ implies linearity of B and since $|[Bx, y]| \leq (M_2 + |a|M_1^2)|x||y|$, we have that $B \in \mathfrak{B}(\mathcal{V})$. Note that $\mathrm{Re}\,[Bx, x] \geq M_3|x|^2$ for $x \in \mathcal{V}$, hence, Corollary 2.3.3 implies

$$|B^{-1}x| \leq |x|/M_3 \quad \text{for all} \quad x \in \mathcal{V}. \tag{2.43}$$

Choose $w \in \mathcal{H}$ and let $g(y) = (y, w)$ for $y \in \mathcal{V}$. Since $|g(y)| \leq M_1|y|\|w\|$, we have, by Lemma 2.2.4, that there exists a unique $v \in \mathcal{V}$ such that $(y, w) = [y, v]$ for all $y \in \mathcal{V}$, moreover, $|v| \leq M_1\|w\|$. Thus $x = B^{-1}v$ is the unique element of $\mathcal{V}$ which satisfies $(y, w) = [y, Bx]$ for all $y \in \mathcal{V}$ and, in view of (2.42), this x is the unique solution of

$$\mathfrak{F}(x,y) - a(x,y) = (w, y) \quad \text{for all} \quad y \in \mathcal{V}.$$

Moreover, (2.43) implies $|x| \leq M_1\|w\|/M_3$. By the definition, $x \in \mathcal{D}(A)$ and $Ax = ax + w$. Thus, $A - a$ is onto and since the numerical range of A lies in Γ, Theorem 1.6.16 implies (1) and that A is closed. Since we have actually shown that $(A-a)^{-1} \in \mathfrak{B}(\mathcal{H}, \mathcal{V})$, the compactness of the identity map $I : \mathcal{V} \to \mathcal{H}$ would imply that $I(A-a)^{-1} \in \mathfrak{B}(\mathcal{H})$ is compact - which proves (2). To show (3), suppose that $u \in \mathcal{V}$ is such that $[z, u] = 0$ for all $z \in \mathcal{D}(A)$. Let $w = B'u$ where B' is the Hilbert space $(\mathcal{V})$ adjoint of B^{-1}; hence

$$\|w\|^2 = (w, B'u) = [v, B'u] = [B^{-1}v, u] = [x, u] = 0,$$

$[u, u] = [Bu, B'u] = [Bu, w] = 0$ proving (3) - which implies that $\mathcal{D}(A)$ is dense also in $\mathcal{H}$. □

EXAMPLE 2.8.3 Let us continue with Example 2.8.1. Since $\|u\|_\infty \leq |u|$ and

$$|u(x) - u(y)| = \left| \int_y^x u'(s)ds \right| \leq |x - y|^{1/2}|u| \quad \text{for} \quad u \in \mathcal{V},$$

the Arzela-Ascoli Theorem 1.1.5 implies that if $\{u_k\}$ is a bounded sequence in $\mathcal{V}$, then a subsequence converges uniformly to some $u \in C[0,1]$ which implies that the identity map from $\mathcal{V}$ to $\mathcal{H}$ is compact. Therefore A has compact resolvent by Theorem 2.8.2.

If $x \in \mathcal{D}(A)$ and $Ax = z$, then obviously

$$(x, A^*y) = (z, y) \quad \text{for all} \quad y \in \mathcal{D}(A^*)$$

and since $A^{**} = A$, (Theorem 2.2.7) the converse is also true. To obtain A^*, define $\mathfrak{F}_1(x,y) = \overline{\mathfrak{F}(y,x)}$ for $x, y \in \mathcal{V}$ and note that hypothesis **H3** is satisfied, with the same constants, when $\mathfrak{F}$ is replaced by $\mathfrak{F}_1$. Let A_1 be the operator associated with $\mathfrak{F}_1$. Thus, $x \in \mathcal{D}(A_1)$ if and only if $x \in \mathcal{V}$ and there exists $z \in \mathcal{H}$ such that

$$\mathfrak{F}(y,x) = (y,z) \quad \text{for all} \quad y \in \mathcal{V}.$$

Moreover, $A_1 x = z$. Note that Theorem 2.8.2 remains true if A is replaced by A_1.

Theorem 2.8.4 $A_1 = A^*$.

PROOF If $x \in \mathcal{D}(A_1), y \in \mathcal{D}(A)$, then $(Ay, x) = \mathfrak{F}(y,x) = (y, A_1 x)$; therefore, $\mathcal{D}(A_1) \subset \mathcal{D}(A^*)$ and $A_1 x = A^* x$ for $x \in \mathcal{D}(A_1)$. Theorems 2.8.2 and 2.2.6 imply $a \in \rho(A_1) \cap \rho(A^*)$ hence $A_1 = A^*$ by Lemma 1.6.14. □

Corollary 2.8.5 *If* $\mathfrak{F}(x,y) = \overline{\mathfrak{F}(y,x)}$ *for all* $x, y \in \mathcal{V}$, *then* A *is self-adjoint.*

Note that

$$\mathcal{N}(A^*) = \{x \in \mathcal{V} \mid \mathfrak{F}(y,x) = 0 \text{ for all } y \in \mathcal{V}\}.$$

So, when the identity map $I : \mathcal{V} \to \mathcal{H}$ is compact, A has compact resolvent hence its range is closed by Theorem 1.7.16 and therefore the Fredholm alternative, Theorem 2.2.8, implies that, for each $z \in \mathcal{H}$, the following statements are equivalent:

(a) there exists $x \in \mathcal{V}$ such that $\mathfrak{F}(x,y) = (z,y)$ for all $y \in \mathcal{V}$

(b) $z \in \mathcal{R}(A)$

(c) $(z,y) = 0$, whenever $y \in \mathcal{V}$ is such that $\mathfrak{F}(w,y) = 0$ for all $w \in \mathcal{V}$.

When A has compact resolvent Theorem 1.7.16 implies also that $\sigma(A) = \sigma_p(A)$ and hence uniqueness implies existence, i.e. if $\mathcal{N}(A) = \{0\}$, then $\mathcal{R}(A) = \mathcal{H}$.

Define a sectorial sesquilinear form $\mathfrak{F}_r$ by

$$\mathfrak{F}_r(x,y) = \frac{\mathfrak{F}(x,y) + \overline{\mathfrak{F}(y,x)}}{2} \quad \text{for} \quad x, y \in \mathcal{V}$$

and note that the hypothesis **H3** is satisfied when $\mathfrak{F}$ is replaced by $\mathfrak{F}_r$. Let A_r be the operator associated with $\mathfrak{F}_r$. Note that Theorem 2.8.2 remains true if A is replaced by A_r. Corollary 2.8.5 implies that A_r is self-adjoint. If $\mathfrak{F}(x,y) = \overline{\mathfrak{F}(y,x)}$ for all $x, y \in \mathcal{V}$, then $A = A_r$.

Theorem 2.8.6 *If for each $x \in \mathcal{V}$ there exists $c < \infty$ such that*

$$|\mathfrak{F}(x,y) - \overline{\mathfrak{F}(y,x)}| \leq c\|y\| \quad \textit{for all} \quad y \in \mathcal{V},$$

then $\mathcal{D}(A) = \mathcal{D}(A^*) = \mathcal{D}(A_r)$, $A_r = (A + A^*)/2$ and there exists $b < \infty$ such that

$$|\mathfrak{F}(x,y) - \overline{\mathfrak{F}(y,x)}| \leq 2b|x|\|y\| \quad \textit{for all} \quad x, y \in \mathcal{V}.$$

PROOF Theorem 1.4.3 and Lemma 2.2.4 imply that for each $x \in \mathcal{V}$ there exists a unique $Tx \in \mathcal{H}$ such that $\mathfrak{F}(x,y) - \overline{\mathfrak{F}(y,x)} = (Tx, y)$ for all $y \in \mathcal{V}$. T is obviously a linear map from $\mathcal{V}$ to $\mathcal{H}$ and it is easy to check that it is also closed and hence bounded by Theorem 1.4.11. This gives b.

If $x \in \mathcal{D}(A)$, then $\mathfrak{F}(y,x) = (y, Ax - Tx)$ for all $y \in \mathcal{V}$ hence $x \in \mathcal{D}(A^*)$.

If $x \in \mathcal{D}(A^*)$, then $\mathfrak{F}(x,y) = (A^*x + Tx, y)$ for all $y \in \mathcal{V}$ hence $x \in \mathcal{D}(A)$, and since $2\mathfrak{F}_r(x,y) = (Ax,y) + (A^*x,y)$ we also have that $x \in \mathcal{D}(A_r)$ and $2A_rx = Ax + A^*x$.

If $x \in \mathcal{D}(A_r)$, then $\mathfrak{F}(x,y) + \overline{\mathfrak{F}(y,x)} = 2(A_rx, y)$ and hence $2\mathfrak{F}(x,y) = 2(A_rx,y) + (Tx,y)$ for all $y \in \mathcal{V}$; therefore, $x \in \mathcal{D}(A)$. □

Theorem 2.8.7 *If $\mathcal{V} \neq \{0\}$ and λ_0 is the largest of real numbers λ such that*

$$\operatorname{Re}\mathfrak{F}(x,x) \geq \lambda\|x\|^2 \quad \textit{for all} \quad x \in \mathcal{V},$$

then $\lambda_0 \in \sigma(A_r)$, $\sigma(A_r) \subset [\lambda_0, \infty)$. Moreover, if $\mu \leq \lambda_0$ and $c = \min\{1, \frac{\lambda_0 - \mu}{\lambda_0 - a}\}$, then

$$\operatorname{Re}\mathfrak{F}(x,x) \geq cM_3|x|^2 + \mu\|x\|^2 \quad \textit{for all} \quad x \in \mathcal{V}.$$

PROOF It is easy to see that λ_0 exists in $\mathbb{R}$ and that $\mathfrak{F}_r(x,x) \geq \lambda_0\|x\|^2$ for all $x \in \mathcal{V}$. Hence $(A_rx, x) \geq \lambda_0\|x\|^2$ for all $x \in \mathcal{D}(A_r)$ and Theorem 2.6.6 implies that $(-\infty, \lambda_0) \in \rho(A_r)$. If $\lambda_0 \notin \sigma(A_r)$, then $(-\infty, \lambda_1) \in \rho(A_r)$ for some $\lambda_1 > \lambda_0$ and Theorem 2.6.6 would imply that

$$\mathfrak{F}_r(x,x) = (A_rx, x) \geq \lambda_1\|x\|^2 \quad \text{for all} \quad x \in \mathcal{D}(A_r),$$

and since $\mathcal{D}(A_r)$ is dense in $\mathcal{V}$, this would contradict the definition of λ_0.

A simple calculation shows that

$$\mathfrak{F}_r(x,x) - \mu\|x\|^2 \geq c(\mathfrak{F}_r(x,x) - a\|x\|^2) \geq cM_3|x|^2 \quad \text{for all} \quad x \in \mathcal{V}$$

which proves the 'moreover' part. □

EXAMPLE 2.8.8 Corollary 2.8.5 implies that A from Examples 2.8.1 and 2.8.3 is self-adjoint. A short calculation gives that the eigenvalues of A are $(n\pi)^2$ and the corresponding eigenvectors are $\phi_n(x) = \sqrt{2}\sin n\pi x$ where $n = 1, 2, \ldots$ Theorems 1.7.16 and 2.8.7 imply that for all $u \in AC[0,1]$ with $u(0) = u(1) = 0$, we have that

$$\int_0^1 |u'(s)|^2 ds \geq \pi^2 \int_0^1 |u(s)|^2 ds.$$

The following Theorem 2.8.9 is the reason why the setting of this Section is sometimes called **variational formulation**.

Theorem 2.8.9 *Suppose* $\lambda \in (-\infty, a + M_3 M_1^{-2})$, $f \in \mathcal{H}$. *Define* $I : \mathcal{V} \to \mathbb{R}$ *by*

$$I(y) = \mathrm{Re}\,\mathfrak{F}(y,y) - \lambda\|y\|^2 - 2\mathrm{Re}\,(f,y) \quad \textit{for all} \quad y \in \mathcal{V}.$$

Then, $I(x) < I(y)$ *for all* $y \in \mathcal{V}\backslash\{x\}$, *where* $x = (A_r - \lambda)^{-1} f$.

PROOF For $y \in \mathcal{V}$, we have that

$$\begin{aligned}
(f, y-x) &= \mathfrak{F}_r(x, y-x) - \lambda(x, y-x) \\
I(y) - I(x) &= \mathfrak{F}_r(y,y) - \lambda\|y\|^2 - \mathfrak{F}_r(x,x) + \lambda\|x\|^2 - 2\mathrm{Re}\,(f, y-x) \\
&= \mathfrak{F}_r(x-y, x-y) - \lambda\|x-y\|^2 \\
&\geq M_3|x-y|^2 + (a-\lambda)\|x-y\|^2 \\
&\geq (M_3 M_1^{-2} + a - \lambda)\|x-y\|^2,
\end{aligned}$$

which proves the assertion of the Theorem. □

The following Lemma and Theorem are the basic tools for obtaining finite element approximations for PDEs. For an example, see Section 2.11.

Lemma 2.8.10 *Suppose that* $\mathcal{U}$ *is a closed (in* $|\cdot|$ *norm) subspace of* $\mathcal{V}$, $f \in \mathcal{H}$, $c \in [0, M_3)$, $\lambda \in \mathbb{K}$, $\mathrm{Re}\,\lambda \leq a + cM_1^{-2}$.

Then there exists a unique $x \in \mathcal{U}$ *such that* $\mathfrak{F}(x,z) = (\lambda x + f, z)$ *for all* $z \in \mathcal{U}$. *Moreover,* $\|x\| \leq \|f\|/(M_3 M_1^{-2} + a - \mathrm{Re}\,\lambda)$, $|x|^2 \leq \|f\|\|x\|/(M_3 - c)$ *and*

$$|x - (A-\lambda)^{-1} f| \leq \frac{M_2 + |\lambda| M_1^2}{M_3 - c} \inf_{z \in \mathcal{U}} |z - (A-\lambda)^{-1} f|. \tag{2.44}$$

PROOF Theorem 2.8.2 implies that there exists a unique $y \in \mathcal{V}$ such that $\mathfrak{F}(y,z) = (\lambda y + f, z)$ for all $z \in \mathcal{V}$. Moreover, $y = (A-\lambda)^{-1}f$.

Let $\overline{\mathcal{U}}$ be the closure of $\mathcal{U}$ in $\mathcal{H}$ and let $f = f_1 + f_2$ where $f_1 \in \overline{\mathcal{U}}$ and $(f_2, z) = 0$ for all $z \in \overline{\mathcal{U}}$ (see Theorem 2.2.2).

Hypotheses $\mathbf{H1}, \mathbf{H2}, \mathbf{H3}$ remain valid if $\mathcal{H}$ is replaced by $\overline{\mathcal{U}}$ and $\mathcal{V}$ by $\mathcal{U}$. Therefore Theorem 2.8.2 implies that there exists a unique $x \in \mathcal{U}$ such that $\mathfrak{F}(x,z) = (\lambda x + f_1, z)$ for all $z \in \mathcal{U}$. Since $(f_1, z) = (f, z)$ for all $z \in \overline{\mathcal{U}}$ we have the existence and uniqueness of x; since $\|f_1\| \le \|f\|$ we also have, from Theorem 2.8.2, the bound on $\|x\|$. The bound on $|x|$ follows from

$$\begin{aligned}(M_3 - c)|x|^2 &\le M_3|x|^2 - cM_1^{-2}\|x\|^2 \le M_3|x|^2 + (a - \operatorname{Re}\lambda)\|x\|^2 \\ &\le \operatorname{Re}(\mathfrak{F}(x,x) - \lambda(x,x)) = \operatorname{Re}(f,x) \le \|f\|\|x\|.\end{aligned}$$

Observe that for all $z \in \mathcal{U}$ we have $\mathfrak{F}(x-y, x-z) = \lambda(x-y, x-z)$,

$$\begin{aligned}(M_3 - c)|x-y|^2 &\le M_3|x-y|^2 + (a - \operatorname{Re}\lambda)\|x-y\|^2 \\ &\le \operatorname{Re}(\mathfrak{F}(x-y, x-y) - \lambda(x-y, x-y)) \\ &= \operatorname{Re}(\mathfrak{F}(x-y, z-y) - \lambda(x-y, z-y)) \\ &\le (M_2 + |\lambda|M_1^2)|x-y||z-y|,\end{aligned}$$

and since $z \in \mathcal{U}$ is arbitrary, inequality (2.44) follows. □

Theorem 2.8.11 *Suppose $\mathcal{V}_1, \mathcal{V}_2, \cdots$ are finite dimensional subspaces of $\mathcal{V}$ such that*

$$\lim_{n\to\infty} \inf_{z\in\mathcal{V}_n} |y - z| = 0 \tag{2.45}$$

for all y in a dense (in $|\cdot|$ norm) subset of $\mathcal{V}$. Choose any $f \in \mathcal{H}$ and $\lambda \in \mathbb{K}$ with $\operatorname{Re}\lambda < a + M_3M_1^{-2}$. Then for each $n \ge 1$, there exists a unique $x_n \in \mathcal{V}_n$ such that

$$\mathfrak{F}(x_n, z) = (\lambda x_n + f, z) \quad \textit{for all} \quad z \in \mathcal{V}_n. \tag{2.46}$$

Moreover,

$$\lim_{n\to\infty} |x_n - (A-\lambda)^{-1}f| = 0.$$

If, in addition, there exist $c_n \in [0,\infty)$ for $n \ge 1$ such that $\lim_{k\to\infty} c_k = 0$ and

$$\inf_{z\in\mathcal{V}_n} |y - z| \le c_n(\|Ay\| + \|y\|) \quad \textit{for} \quad n \ge 1, y \in \mathcal{D}(A), \tag{2.47}$$

then $(A-\lambda)^{-1}$ is a compact linear operator on $\mathcal{H}$ and there exists $b \in (0,\infty)$ which depends only on λ, a, M_1, M_2 and M_3, such that

$$|x_n - (A-\lambda)^{-1}f| \le c_n b\|f\| \quad \textit{for } n \ge 1. \tag{2.48}$$

PROOF Since (2.45) holds for y in a dense subset of $\mathcal{V}$, it holds for all y in $\mathcal{V}$. Except for compactness of $(A-\lambda)^{-1}$, everything follows immediately from Lemma 2.8.10. Lemma 2.8.10 also implies that $x_n = B_n f$ for some $B_n \in \mathfrak{B}(\mathcal{H}, \mathcal{V})$ and since ranges of B_n are finite dimensional, we have that B_n are compact linear operators on $\mathcal{H}$ and hence (2.48) and Theorem 1.7.4 imply that $(A-\lambda)^{-1}$ is compact on $\mathcal{H}$. □

Observe that the existence of error bounds, as in (2.48), with $c_k \to 0$ also implies that A has compact resolvent as well as estimates as in (2.47).

If λ is a scalar and $\operatorname{Re}\lambda < a + M_3 M_1^{-2}$, then Theorem 2.8.2 implies that $A-\lambda$ is m-accretive and that $0 \in \rho(A-\lambda)$, hence, Theorem 2.3.5 implies existence of $(A-\lambda)^{1/2}$, the unique m-accretive square root of $A-\lambda$. Define $G = (A_r - a)^{1/2}$. See also Corollary 6.1.14.

Theorem 2.8.12 *G is a self-adjoint, m-accretive operator in $\mathcal{H}$, $\mathcal{D}(G) = \mathcal{V}$, $0 \in \rho(G)$,*

$$M_3^{1/2}|x| \le \|Gx\| \le (M_2 + |a|M_1^2)^{1/2}|x| \quad \textit{for} \quad x \in \mathcal{V},$$

$$\mathfrak{F}_r(x,y) = (Gx, Gy) + a(x,y) \quad \textit{for} \quad x, y \in \mathcal{V}.$$

Moreover, there exists $B \in \mathfrak{B}(\mathcal{H})$ such that $B + B^* = 2 + 2aG^{-2}$, $A = GBG$ and

$$\mathfrak{F}(x,y) = (BGx, Gy) \quad \textit{for} \quad x, y \in \mathcal{V}.$$

PROOF Since G is symmetric, Lemma 2.6.1 implies that G is self-adjoint. If $x \in \mathcal{D}(A_r)$, then

$$\|Gx\|^2 = ((A_r - a)x, x) = \mathfrak{F}_r(x,x) - a\|x\|^2 \ge M_3|x|^2 \tag{2.49}$$

$$\|Gx\|^2 \le M_2|x|^2 - a\|x\|^2 \le (M_2 + |a|M_1^2)|x|^2. \tag{2.50}$$

If $x \in \mathcal{D}(G)$, then Theorem 2.3.5 implies existence of $x_n \in \mathcal{D}(A_r)$ such that $x_n \to x$ and $Gx_n \to Gx$. (2.49) implies that x_n form a Cauchy sequence in $\mathcal{V}$ hence $|x_n - y| \to 0$ for some $y \in \mathcal{V}$ and since $\|x_n - y\| \le M_1|x_n - y|$, we have that $x = y$ and therefore $\mathcal{D}(G) \subset \mathcal{V}$.

If $x \in \mathcal{V}$, then Theorem 2.8.2 implies existence of $x_n \in \mathcal{D}(A_r)$ such that $|x_n - x| \to 0$. (2.50) implies that Gx_n converge in $\mathcal{H}$ and, since G is closed, we have that $x \in \mathcal{D}(G)$. Thus $\mathcal{D}(G) = \mathcal{V}$. Taking the limit in

$$(Gx_n, Gy) = ((A_r - a)x_n, y) = \mathfrak{F}_r(x_n, y) - a(x_n, y) \quad \text{for} \quad y \in \mathcal{V}$$

proves that $\mathfrak{F}_r(x,y) = (Gx, Gy) + a(x,y)$ for $x, y \in \mathcal{V}$.

Define $\mathfrak{F}_2(x,y) = \mathfrak{F}(G^{-1}x, G^{-1}y)$ for $x, y \in \mathcal{H}$. Note that the hypotheses **H1**, **H2**, **H3** can be satisfied with $\mathcal{H}$ in place of $\mathcal{V}$ and $\mathfrak{F}_2$ in place of $\mathfrak{F}$. Let

B be the operator associated with $\mathfrak{F}_2$. $|\mathfrak{F}_2(x,y)| \le M_2 M_3^{-1}\|x\|\|y\|$ implies that $\mathcal{D}(B) = \mathcal{H}$ and $\|B\| \le M_2 M_3^{-1}$. Obviously $\mathfrak{F}(x,y) = (BGx, Gy)$ for $x, y \in \mathcal{V}$ which implies that $A = GBG$ and, since

$$2\mathfrak{F}_r(x,y) = ((B+B^*)Gx, Gy) = 2(Gx, Gy) + 2a(x,y),$$

we have that $B + B^* = 2 + 2aG^{-2}$. □

Corollary 2.8.13 *There exists a self-adjoint operator C in $\mathcal{H}$ such that C is m-accretive, $0 \in \rho(C)$, $\mathcal{D}(C) = \mathcal{V}$ and*

$$(Cx, Cy) = [x,y] \quad \textit{for all} \quad x, y \in \mathcal{V}.$$

PROOF Apply Theorem 2.8.12 when $\mathfrak{F}(x,y) = [x,y]$. □

2.9 Example: Harmonic Oscillator and Hermite Functions

To a particle subject to quadratic potential in one dimension (harmonic oscillator), one can assign the following quantum mechanical Hamiltonian:

$$(H_{osc}\psi)(x) = -\psi''(x) + x^2\psi(x), \quad x \in \mathbb{R}.$$

Since $|\psi|^2$ is a probability density, we have $\psi \in L^2(\mathbb{R})$. It can be easily seen that H_{osc} is symmetric in $L^2(\mathbb{R})$ provided that functions in its domain decay sufficiently rapidly at $\pm\infty$ and are smooth enough. However, we need to specify its domain so that H_{osc} is a self-adjoint operator. It is not obvious how to do this. Here, this will be accomplished by using techniques described in the section on Sectorial Forms. It will be also shown that H_{osc} has a complete set of eigenvectors - called Hermite functions.

Let $\mathcal{H} = L^2(\mathbb{R})$ and let $(\cdot,\cdot)$, $\|\cdot\|$ denote the usual inner product and norm on $L^2(\mathbb{R})$. Define also

$$\mathcal{AC} = \{\, f : \mathbb{R} \to \mathbb{C} \mid f \in AC[-a,a] \text{ for all } a > 0 \,\},$$

$$\mathcal{V} = \{\, f \in \mathcal{AC} \mid \int_{-\infty}^{\infty} (|f'(x)|^2 + (1+x^2)|f(x)|^2)dx < \infty \,\},$$

$$[f,g] = \int_{-\infty}^{\infty} (f'(x)\overline{g'(x)} + (1+x^2)f(x)\overline{g(x)})dx, \quad |f| = [f,f]^{1/2} \quad \text{for all} \quad f, g \in \mathcal{V}.$$

Example 1.3.4 shows that $\mathcal{V}$ is dense in $\mathcal{H}$.

Lemma 2.9.1 *If $f \in \mathcal{V}$, then $|f(x)| \le (1+x^2)^{-1/4}|f|$ for all $x \in \mathbb{R}$.*

PROOF If $0 \leq x \leq a$, then

$$|f(a)|^2 - |f(x)|^2 = 2\mathrm{Re} \int_x^a f'\overline{f}$$

and, since $f'\overline{f} \in L^1(\mathbb{R})$, the limit of the right-hand side exists as $a \to \infty$, hence, the limit of $|f(a)|^2$ as $a \to \infty$ has to exist and, since $f \in L^2(\mathbb{R})$, the limit has to be 0. Thus

$$\begin{aligned} |f(x)|^2 &= -2\mathrm{Re} \int_x^\infty f'(s)\overline{f(s)}ds \\ &\leq \int_x^\infty (1+s^2)^{-1/2}(|f'(s)|^2 + (1+s^2)|f(s)|^2)ds \\ &\leq (1+x^2)^{-1/2} \int_x^\infty (|f'(s)|^2 + (1+s^2)|f(s)|^2)ds. \end{aligned}$$

A similar argument can be applied when $x \leq 0$. □

Lemma 2.9.2 *$\mathcal{V}$ is a Hilbert space.*

PROOF Let $\{f_n\}$ be a Cauchy sequence in $\mathcal{V}$. Lemma 2.9.1 implies that for each $x \in \mathbb{R}$ there exists $\lim_{n\to\infty} f_n(x) = f(x)$. Since $\{f_n'\}$ is Cauchy in $L^2(\mathbb{R})$, there exists $g \in L^2(\mathbb{R})$ such that $\|f_n' - g\| \to 0$ hence $f_n(x) - f_n(0) = \int_0^x f_n'$ implies $f(x) - f(0) = \int_0^x g$. Thus $f \in \mathcal{AC}$, $f' = g$ a.e.. Since the functions $x \to (1+x^2)^{1/2} f_n(x)$ form a Cauchy sequence in $L^2(\mathbb{R})$, they converge in $L^2(\mathbb{R})$ to $(1+x^2)^{1/2} f(x)$. Thus, $f \in \mathcal{V}$ and $|f_n - f| \to 0$ as $n \to \infty$. □

Define $\mathfrak{F} : \mathcal{V} \times \mathcal{V} \to \mathbb{C}$ by

$$\mathfrak{F}(f,g) = [f,g].$$

The assumptions in the Sectorial Forms section are clearly satisfied ($M_1 = 1$, $M_2 = 1$, $M_3 = 1$, $a = 0$). Let A be the self-adjoint operator (Corollary 2.8.5) associated with $\mathfrak{F}$. The following two Lemmas characterize A explicitly and suggest that the definition $H_{osc} = A - 1$ is the right one.

Lemma 2.9.3 *If $f \in \mathcal{V}$, $f' \in \mathcal{AC}$ and $g \in \mathcal{H}$, where $g(x) = -f''(x) + (1+x^2)f(x)$, then $f \in \mathcal{D}(A)$ and $Af = g$.*

PROOF Choose $v \in \mathcal{V}$, $a < b$ and note that

$$\int_a^b (f'(x)\overline{v'(x)} + (1+x^2)f(x)\overline{v(x)})dx = f'(b)\overline{v(b)} - f'(a)\overline{v(a)} + \int_a^b g(x)\overline{v(x)}dx.$$

Therefore the limits of $f'(x)\overline{v(x)}$ have to exist as $x \to \pm\infty$ and, since $f'\overline{v} \in L^1(\mathbb{R})$, the limits have to be 0. Therefore $\mathfrak{F}(f,v) = (g,v)$ for all $v \in \mathcal{V}$. □

Lemma 2.9.4 *$f \in \mathcal{D}(A)$ if and only if*

$$f \in \mathcal{V},\ f' \in \mathcal{AC} \text{ and } g \in \mathcal{H}, \text{ where } g(x) = -f''(x) + (1+x^2)f(x).$$

PROOF Suppose $f \in \mathcal{D}(A)$. $\mathfrak{F}(f,v) = (Af, v)$ for all $v \in \mathcal{V}$ implies that

$$\int_{-\infty}^{\infty} f'(x)\psi'(x)dx = \int_{-\infty}^{\infty} ((Af)(x) - (1+x^2)f(x))\psi(x)dx$$

for all $\psi \in C_0^1(\mathbb{R})$. Thus Lemma 2.4.5 implies that $f' \in \mathcal{AC}$ and $f''(x) = -(Af)(x) + (1+x^2)f(x)$. □

Lemma 2.9.5 *The identity map $I : \mathcal{V} \to \mathcal{H}$ is compact.*

PROOF Choose $f_n \in \mathcal{V}$ such that $\sup_n |f_n| = M < \infty$. Since

$$|f_n(x) - f_n(y)| = |\textstyle\int_y^x f_n'| \le |x-y|^{1/2}M,$$

$\{f_n\}$ is an equicontinuous sequence. Lemma 2.9.1 implies its boundedness; thus the Arzela-Ascoli Theorem 1.1.5 implies that there exists a subsequence $\{f_{n_k}\}$ converging uniformly on compact subsets of $\mathbb{R}$ to some $g \in C(\mathbb{R})$. Using the fact that f_n decay uniformly at infinity enables us to conclude that the subsequence $\{f_{n_k}\}$ converges uniformly on $\mathbb{R}$ to g. Since

$$\int_{-a}^{a} (1+x^2)|f_n(x)|^2 dx \le M^2 \quad \text{for all} \quad a > 0,\ n \ge 1,$$

we have that

$$\int_{-a}^{a} (1+x^2)|g(x)|^2 dx \le M^2 \quad \text{for all} \quad a > 0.$$

Therefore

$$\int_{-\infty}^{\infty} |g - f_{n_k}|^2 \le$$

$$\|g - f_{n_k}\|_\infty \left(\int_{-\infty}^{\infty} \frac{dx}{1+x^2}\right)^{1/2} \left(\int_{-\infty}^{\infty} (1+x^2)|g(x) - f_{n_k}(x)|^2 dx\right)^{1/2}$$

implies $\|g - f_{n_k}\|^2 \le \|g - f_{n_k}\|_\infty \pi^{1/2} 2M \to 0$. □

Theorem 2.8.2 implies that A^{-1} is compact, hence, Theorems 2.6.8, 1.7.16 apply. In particular, $\sigma(A) = \sigma_p(A) \subset [1, \infty)$ and the normalized eigenvectors of A form a complete orthonormal set in $\mathcal{H}$. We proceed now to find eigenvalues and eigenvectors explicitly.

Define polynomials h_n by

$$h_0(x) = 1, \quad h_{n+1}(x) = 2xh_n(x) - h_n'(x) \quad \text{for } x \in \mathbb{R}, n = 0, 1, 2, \ldots.$$

Define **Hermite functions** ϕ_n by

$$\phi_n(x) = h_n(x)e^{-x^2/2} \quad \text{for } x \in \mathbb{R}, n \geq 0.$$

Note that $\phi_n \in \mathcal{D}(A)$, $\phi_{n+1}(x) = x\phi_n(x) - \phi_n'(x)$ for $n \geq 0$, $x \in \mathbb{R}$. By induction,

$$A\phi_n = (2n + 2)\phi_n \quad \text{for } n \geq 0.$$

In the next two Lemmas it will be shown that $\{\phi_n\}$ are the only eigenvectors of A.

Lemma 2.9.6 *Suppose $u \in \mathcal{D}(A)$ and $Au = \lambda u$ for some $\lambda \in \mathbb{R}$. Let $w(x) = u'(x) + xu(x)$. Then $w \in \mathcal{D}(A)$, $Aw = (\lambda - 2)w$ and $w'(x) - xw(x) = (2 - \lambda)u(x)$.*

PROOF It is easy to see that $w \in \mathcal{H} \cap \mathcal{AC}$, $w'(x) = xw(x) + (2 - \lambda)u(x)$, $w''(x) = (x^2 + 3 - \lambda)w(x)$. Thus, all that has to be shown is that $w \in \mathcal{V}$.

Let $f(x) = |w(x)|^2$. Note that $f'' = 2\text{Re}\,(w''\overline{w} + |w'|^2)$. Thus

$$f'(b) - f'(a) = 2\int_a^b (|w'(x)|^2 + (x^2 + 3 - \lambda)|w(x)|^2)dx.$$

Hence, as $b \to \infty$ the right-hand-side increases to some number in $(-\infty, \infty]$ and, since $f \in L^1(\mathbb{R})$, the limit cannot be $+\infty$. The limit of the right-hand-side must similarly be finite as $a \to -\infty$, and therefore $w \in \mathcal{V}$. □

Lemma 2.9.7 *If $u \in \mathcal{D}(A)$, $u \neq 0$, $\lambda \in \mathbb{R}$ and $Au = \lambda u$, then $u = c\phi_m$ and $\lambda = 2m + 2$ for some $c \in \mathbb{C}$ and some integer $m \geq 0$.*

PROOF Let $u_0 = u$, $u_{n+1} = u_n' + xu_n$ for $n \geq 0$. Note $Au_n = (\lambda - 2n)u_n$ for $n \geq 0$. Since $\sigma(A) \subset [1, \infty)$, there exists $m \geq 0$ such that $u_m \neq 0$ and $0 = u_{m+1} = u_{m+2} = \cdots$ Thus $0 = u_m' + xu_m$ implies $u_m(x) = c_m e^{-x^2/2}$ for some constant $c_m \neq 0$. Since $u_{n+1}' - xu_{n+1} = (2 + 2n - \lambda)u_n$ for $n \geq 0$, we have that $\lambda = 2m + 2$ and the induction assumption $u_{n+1} = c_{n+1}\phi_{m-n-1}$ implies $u_n = \phi_{m-n}c_{n+1}/(2m - 2n)$. Therefore $u_n = \phi_{m-n}c_m/(2^{m-n}(m - n)!)$ for $0 \leq n \leq m$. □

The following Theorem summarizes the above results. A different proof of completeness of normalized Hermite functions can be found in Chapter 3, Exercise 1.

Theorem 2.9.8 *A is a self-adjoint operator with compact resolvent, $\sigma(A) = \sigma_p(A) = \{2, 4, 6, \cdots\}$ and $\{\phi_n/\|\phi_n\|\}_{n=0}^{\infty}$ is a complete orthonormal set in $L^2(\mathbb{R})$.*

2.10 Example: Completeness of Bessel Functions

If $u = u(r)$, $r = (x_1^2 + \cdots + x_n^2)^{1/2}$, then

$$\Delta u \equiv \frac{\partial^2 u}{\partial x_1^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2} = \frac{d^2 u}{dr^2} + \frac{n-1}{r}\frac{du}{dr}. \tag{2.51}$$

A natural Hilbert space in which to study the Laplacian Δ is $L^2(\Omega)$, where Ω is a unit ball in $\mathbb{R}^n$. However, if one considers only radially symmetric functions u, then the square of the $L^2(\Omega)$ norm of u is proportional to

$$\int_0^1 r^{n-1}|u(r)|^2 dr.$$

Multiplying (2.51) by $r^{n-1}\overline{v(r)}$ and integrating by parts gives

$$\int_0^1 r^{n-1}(\Delta u)(r)\,\overline{v(r)}dr = -\int_0^1 r^{n-1}u'(r)\overline{v'(r)}dr,$$

provided that $v(1) = 0$ and that u and v are smooth enough and $n > 1$. This motivates the following definitions.

Fix any $\nu \geq 0$. Let $\mathcal{H}$ denote the set of all equivalence classes of measurable functions f such that

$$\int_0^1 x^{1+2\nu}|f(x)|^2 dx < \infty.$$

$\mathcal{H}$ is a Hilbert space with an inner product

$$(f, g) = \int_0^1 x^{1+2\nu} f(x)\overline{g(x)}dx.$$

Let $\mathcal{AC}$ be the set of all $u \in C(0,1]$ such that $u \in AC[\varepsilon, 1]$ for all $\varepsilon \in (0,1)$. Define $\mathcal{V}$ to be the set of all $u \in \mathcal{AC}$ such that $u(1) = 0$ and

$$\int_0^1 x^{1+2\nu}|u'(x)|^2 dx < \infty.$$

Define

$$[u, v] = \int_0^1 x^{1+2\nu}u'(x)\overline{v'(x)}dx, \quad |u| = [u, u]^{1/2} \quad \text{for} \quad u, v \in \mathcal{V}.$$

Lemma 2.10.1 *$\mathcal{V}$ is a dense subspace of $\mathcal{H}$. If $u \in \mathcal{V}$, then $\|u\| \leq |u|$ and*

$$x^{1+2\nu}|u(x)|^2 < |u|^2 \quad \textit{for} \quad x \in (0,1]. \tag{2.52}$$

PROOF For $u \in \mathcal{V}$ and $x \in (0,1]$, we have

$$\begin{aligned} u(x) &= -\int_x^1 s^{1/2+\nu} u'(s) s^{-1/2-\nu} ds \\ |u(x)|^2 &\leq \int_x^1 s^{1+2\nu} |u'(s)|^2 ds \int_x^1 s^{-1-2\nu} ds. \end{aligned}$$

Using Example 1.3.4 it is easy to see that $\mathcal{V}$ is a dense subspace of $\mathcal{H}$. □

Lemma 2.10.2 *$\mathcal{V}$ is a Hilbert space.*

PROOF If $\{u_n\}$ is a Cauchy sequence in $\mathcal{V}$, then $s^{1/2+\nu} u_n'(s)$ converge in $L^2(0,1)$ to some g hence

$$u_n(x) = -\int_x^1 s^{1/2+\nu} u_n'(s) s^{-1/2-\nu} ds \to -\int_x^1 g(s) s^{-1/2-\nu} ds \equiv u.$$

It is easy to see that $u \in \mathcal{V}$ and $\lim_{n\to\infty} |u_n - u| = 0$. □

Let $\mathfrak{F}(u,v) = [u,v]$ for $u, v \in \mathcal{V}$. The assumptions in the Sectorial Forms section are clearly satisfied with $M_1 = M_2 = M_3 = 1, a = 0$. Let A be the self-adjoint operator (Corollary 2.8.5) associated with $\mathfrak{F}$. Theorem 2.8.2 implies that $\rho(A) \supset (-\infty, 1)$. The following two Lemmas give an explicit representation of A.

Lemma 2.10.3 *If $u \in \mathcal{D}(A)$, then $u \in \mathcal{V}$, $u' \in \mathcal{AC}$, $\lim_{x\to 0+} |u'(x)| x^\nu = 0$ and*

$$(Au)(x) = -u''(x) - \frac{1+2\nu}{x} u'(x) \quad \textit{for almost all } x \in (0,1).$$

PROOF If $u \in \mathcal{D}(A)$, then for all $v \in \mathcal{V}$

$$\int_0^1 x^{1+2\nu} u'(x) \overline{v'(x)} dx = \int_0^1 x^{1+2\nu} (Au)(x) \overline{v(x)} dx.$$

Lemma 2.4.5 implies that for any $c \in (0,1)$ there exists k such that

$$x^{1+2\nu} u'(x) = k - \int_c^x s^{1+2\nu} (Au)(s) ds \quad \text{for} \quad x \in (0,1] \tag{2.53}$$

and, since the integrand is in $L^1(0,1)$, we may take $c = 0$. If $k \neq 0$, then

$$x^{1+2\nu} |u'(x)|^2 \geq x^{-1-2\nu} |k/2|^2 \quad \text{for } x \text{ close to } 0,$$

which contradicts the fact that $u \in \mathcal{V}$. (2.53) with $c = k = 0$ implies

$$|x^\nu u'(x)|^2 \leq (2+2\nu)^{-1} \int_0^x s^{1+2\nu} |(Au)(s)|^2 ds.$$

The derivative of (2.53) gives the expression for Au. □

Lemma 2.10.4 *If $u \in \mathcal{V}$, $u' \in \mathcal{AC}$, $\lim_{x\to 0+} |u'(x)|x^\nu = 0$ and $f \in \mathcal{H}$ where $f(x) = -u''(x) - (1+2\nu)u'(x)/x$, then $u \in \mathcal{D}(A)$.*

PROOF If $v \in \mathcal{V}$, then for $x \in (0,1)$ we have

$$-x^{1/2}\, x^\nu u'(x)\, x^{1/2+\nu}\overline{v(x)} = \int_x^1 s^{1+2\nu} u'(s)\overline{v'(s)}ds - \int_x^1 s^{1+2\nu} f(s)\overline{v(s)}ds.$$

(2.52) implies that the left hand side approaches 0 as $x \to 0$ hence $[u,v] = (f,v)$. □

Lemma 2.10.5 *A^{-1} is compact.*

PROOF Let $\{u_n\}$ be a bounded sequence in $\mathcal{V}$, $M = \sup_n |u_n|$. Note

$$|u_n(x) - u_n(y)| = \left| \int_y^x s^{1/2+\nu} u_n'(s)\; s^{-1/2-\nu}\, ds \right| \le M\rho^{-1/2-\nu}|x-y|^{1/2}$$

for $x, y \ge \rho > 0$; (2.52) implies pointwise boundedness hence the Arzela-Ascoli Theorem 1.1.5 implies that a subsequence $\{u_{n_i}\}$ converges pointwise to some $u \in C(0,1]$. (2.52) implies

$$x^{1+2\nu}|u_n(x) - u_m(x)|^2 < 4M^2 \quad \text{for} \quad x \in (0,1],\ n,m \ge 1$$

hence

$$x^{1+2\nu}|u_{n_i}(x) - u(x)|^2 \le 4M^2 \quad \text{for} \quad x \in (0,1],\ i \ge 1$$

and the DCT implies $\lim_{i\to\infty} \|u_{n_i} - u\| = 0$. Thus the identity map from $\mathcal{V}$ to $\mathcal{H}$ is compact. Theorem 2.8.2 implies that A^{-1} is compact. □

Since each eigenvector u of A satisfies a 2nd order ODE with $u(1) = 0$, we see (Example 1.6.9) that u is uniquely determined by the value of $u'(1)$. Hence $\dim \mathcal{N}(A-\lambda) = 1$ for each eigenvalue λ. The above observations, together with Theorems 2.6.8 and 1.7.16, imply:

Theorem 2.10.6 *A is a self-adjoint operator in $\mathcal{H}$ and*

$$\sigma(A) = \sigma_p(A) = \{\lambda_1, \lambda_2, \ldots\}$$

for some positive numbers $\lambda_1 < \lambda_2 < \cdots$ with $\lim_{n\to\infty} \lambda_n = \infty$. $\dim \mathcal{N}(A-\lambda_n) = 1$ for $n \ge 1$ and, if $\phi_n \in \mathcal{N}(A-\lambda_n)$ with $\|\phi_n\| = 1$, then $\{\phi_1, \phi_2, \ldots\}$ is a complete orthonormal set in $\mathcal{H}$.

We shall now examine eigenvalues and eigenvectors of A in more detail. The **Bessel function** J_ν is defined by

$$J_\nu(z) = (z/2)^\nu \sum_{n=0}^{\infty} \frac{(-z^2/4)^n}{n!\Gamma(n+1+\nu)} \tag{2.54}$$

for $|\arg(z)| < \pi$. For $\lambda > 0$, $x > 0$ define

$$v_\lambda(x) = (x\lambda^{1/2}/2)^{-\nu} J_\nu(x\lambda^{1/2}) = \sum_{n=0}^{\infty} \frac{(-x^2\lambda/4)^n}{n!\Gamma(n+1+\nu)};$$

a simple verification shows that, for $x > 0$,

$$-v_\lambda''(x) - (1+2\nu)v_\lambda'(x)/x = \lambda v_\lambda(x).$$

Lemma 2.10.4 implies that v_λ is an eigenvector of A when $J_\nu(\lambda^{1/2}) = 0$. Now suppose that $\lambda \in \sigma_p(A)$, $u \in \mathcal{D}(A)$ $Au = \lambda u$ and let $w = u'v_\lambda - uv_\lambda'$. Note that $w'(x) = -(1+2\nu)w(x)/x$ hence for some constant c

$$x^{1+2\nu}(u'(x)v_\lambda(x) - u(x)v_\lambda'(x)) = \mathrm{c} \quad \text{for} \quad x \in (0,1].$$

If in this equation we use the bounds on $u(x)$ and $u'(x)$, as given by (2.52) and Lemma 2.10.3, as well as the bounds on v_λ, v_λ' that follow from the Taylor series representation of v_λ, we conclude that $c = 0$. Hence $u'v_\lambda = uv_\lambda'$ which implies that $u = c_1 v_\lambda$ on $(0,1]$. This proves

Lemma 2.10.7 *$\lambda \in \sigma_p(A)$, $u \in \mathcal{D}(A)$ and $Au = \lambda u$ if and only if $J_\nu(\lambda^{1/2}) = 0$, $\lambda > 0$ and $u(x) = \mathrm{c}\, x^{-\nu} J_\nu(x\lambda^{1/2})$ for some constant c and all $x \in (0,1]$.*

Let $j_{\nu,n}$ be the nth positive zero of J_ν. Since $j_{\nu,n}^2$ are eigenvalues of A, we have by Theorem 2.10.6 that $\lim_{n\to\infty} j_{\nu,n} = \infty$. This Theorem and the above Lemma imply

Theorem 2.10.8 *If $\int_0^1 x|f(x)|^2 dx < \infty$, then*

$$\lim_{n\to\infty} \int_0^1 x \left| f(x) - \sum_{i=1}^{n} c_i J_\nu(j_{\nu,i}x) \right|^2 dx = 0$$

where

$$c_i = \frac{\int_0^1 x f(x) J_\nu(j_{\nu,i}x) dx}{\int_0^1 x J_\nu(j_{\nu,i}x)^2 dx}.$$

Evaluation of $\int_0^1 xJ_\nu(j_{\nu,i}x)^2dx$ can be simplified as follows. One verifies

$$(xJ'_\nu(x))^2 = (\nu^2 - x^2)J_\nu(x)^2 + 2\int_0^x sJ_\nu(s)^2ds$$

by differentiation, using the fact that

$$x^2J''_\nu(x) + xJ'_\nu(x) + (x^2 - \nu^2)J_\nu(x) = 0.$$

Therefore

$$\int_0^1 xJ_\nu(j_{\nu,i}x)^2dx = \frac{1}{2}J'_\nu(j_{\nu,i})^2.$$

Using (2.54) it is easy to see that $xJ'_\nu(x) = \nu J_\nu(x) - xJ_{\nu+1}(x)$ which implies

$$\int_0^1 xJ_\nu(j_{\nu,i}x)^2dx = \frac{1}{2}J_{\nu+1}(j_{\nu,i})^2.$$

2.11 Example: Finite Element Method

Theorem 2.8.11 gives the mathematical foundation of the **finite element** method for elliptic type problems. In order to demonstrate its usage we shall consider the problem of finding finite element approximations of u which is required to satisfy

$$-u''(x) + q(x)u(x) = f(x) \quad \text{for} \quad 0 \le x \le 1 \tag{2.55}$$

$$u(0) = 0,\ u'(1) + \gamma u(1) = 0 \tag{2.56}$$

with given $f \in L^2(0,1)$, $q \in L^1(0,1)$, $\gamma \in \mathbb{C}$. We assume for simplicity that $\operatorname{Re}\gamma \ge 0$ and $\operatorname{Re} q \ge 0$.

To obtain the variational formulation of the problem (i.e., to find the appropriate sectorial form) multiply (2.55) by $\overline{v}$, assuming $v(0) = 0$ just as $u(0) = 0$, and integrate by parts to obtain

$$-u'(1)\overline{v(1)} + \int_0^1 (u'(x)\overline{v'(x)} + q(x)u(x)\overline{v(x)})dx = \int_0^1 f(x)\overline{v(x)}dx. \tag{2.57}$$

This suggests that we define $\mathcal{H} = L^2(0,1)$ with $(\cdot,\cdot)$, $\|\cdot\|$ representing the usual inner product and norm on $L^2(0,1)$ and

$$\mathcal{V} = \{\, v \in AC[0,1] \mid v(0) = 0, v' \in L^2(0,1) \,\}$$

$$[v,w] = \int_0^1 v'(x)\overline{w'(x)}dx, \quad |v| = [v,v]^{1/2} \quad \text{for all} \quad v,w \in \mathcal{V}.$$

$\mathcal{V}$ is dense in $\mathcal{H}$, see Example 1.3.4, and it can be easily shown that $\mathcal{V}$ is a Hilbert space with an inner product $[\cdot,\cdot]$. The left hand side of (2.57) does not yet represent

the appropriate sectorial form because $u'(1)$ cannot be bounded by $|u|$; however, $u'(1)$ can be replaced by $-\gamma u(1)$ because of (2.56), and since $\|u\|_\infty \le |u|$ for every $u \in \mathcal{V}$, we now have the boundedness of the form. Thus the appropriate $\mathfrak{F} : \mathcal{V} \times \mathcal{V} \to \mathbb{C}$ is

$$\mathfrak{F}(w,v) = \gamma w(1)\overline{v(1)} + \int_0^1 (w'(x)\overline{v'(x)} + q(x)w(x)\overline{v(x)})dx. \tag{2.58}$$

Using these definitions, the assumptions of the section on Sectorial Forms are satisfied with $M_1 = 1$, $M_2 = 1 + |\gamma| + \|q\|_1$, $a = 0$, $M_3 = 1$. The linear operator A associated with $\mathfrak{F}$ has a bounded inverse (Theorem 2.8.2) and by using Lemma 2.4.5, one can easily obtain the following explicit representation of A:

$$Av = -v'' + qv \quad \text{for} \quad v \in \mathcal{D}(A)$$

$$\mathcal{D}(A) = \{\, v \in \mathcal{V} \mid v' \in AC[0,1],\ -v'' + qv \in L^2(0,1),\ v'(1) + \gamma v(1) = 0 \,\},$$

which confirms correctness of the variational formulation.

For $n \ge 1$ define $\mathcal{V}_n \subset \mathcal{V}$ as follows: $v \in \mathcal{V}_n$ if and only if

$$v \in C[0,1],\ v(0) = 0,\ v \text{ is linear on } [\tfrac{i}{n}, \tfrac{i+1}{n}] \text{ for each } i = 0, 1, \ldots, n-1.$$

Lemma 2.11.1 *If $v, v' \in AC[0,1]$, $v'' \in L^2(0,1)$, $n \ge 1$ and*

$$g(x) = v(\tfrac{i}{n}) + (nx - i)(v(\tfrac{i+1}{n}) - v(\tfrac{i}{n})) \quad \textit{for} \quad \tfrac{i}{n} \le x \le \tfrac{i+1}{n},\ 0 \le i \le n-1,$$

then $g \in AC[0,1]$ and $\|g' - v'\|_2 \le \frac{1}{n}\|v''\|_2$.

PROOF If $\frac{i}{n} \le x \le \frac{i+1}{n}$, $0 \le i \le n-1$, then

$$g'(x) - v'(x) = n \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left(\int_x^s v''(t)dt \right) ds$$

$$|g'(x) - v'(x)| \le n \int_{\frac{i}{n}}^{\frac{i+1}{n}} |s - x|^{1/2} \left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} |v''|^2 \right)^{1/2} ds \le n^{-1/2} \left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} |v''|^2 \right)^{1/2}$$

$$\int_{\frac{i}{n}}^{\frac{i+1}{n}} |g' - v'|^2 \le n^{-2} \int_{\frac{i}{n}}^{\frac{i+1}{n}} |v''|^2.$$

□

The above Lemma implies

$$\lim_{n \to \infty} \inf_{w \in \mathcal{V}_n} |w - v| = 0$$

whenever $v \in \mathcal{D} \equiv \{v \in \mathcal{V} | v' \in AC[0,1], v'' \in L^2(0,1)\}$. If $q = 0$, then $\mathcal{D}(A) \subset \mathcal{D}$ and, since Theorem 2.8.2 implies that $\mathcal{D}(A)$ is dense in $\mathcal{V}$, we have that $\mathcal{D}$ is dense in $\mathcal{V}$ and therefore Theorem 2.8.11 applies.

Fix $f \in \mathcal{H}$, $n \geq 1$. Let $w_n \in \mathcal{V}_n$ denote the nth finite element **Galerkin approximation** of $u = A^{-1}f$ defined by

$$\mathfrak{F}(w_n, v) = (f, v) \quad \text{for all} \quad v \in \mathcal{V}_n. \tag{2.59}$$

Since $w_n \in \mathcal{V}_n$, there exist scalars $c_1, \ldots, c_n$ such that

$$w_n(x) = c_i + (nx - i)(c_{i+1} - c_i) \quad \text{for } \tfrac{i}{n} \leq x \leq \tfrac{i+1}{n},\ 0 \leq i \leq n-1$$

where $c_0 = 0$. Note $w_n(i/n) = c_i$ for $0 \leq i \leq n$. The **hat function** ψ is defined by

$$\psi(x) = \begin{cases} 1 - |x| & \text{if } -1 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Using $v(x) = \psi(nx - k)$ in (2.59) with $k = 1, \ldots, n$ gives the following n equations,

$$\alpha_k c_k - \beta_{k-1} c_{k-1} - \beta_k c_{k+1} = \int_{-1}^{1} f(\tfrac{k+s}{n})\psi(s)ds \quad \text{for} \quad k = 1, \ldots, n-1$$

$$\alpha_n c_n - \beta_{n-1} c_{n-1} = \int_{-1}^{0} f(\tfrac{n+s}{n})\psi(s)ds,$$

where

$$\alpha_k = 2n^2 + \int_{-1}^{1} q(\tfrac{k+s}{n})\psi(s)^2 ds \quad \text{for} \quad k = 1, \ldots, n-1$$

$$\beta_k = n^2 - \int_{0}^{1} q(\tfrac{k+s}{n}) s\psi(s) ds \quad \text{for} \quad k = 0, \ldots, n-1$$

$$\alpha_n = n^2 + n\gamma + \int_{-1}^{0} q(\tfrac{n+s}{n})\psi(s)^2 ds,$$

which completely determine $c_1, \ldots, c_n$ and hence w_n. Theorem 2.8.11 implies that

$$\|w_n - u\|_\infty \leq |w_n - u| \to 0 \quad \text{as } n \to \infty.$$

This convergence result can be strengthened when $q \in L^2(0,1)$. If $q \in L^2(0,1)$, then

$$\begin{aligned} \|v''\|_2 &\leq \|Av\| + \|q\|_2 \|v\|_\infty \leq \|Av\| + \|q\|_2 |v| \\ &\leq \|Av\| + \|q\|_2 (\mathrm{Re}\, \mathfrak{F}(v,v))^{1/2} = \|Av\| + \|q\|_2 (\mathrm{Re}\,(Av, v))^{1/2} \\ &\leq (1 + \|q\|_2)\|Av\| + \|q\|_2 \|v\| \end{aligned}$$

for $v \in \mathcal{D}(A)$, hence Lemma 2.11.1 implies (2.47). Therefore there exists a constant C, independent of n and f, such that

$$\|w_n - u\|_\infty \leq |w_n - u| \leq \frac{C}{n}\|f\|.$$

2.12 Friedrichs Extension

Throughout this section it will be assumed that:

(**S1**) $\mathcal{H}$ is a *complex* Hilbert space

(**S2**) S is a densely defined linear operator in $\mathcal{H}$

(**S3**) there exist $r \in \mathbb{R}$ and $M \in (0, \infty)$ such that

$$|\mathrm{Im}\,(Sx, x)| \leq M\,\mathrm{Re}\,(Sx - rx, x) \quad \text{for all} \quad x \in \mathcal{D}(S).$$

The purpose of this section is to prove:

Theorem 2.12.1 *There exists a subspace $\mathcal{V}$ of $\mathcal{H}$, an inner product $[\cdot,\cdot]$ on $\mathcal{V}$ with the corresponding norm $|\cdot|$ and a sectorial sesquilinear form $\mathfrak{F}$ on $\mathcal{V}$ such that the assumptions* **H1**, **H2**, **H3** *in the Sectorial Forms section are satisfied and that also the following hold:*

(a) $\mathcal{D}(S)$ is dense subspace in $\mathcal{V}$ (in $|\cdot|$ norm)

(b) $[x, y] = (1/2)(\,(Sx, y) + (x, Sy)\,) + (1 - r)(x, y)$ for all $x, y \in \mathcal{D}(S)$

(c) $\mathfrak{F}(x, y) = (Sx, y)$ for all $x, y \in \mathcal{D}(S)$.

The linear operator A associated with $\mathfrak{F}$ is obviously an extension of S and is called the **Friedrichs extension** of S. Continuity of $\mathfrak{F}$ and the fact that $\mathcal{D}(S)$ is dense in $\mathcal{V}$ also imply that the numerical range of $\mathfrak{F}$, and hence of A, also lies in $\Gamma = \{\,\zeta \in \mathbb{C}\,|\,|\mathrm{Im}\,\zeta| \leq M(\mathrm{Re}\,\zeta - r)\,\}$. Theorems 1.6.16 and 2.8.2 imply that $\sigma(A) \subset \Gamma$ and $\|(A - \lambda)^{-1}\| \leq 1/\mathrm{dist}(\lambda, \Gamma)$ for $\lambda \in \mathbb{C}\backslash\Gamma$. Also, if S is symmetric, then A is self-adjoint by Corollary 2.8.5 since $\mathfrak{F}(x, y) = \overline{\mathfrak{F}(y, x)}$ for all $x, y \in \mathcal{D}(S)$ and hence for all $x, y \in \mathcal{V}$.

Observe also that if $\mathcal{R}(S - \lambda) = \mathcal{H}$ for some $\lambda \in \mathbb{C}\backslash\Gamma$, then $\sigma(S) \subset \Gamma$ by Theorem 1.6.16 and hence Lemma 1.6.14 implies that $S = A$ in this case.

A combination of Theorems 2.8.11 and 2.12.1, stated as Theorem 2.12.6 below, gives a proof of convergence of Galerkin approximations under assumptions that are elementary to verify in applications.

Now to the proof of Theorem 2.12.1. Define an inner product $[\cdot,\cdot]$ on $\mathcal{D}(S)$ by (b) of Theorem 2.12.1 and note that

$$|x|^2 = \mathrm{Re}\,(Sx - rx + x, x) \geq \|x\|^2 \quad \text{for all} \quad x \in \mathcal{D}(S). \tag{2.60}$$

Lemma 2.12.2 $|(Sx - rx + x, y)| \leq (1 + M)|x||y|$ *for all $x, y \in \mathcal{D}(S)$.*

PROOF Define $R(x) = (Sx - rx + x, x)$ and note that, for all $x, y \in \mathcal{D}(S)$,

$$\begin{aligned} & 4|\mathrm{Re}\,(Sx - rx + x, y)| \\ &= |\mathrm{Re}\,R(x+y) - \mathrm{Re}\,R(x-y) - \mathrm{Im}\,R(x+iy) + \mathrm{Im}\,R(x-iy)| \\ &\leq \mathrm{Re}\,R(x+y) + \mathrm{Re}\,R(x-y) + M(\mathrm{Re}\,R(x+iy) + \mathrm{Re}\,R(x-iy)) \\ &= 2(1+M)(|x|^2 + |y|^2). \end{aligned}$$

Choosing suitable $\alpha \in \mathbb{C}$ and replacing x by αx gives

$$2|(Sx - rx + x, y)| \leq (1+M)(|x|^2 + |y|^2).$$

Replacing x by tx and y by $(1/t)y$ gives

$$2|(Sx - rx + x, y)| \leq (1+M)(t^2|x|^2 + t^{-2}|y|^2),$$

and choosing $t \in (0, \infty)$ so that the right hand side is minimized gives the claimed bound. □

Define $\mathcal{V} \subset \mathcal{H}$ as follows. $x \in \mathcal{V}$ if and only if there exists a sequence $x_1, x_2, \ldots$ in $\mathcal{D}(S)$ such that

$$\lim_{n\to\infty} \|x_n - x\| = \lim_{n,m\to\infty} |x_n - x_m| = 0;$$

in such cases we will write $x_n \overset{in\mathcal{D}(S)}{\longrightarrow} x$.

Lemma 2.12.3 *If* $x_n \overset{in\mathcal{D}(S)}{\longrightarrow} 0$, *then* $\lim_{n\to\infty} |x_n| = 0$.

PROOF

$$\begin{aligned} |x_n|^2 &= \mathrm{Re}\,((S - r + 1)(x_n - x_m), x_n) + \mathrm{Re}\,((S - r + 1)x_m, x_n) \\ &\leq (1+M)|x_n - x_m||x_n| + \|(S - r + 1)x_m\|\|x_n\|. \end{aligned}$$

□

For $x, y \in \mathcal{V}$ define $[x, y]$ as follows. Choose $x_n \overset{in\mathcal{D}(S)}{\longrightarrow} x$, $y_n \overset{in\mathcal{D}(S)}{\longrightarrow} y$. It is easy to see that $\lim_{n\to\infty}[x_n, y_n]$ exists and that Lemma 2.12.3 implies that it does not depend on the particular choice of x_n, y_n. Define $[x, y] = \lim_{n\to\infty}[x_n, y_n]$. $[\cdot, \cdot]$ is an inner product on $\mathcal{V}$ and (2.60) implies

$$|x| \equiv [x, x]^{1/2} \geq \|x\| \quad \text{for all} \quad x \in \mathcal{V}.$$

Lemma 2.12.4 $\mathcal{D}(S)$ *is dense in* $\mathcal{V}$, *i.e., if* $x_n \overset{in\mathcal{D}(S)}{\longrightarrow} x$, *then* $\lim_{n\to\infty} |x_n - x| = 0$.

PROOF Observe first that $\lim_{n\to\infty}|x_n| = |x|$ and $\lim_{n\to\infty}|x_n \pm x_m| = |x \pm x_m|$ for $m \geq 1$. Pick $\varepsilon > 0$ and let N be such that $|x_n - x_m|^2 < \varepsilon$ and $||x_n|^2 - |x|^2| < \varepsilon$ for all $n, m \geq N$. Note that for $n, m \geq N$ we have

$$|x_n - x_m|^2 + |x_n + x_m|^2 = 2|x_n|^2 + 2|x_m|^2$$

$$||x_n + x_m|^2 - 4|x|^2| < 5\varepsilon$$

$$||x_n + x|^2 - 4|x|^2| \leq 5\varepsilon$$

$$|x_n - x|^2 + |x_n + x|^2 = 2|x_n|^2 + 2|x|^2$$

$$|x_n - x|^2 < 7\varepsilon.$$

□

Lemma 2.12.5 *$\mathcal{V}$ is a Hilbert space with inner product $[\cdot, \cdot]$.*

PROOF Suppose $\{x_n\}$ is a Cauchy sequence in $\mathcal{V}$. Pick $y_n \in \mathcal{D}(S)$ such that $|x_n - y_n| < 1/n$ for $n \geq 1$ (Lemma 2.12.4). Hence

$$\|y_n - y_m\| \leq |y_n - y_m| < \frac{1}{n} + \frac{1}{m} + |x_n - x_m|$$

and therefore there exists $y \in \mathcal{H}$ such that $y_n \overset{in\mathcal{D}(S)}{\longrightarrow} y$, hence, $y \in \mathcal{V}$ and

$$|x_n - y| < \frac{1}{n} + |y_n - y| \to 0 \quad \text{as } n \to \infty.$$

□

Lemma 2.12.2 implies that $|(Sx, y)| \leq M_2|x||y|$ for $x, y \in \mathcal{D}(S)$, where $M_2 = M + 1 + |r - 1|$. Define $\mathfrak{F}(x, y)$ for $x, y \in \mathcal{V}$ by

$$\mathfrak{F}(x, y) = \lim_{n\to\infty}(Sx_n, y_n)$$

where $x_n \overset{in\mathcal{D}(S)}{\longrightarrow} x$, $y_n \overset{in\mathcal{D}(S)}{\longrightarrow} y$. The assumptions $\mathbf{H1}, \mathbf{H2}, \mathbf{H3}$ in the section on Sectorial Forms are thus satisfied with $M_1 = M_3 = 1$, $a = r - 1$.

This completes the proof of Theorem 2.12.1.

The following Theorem, with all assumptions explicitly stated, follows directly from Theorems 2.8.11 and 2.12.1.

Theorem 2.12.6 *Suppose that*

(a) $\mathcal{H}$ is a complex Hilbert space

(b) $\phi_n \in \mathcal{H}$ *for* $n \geq 1$, $\mathcal{V}_n = \text{span}\{\phi_1, \ldots, \phi_n\}$ *and* $\cup_{n=1}^{\infty} \mathcal{V}_n$ *is dense in* $\mathcal{H}$

(c) S *is a linear operator in* $\mathcal{H}$ *with domain* $\mathcal{D}(S) = \cup_{n=1}^{\infty} \mathcal{V}_n$ *and there exist* $r \in \mathbb{R}$ *and* $M \in (0, \infty)$ *such that*

$$|\text{Im}\,(Sx, x)| \leq M\text{Re}\,(Sx - rx, x) \quad \textit{for all} \quad x \in \mathcal{D}(S)$$

(d) $f \in \mathcal{H}$ *and* $\lambda \in \mathbb{C}$ *with* $\text{Re}\,\lambda < r$.

Then for each $n \geq 1$ *there exists a unique* $x_n \in \mathcal{V}_n$ *such that*

$$(Sx_n - \lambda x_n, \phi_k) = (f, \phi_k) \quad \textit{for} \quad k = 1, \ldots, n.$$

Moreover,

$$\lim_{n\to\infty} \|x_n - x\| \leq \lim_{n\to\infty} |x_n - x| = 0$$

where $x = (A - \lambda)^{-1} f$ *and* A *is the Friedrichs extension of* S.

EXAMPLE 2.12.7 Let $\Omega = (0, \pi) \times (0, \pi)$, $\mathcal{H} = L^2(\Omega)$ and

$$Su = -au_{xx} - bu_{xy} - cu_{yy} + du_x + eu_y + fu,$$

where a, b, c, d, e, f are complex valued functions on Ω such that

$$a, b, c \in C_B^1(\Omega), \qquad d, e, f \in L^\infty(\Omega).$$

To make S an elliptic operator assume there exists $\delta > 0$ such that

$$\text{Re}\,(a(x,y)\zeta^2 + b(x,y)\zeta\eta + c(x,y)\eta^2) \geq \delta(\zeta^2 + \eta^2) \quad \text{for} \quad (x,y) \in \Omega,\ \zeta, \eta \in \mathbb{R}. \quad (2.61)$$

Let $\phi_{km}(x,y) = \sin kx \sin my$ for $k, m \geq 1$ and let the domain of S consist of all finite linear combinations of ϕ_{km}. We know that $\mathcal{D}(S)$ is dense in $\mathcal{H}$ - see Example 2.1.8 and Theorem 2.1.10.

To show **S3**, observe that integration by parts implies, for $u \in \mathcal{D}(S)$,

$$\begin{aligned}
\int bu_{xy}\overline{u} &= -\int (b\overline{u})_y u_x = -\int b_y \overline{u} u_x - \int b\overline{u}_y u_x \\
&= -\int (b\overline{u})_x u_y = -\int b_x \overline{u} u_y - \int b\overline{u}_x u_y \\
&= -\tfrac{1}{2}\int b_x \overline{u} u_y - \tfrac{1}{2}\int b_y \overline{u} u_x - \int b\text{Re}\,(\overline{u}_x u_y)
\end{aligned}$$

where $\int$ denotes the double integral over Ω, hence,

$$(Su, u) = \int (a\overline{u})_x u_x - bu_{xy}\overline{u} + (c\overline{u})_y u_y + (du_x + eu_y + fu)\overline{u}$$

$$= \int a|u_x|^2 + c|u_y|^2 + b\text{Re}\,(\overline{u}_x u_y) + (d + a_x + b_y/2)u_x\overline{u} + (e + c_y + b_x/2)u_y\overline{u} + f|u|^2.$$

(2.61) implies

$$\mathrm{Re}\,(a|u_x|^2 + c|u_y|^2 + b\mathrm{Re}\,(\overline{u}_x u_y)) \geq \delta(|u_x|^2 + |u_y|^2)$$

hence

$$\mathrm{Re}\,(Su,u) \geq \delta\|\nabla u\|^2 - E\|\nabla u\|\|u\| + f_m\|u\|^2 \tag{2.62}$$

where $\|\nabla u\|^2 = \|u_x\|^2 + \|u_y\|^2$ and $f_m \in \mathbb{R}$, $E \in [0,\infty)$ are such that a.e.

$$f_m \leq \mathrm{Re}\, f, \quad E^2 \geq |d + a_x + b_y/2|^2 + |e + c_y + b_x/2|^2.$$

Obviously

$$|\mathrm{Im}\,(Su,u)| \leq C_1\|\nabla u\|^2 + E\|\nabla u\|\|u\| + C_2\|u\|^2$$

for some constants C_1, C_2, hence, (2.62) implies that if

$$r < f_m - \frac{E^2}{4\delta},$$

then there exists $M \in (0,\infty)$ such that

$$|\mathrm{Im}\,(Su,u)| \leq M\mathrm{Re}\,(Su - ru, u) \quad \text{for all} \quad u \in \mathcal{D}(S).$$

Better estimates of r are possible. For example, just using the fact that $\|\nabla u\| \geq \sqrt{2}\|u\|$ for $u \in \mathcal{D}(S)$ implies that we can take $r < f_m + 2\delta - \sqrt{2}E$ when $E < 2\sqrt{2}\delta$.

Choose $\lambda < r$ and $g \in L^2(\Omega)$. The Galerkin approximation

$$u_n = \sum_{k=1}^{n}\sum_{m=1}^{n} c_{km}\phi_{km}$$

is determined by

$$(Su_n - \lambda u_n, \phi_{km}) = (g, \phi_{km}) \quad \text{for} \quad 1 \leq k,m \leq n$$

and, by Theorem 2.12.6, they converge to $(A-\lambda)^{-1}g$ where A is the Friedrichs extension of S.

EXAMPLE 2.12.8 Various singular problems can also be handled by Theorem 2.12.6. Let us consider, for example, the problem of finding u such that

$$-u''(x) - \frac{4}{x}u'(x) + \left(\frac{4}{x^2} - 1\right)u(x) = f(x) \quad \text{for} \quad 0 < x < 1$$

$$u(0) = u(1) = 0.$$

Let the Hilbert space $\mathcal{H}$ consist of all measurable functions $g : (0,1) \to \mathbb{C}$ which satisfy

$$\int_0^1 x^4|g(x)|^2dx < \infty,$$

and let its inner product be given by

$$(h,g) = \int_0^1 x^4 h(x)\overline{g(x)}\,dx \quad \text{for} \quad h,g \in \mathcal{H}.$$

Choose the basis functions to be

$$\phi_k(x) = x^k(1-x) \quad \text{for} \quad k \geq 1$$

and let the $\mathcal{D}(S)$ be the set of all finite linear combinations of ϕ_k. Taking $h(x) = x^3(1-x)$ in Example 2.1.9 gives that $\mathcal{D}(S)$ is dense in $\mathcal{H}$. Define

$$(Su)(x) = -u''(x) - \frac{4}{x}u'(x) + \left(\frac{4}{x^2} - 1\right)u(x) \quad \text{for} \quad u \in \mathcal{D}(S).$$

It is easy to see that $(Su - 3u, u) \in [0, \infty)$ for all $u \in \mathcal{D}(S)$ and hence we can apply Theorem 2.12.6 with $\lambda = 0$ and any $f \in \mathcal{H}$. In particular, the coefficients $z = (z_1, \ldots, z_n)^\perp$ of the nth Galerkin approximation $u_n = z_1\phi_1 + \cdots + z_n\phi_n$ are given by $z = A_n^{-1}c_n$ where $c_n = ((f, \phi_1), \ldots, (f, \phi_n))^\perp$ and the entries of the $n \times n$ matrix A_n are

$$(A_n)_{ij} = (S\phi_j, \phi_i) = \frac{2ij + 4k}{k(k-1)(k-2)} - \frac{2}{k(k+1)(k+2)} \qquad (k = i + j + 5).$$

Convergence can be tested by choosing the solution to be $u(x) = \sin \pi x$. It is found that the error rapidly decreases with n; for example, $\|u - u_9\|_\infty$ is approximately 10^{-8}.

2.13 Exercises

1. Show that every orthonormal set is contained in a complete orthonormal set.

2. Suppose that $f \in C(\mathbb{R})$ and that it has period 2π. Show that the arithmetic means of the partial sums of the Fourier series of f converge uniformly to f. *Hint:* find Q_n such that

$$\frac{P_0(x) + \cdots + P_{n-1}(x)}{n} = \int_{-\pi}^{\pi} f(x-s)Q_n(s)ds$$

where P_n are the partial sums, as in Example 2.1.7, and consider the integrand on intervals $I = (-\delta, \delta)$, $(-\pi, \pi)\backslash I$ separately.

3. Prove completeness of the normalized Legendre polynomials in $L^2(-1, 1)$.

4. Prove that if M is a closed subspace of a Hilbert space H and $x_0 \in H$, then

$$\operatorname{dist}(x_0, M) = \max\{|(x_0, y)| \mid y \in M^\perp, \|y\| = 1\}.$$

5. Provide a detailed proof of Theorem 2.2.6.

6. If S is a densely defined accretive linear operator such that $(Su, v) = -(u, Sv)$ for all $u, v \in \mathcal{D}(S)$, and if T is an m-accretive extension of S as in Theorem 2.3.6, then show that

$$(Tu, v) = -(u, Sv) \quad \text{for all} \quad u \in \mathcal{D}(T),\ v \in \mathcal{D}(S).$$

7. Suppose that $\mathcal{V}_1 \subset \mathcal{V}_2 \subset \cdots$ are finite dimensional subspaces of a Hilbert space H and that $\mathcal{V} = \cup_{n=1}^{\infty} \mathcal{V}_n$ is dense in H. Let $S : \mathcal{V} \to H$ be linear and such that for some $c > 0$,

$$\mathrm{Re}\,(Su, u) \geq c\|u\|^2 \quad \text{for all} \quad u \in \mathcal{V}.$$

Show that there exists a linear operator T which is an extension of S such that $(-\infty, c) \subset \rho(T)$, $T - c$ is accretive and for each $f \in H$ we can approximate $T^{-1}f$ as follows: for each $n \geq 1$, let $x_n, y_n \in \mathcal{V}_n$ be such that

$$(Sx_n, Sv) = (f, Sv) \quad \text{for all} \quad v \in \mathcal{V}_n$$

$$(y_n, v) = (f - Sx_n, v) \quad \text{for all} \quad v \in \mathcal{V}_n.$$

Then

$$\lim_{n\to\infty} x_n + \frac{1}{2c} y_n = T^{-1}f.$$

8. Can you prove the existence of a weak solution of $\Delta u = f$ when $\Omega \subset \mathbb{R}^2$ is the region between the lines $y = x \pm 1$?

9. Let $\Omega = \{\, (x,t) \mid x \in (0,1),\ t \in \mathbb{R} \,\}$, $f \in L^2(\Omega)$. Show that there exists a unique $u \in L^2(\Omega)$ such that

$$\int_\Omega u(\varphi_t + \varphi_{xx}) = \int_\Omega f\varphi$$

for all φ that are finite linear combinations of functions like $a(t)\sin n\pi x$, with $a \in C_0^1(\mathbb{R})$, $n \geq 1$.

10. In $L^2(0,1)$ define $Au = -u''$ with

$$\mathcal{D}(A) = \{\, u \in AC[0,1] \mid u' \in AC[0,1],\ u'' \in L^2(0,1),\ u(0) = u(1) = 0 \,\}.$$

Starting directly from the definition show that $\mathcal{D}(A) = \mathcal{D}(A^*)$.

11. Suppose that T is a bounded self-adjoint operator with $\sigma(T) \subset [a,b]$ for some $0 < a \leq b < \infty$. Define $A_0 = \frac{2}{a+b} I$ and $A_{n+1} = 2A_n - A_n^2 T$ for $n \geq 0$. Show that

$$\|A_n - T^{-1}\| \leq \frac{1}{a}\left(\frac{b-a}{b+a}\right)^{2^n} \quad \text{for} \quad n \geq 0.$$

12. Suppose that T is a bounded self-adjoint operator in a Hilbert space H with $\sigma(T) \subset [\lambda, \mu]$ for some $0 < \lambda < \mu < \infty$ and let $y \in H$ be given. The following is the **conjugate gradient method** of finding $x = T^{-1}y$. Choose $x_0 \in H$ and let $r_0 = p_0 = y - Tx_0$. If $r_k = y - Tx_k \neq 0$ and $p_k \neq 0$, define

$$x_{k+1} = x_k + a_k p_k \quad \text{where} \quad a_k = \frac{\|r_k\|^2}{(Tp_k, p_k)}$$

$$p_{k+1} = r_{k+1} + b_k p_k \quad \text{where} \quad r_{k+1} = y - Tx_{k+1},\ b_k = \frac{\|r_{k+1}\|^2}{\|r_k\|^2}$$

and show that $(r_i, r_j) = (p_i, r_j) = (Tp_i, p_j) = 0$ for $0 \le i < j \le k+1$ and that if $r_{k+1} \neq 0$, then $p_{k+1} \neq 0$. Note that the orthogonality of residuals r_i implies that, in a finite dimensional space, this process gives 0 residual in a finite number of steps. Show also that if $e_i = (T(x - x_i), x - x_i)$, then $e_{k+1} = e_k - a_k \|r_k\|^2$ and $e_{k+1} \le e_k(1 - \lambda/\mu)$.

13. Suppose that V is a vector space over $\mathbb{C}$ and that the maps $a, b : V \times V \to \mathbb{C}$ have the following properties

    (1) $b(\alpha x + \beta y, z) = \alpha b(x, z) + \beta b(y, z)$
    (2) $b(z, \alpha x + \beta y) = \overline{\alpha} b(z, x) + \overline{\beta} b(z, y)$
    (3) $a(\alpha x + \beta y, z) = \alpha a(x, z) + \beta a(y, z)$
    (4) $\overline{a(x, y)} = a(y, x)$
    (5) $a(x, x) \ge 0$

    for all $\alpha, \beta \in \mathbb{C}$ and all $x, y, z \in V$. Suppose also that there exists $M < \infty$ such that

    $$|b(x, x)| \le M a(x, x) \quad \text{for all} \quad x \in V.$$

    Show that

    $$|b(x, y)| \le 2M \sqrt{a(x, x) a(y, y)} \quad \text{for all} \quad x, y \in V.$$

    Can the same be done if $\mathbb{C}$ is replaced by $\mathbb{R}$? If so, prove it; if not, give a counterexample.

14. Suppose $f \in AC[0, 1]$, $f(0) = f(1) = 0$, f is real valued, $u \in C^1[0, 1]$, $u' \in AC[0, 1]$, $u'(0) = u'(1) = 0$, $u \not\equiv 0$ and $\lambda \in \mathbb{C}$ satisfy

    $$u'' + fu' + (f' + \lambda)u = 0 \quad \text{a.e. on } [0, 1].$$

    Show that $\lambda \in [0, \infty)$.

15. Consider the eigenvalue problem

    $$fu'' = -\lambda u \quad \text{a.e. on } [0, 1],\ u \not\equiv 0,$$

    $$u'(0) = u(1), \quad u(0) = -u'(1),$$

    where $f \in L^\infty(0, 1)$ is real valued and $\inf_x f(x) > 0$. Show that the set of all eigenvalues λ can be written as $\{\lambda_1, \lambda_2, \ldots\} \subset \mathbb{R}$ for some $\lambda_1 < \lambda_2 < \ldots$ such that $\lim_{n \to \infty} \lambda_n = \infty$ and that the set of all finite linear combinations of the eigenvectors u is dense in $L^2(0, 1)$.

16. Find the largest real number λ such that

$$\int_0^1 |u'(x)|^2 dx + 4\int_1^2 |u'(x)|^2 dx \geq \lambda \int_0^2 |u(x)|^2 dx$$

for all $u \in AC[0,2]$ with $u(0) = u(2) = 0$. (Answer: $(2\tan^{-1}\sqrt{5})^2$)

17. Define a linear operator A in $L^2(0,1)$ as follows: $u \in \mathcal{D}(A)$ if and only if

$$u \in AC[0,1],\ u(1) = 0 \text{ and } f \in AC(0,1],\ f' \in L^2(0,1) \text{ where } f(x) = xu'(x).$$

For $u \in \mathcal{D}(A)$ let

$$(Au)(x) = -xu''(x) - u'(x).$$

Show that

(a) A is a self-adjoint operator

(b) $0 \in \rho(A)$ and A^{-1} is compact

(c) $\sigma(A) = \sigma_p(A) = \{\lambda_1, \lambda_2, \cdots\}$ for some $0 < \lambda_1 < \lambda_2 < \cdots$, $\lim_{n\to\infty} \lambda_n = \infty$

(d) $\dim \mathcal{N}(A - \lambda_n) = 1$ for all $n \geq 1$

(e) normalized eigenvectors of A form a complete orthonormal set in $L^2(0,1)$.

18. For $u \in AC[0,1]$ with $u(1) = 0$ define

$$I(u) = \int_0^1 x|u'(x)|^2 dx + \text{Re} \int_0^1 u(x)dx.$$

Find the minimum value that I attains.

19. Suppose that B is closed, densely defined, linear operator in a Hilbert space $\mathcal{H}$. Define $\mathcal{V} = \mathcal{D}(B)$ and

$$[u,v] = (Bu, Bv) + (u,v), \quad |u| = [u,u]^{1/2}$$

$$\mathfrak{F}(u,v) = (Bu, Bv)$$

for $u, v \in \mathcal{V}$. Show that assumptions **H**1, **H**2 and **H**3 are satisfied and that B^*B is the linear operator associated with $\mathfrak{F}$. Show that B^*B is self-adjoint and that $\sigma(B^*B) \subset [0,\infty)$.

20. For $f \in L^2(\mathbb{R})$ find the finite element approximation $u_n \in \mathcal{V}_n$ of $u \in L^2(\mathbb{R})$ satisfying

$$-u''(x) + u(x) = f(x) \quad \text{for} \quad x \in \mathbb{R}.$$

Let $\mathcal{V}_n$ consist of $v \in C(\mathbb{R})$ such that $v(x) = 0$ for $x \leq a_n$ or $x \geq b_n$ and

$$v \text{ is linear on } [x_k^n, x_{k+1}^n] \text{ for } 0 \leq k \leq N_n$$

where $N_n \geq 1$, $-\infty < a_n < b_n < \infty$, $\delta_n = \frac{b_n - a_n}{N_n + 1}$, $x_k^n = a_n + k\delta_n$ are such that

$$\lim_{n\to\infty} a_n = -\infty, \quad \lim_{n\to\infty} b_n = \infty, \quad \lim_{n\to\infty} \delta_n = 0.$$

Prove convergence of the approximations. Show that there do not exist $c_n \to 0$ such that $\|u - u_n\|_2 \leq c_n \|f\|_2$ for all $f \in L^2(\mathbb{R})$.

Chapter 3

Sobolev Spaces

3.1 Introduction

In all our examples of sectorial forms, the norm in the Hilbert space $\mathcal{V}$ involved derivatives of functions. Since, in the examples so far, the functions were functions of one variable defined on an interval, this did not pose any real challange. However, normed spaces of functions of several variables, with norms involving derivatives of functions, are much more difficult to understand and are the the main subject of this chapter. For motivation purposes, it may be a good idea to peek ahead at Section 3.7, to see how the understanding of such spaces can be used to define sectorial forms associated with elliptic problems.

To avoid constant repetition, it is assumed, throughout this chapter and unless specifically restricted, that Ω is an arbitrary nonempty open set in $\mathbb{R}^n$, where $n \in \mathbb{N}$ is also arbitrary. Functions defined on Ω are assumed to be complex valued.

Recall if A and B are subsets of $\mathbb{R}^n$, then $A+B$ is open if A is open; and that $A+B$ is compact if A and B are compact. The boundary of A, denoted by ∂A, is defined to be $\overline{A} \cap \overline{A^c}$. If $K \subset \Omega \neq \mathbb{R}^n$ and K is a nonempty compact set, then $d = \text{dist}(K, \Omega^c) > 0$; moreover, if $\varepsilon \in (0, d)$, then $C = K + \overline{B(0,\varepsilon)} \subset \Omega$ and $\text{dist}(C, \Omega^c) = d - \varepsilon$.

For any two functions f, g on $\mathbb{R}^n$ define their **convolution** $f * g$ by

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy$$

for those x for which the integral exists.

If $\varphi \in C_0(\mathbb{R}^n)$ and $f \in L^1_{loc}(\mathbb{R}^n)$, then it is easy to see that $\varphi * f \in C(\mathbb{R}^n)$. If $\varphi \in C^1_0(\mathbb{R}^n)$, then $(D_i\varphi) * f$ is continuous and, since

$$\int_a^b dx_i \int_{\mathbb{R}^n} (D_i\varphi)(x-y)f(y)dy = (\varphi * f)(x)|_{x_i=a}^{x_i=b},$$

we see that $(D_i\varphi) * f = D_i(\varphi * f)$. Repeating this argument proves the following:

Theorem 3.1.1 *If $\varphi \in C_0^m(\mathbb{R}^n)$, $m \geq 0$, $f \in L^1_{loc}(\mathbb{R}^n)$, then $\varphi * f \in C^m(\mathbb{R}^n)$ and*

$$D^\alpha(\varphi * f) = (D^\alpha \varphi) * f$$

for every multi-index α with $|\alpha| \leq m$.

J_ε, with $\varepsilon \in (0, \infty)$, is said to be a **mollifier** if

$$J_\varepsilon \in C_0^\infty(\mathbb{R}^n), \quad J_\varepsilon \geq 0, \quad J_\varepsilon(x) = 0 \text{ for } |x| \geq \varepsilon, \quad \int_{\mathbb{R}^n} J_\varepsilon = 1.$$

Mollifiers can be constructed as follows. Let $f(x) = \exp(1/x)$ for $x < 0$ and $f(x) = 0$ for $x \geq 0$. It is easy to see that $f \in C^\infty(\mathbb{R})$. Let $\phi(x) = f(|x|^2 - 1)$ for $x \in \mathbb{R}^n$. Note that $\phi \in C_0^\infty(\mathbb{R}^n)$ and that we can define for every $\varepsilon > 0$ a mollifier $J_\varepsilon(x) = c^{-1}\varepsilon^{-n}\phi(x/\varepsilon)$ where $c = \int_{\mathbb{R}^n} \phi$.

If J_ε is a mollifier, $f \in L^p(\Omega)$, $1 \leq p \leq \infty$, then we define $J_\varepsilon * f : \Omega \to \mathbb{C}$ by

$$(J_\varepsilon * f)(x) = \int_\Omega J_\varepsilon(x - y) f(y) dy, \quad x \in \Omega.$$

Theorem 3.1.1 implies that $J_\varepsilon * f \in C^\infty(\Omega)$.

Theorem 3.1.2 *If $K \subset \Omega$ and K is a compact set, then there exists $\varphi \in C_0^\infty(\Omega)$ such that $0 \leq \varphi(x) \leq 1$ for all $x \in \Omega$ and $\varphi(y) = 1$ for all $y \in K$.*

PROOF If K is empty, take $\varphi = 0$. Assume that K is not empty. Let $d = 1$ if $\Omega = \mathbb{R}^n$ and otherwise let $d = \operatorname{dist}(K, \Omega^c)/3$. Let f be the characteristic function of the set $C = K + B(0, d)$ and define $\varphi = J_d * f$, where J_d is a mollifier. $\varphi \in C^\infty(\mathbb{R}^n)$ by Theorem 3.1.1 and

$$\varphi(x) = \int_{B(x,d) \cap C} J_d(x - y) dy, \qquad x \in \mathbb{R}^n.$$

Since $B(x, d) \cap C = B(x, d)$ for $x \in K$ and $B(x, d) \cap C$ is empty when

$$x \notin K + B(0, 2d) \subset \Omega$$

the proof is complete. □

Theorem 3.1.3 (Partition of Unity) *If Γ is a collection of open sets in $\mathbb{R}^n$ whose union is Ω, then there exists $\{\psi_1, \psi_2, \ldots\} \subset C_0^\infty(\Omega)$ with $\psi_i \geq 0$ such that*

(a) each ψ_i has its support in some member of Γ

(b) $\sum_{i=1}^\infty \psi_i(x) = 1$ for every $x \in \Omega$

(c) *for every compact set* $K \subset \Omega$ *there corresponds an integer* m *and open set* $W \supset K$ *such that* $\psi_1(x) + \cdots + \psi_m(x) = 1$ *for every* $x \in W$.

PROOF Let S be a countable dense subset of Ω. The collection of balls $B(x,r)$ with $x \in S$, r a positive rational number, and $\overline{B(x,r)} \subset \mathcal{O}$ for some $\mathcal{O} \in \Gamma$, is a countable collection. Hence, it can be denoted by $\{V_1, V_2, \ldots\}$. It is easy to see that $\Omega = \cup V_i$. Choose $\mathcal{O}_i \in \Gamma$ so that $\overline{V_i} \subset \mathcal{O}_i$.

By Theorem 3.1.2, $\varphi_1, \varphi_2, \ldots$ can be chosen so that $\varphi_i \in C_0^\infty(\mathcal{O}_i)$, $0 \le \varphi_i \le 1$ in $\mathcal{O}_i$ and $\varphi_i = 1$ in V_i. Let $\varphi_i(x) = 0$ for $x \in \Omega \backslash \mathcal{O}_i$. Hence, $\varphi_i \in C_0^\infty(\Omega)$. Define $\psi_1 = \varphi_1$ and

$$\psi_{i+1} = (1 - \varphi_1) \cdots (1 - \varphi_i)\varphi_{i+1} \quad \text{for} \quad i \ge 1.$$

Obviously $\psi_i \in C_0^\infty(\Omega)$, $\psi_i \ge 0$ and $\operatorname{supp}(\psi_i) \subset \mathcal{O}_i$. By induction,

$$\psi_1 + \cdots + \psi_i = 1 - (1 - \varphi_1) \cdots (1 - \varphi_i) \quad \text{for} \quad i \ge 1.$$

Therefore, for each $m \ge 1$,

$$\psi_1(x) + \cdots + \psi_m(x) = 1 \quad \text{if} \quad x \in V_1 \cup \cdots \cup V_m.$$

This gives (b). If K is compact, $K \subset \Omega$, then $K \subset V_1 \cup \cdots \cup V_m$ for some m, and (c) follows. □

REMARK In the above proof it has been also shown that for each Ω there exist nonempty open balls $V_1, V_2, \ldots$ such that $\Omega = \cup V_i = \cup \overline{V_i}$. This fact will be often used.

Theorem 3.1.4 *If* $u \in C_0^m(\Omega)$, $m \ge 0$ *and* J_ε *is a mollifier, then*

(a) $\operatorname{supp}(J_\varepsilon * u) \subset \operatorname{supp}(u) + \overline{B(0,\varepsilon)}$

(b) $J_\varepsilon * u \in C_0^\infty(\Omega)$ *for all sufficiently small* ε

(c) $\lim_{\varepsilon \to 0} \|J_\varepsilon * u - u\|_{m,\infty} = 0$.

PROOF We may assume that $u \not\equiv 0$. Hence, $\operatorname{supp}(u)$ is not empty. Since

$$(J_\varepsilon * u)(x) = \int_{\operatorname{supp}(u) \cap B(x,\varepsilon)} J_\varepsilon(x - y)u(y)dy \quad \text{for} \quad x \in \Omega,$$

it follows that $\operatorname{supp}(J_\varepsilon * u) \subset \operatorname{supp}(u) + \overline{B(0,\varepsilon)}$.

Let $u = 0$ in $\mathbb{R}^n \backslash \Omega$ and note

$$\begin{aligned}(J_\varepsilon * u)(x) - u(x) &= \int_{\mathbb{R}^n} J_\varepsilon(x - y)u(y)dy - \int_{B(x,\varepsilon)} J_\varepsilon(x - y)u(x)dy \\ &= \int_{B(x,\varepsilon)} J_\varepsilon(x - y)(u(y) - u(x))dy.\end{aligned}$$

Therefore

$$\sup_{\Omega} |J_\varepsilon * u - u| \leq \sup_{|x-y|<\varepsilon} |u(y) - u(x)|$$

and, since u is uniformly continuous, (c) follows when $m = 0$. The general case follows by applying this result to each $D^\alpha u$, $|\alpha| \leq m$, since

$$D^\alpha(J_\varepsilon * u) = D^\alpha(u * J_\varepsilon) = (D^\alpha u) * J_\varepsilon = J_\varepsilon * (D^\alpha u) \overset{\varepsilon\to 0}{\longrightarrow} D^\alpha u.$$

□

Lemma 3.1.5 (Young) *If $p, q, r \in [1, \infty]$, $1/p + 1 = 1/q + 1/r$, $f \in L^q(\mathbb{R}^n)$ and $g \in L^r(\mathbb{R}^n)$, then $f * g$ exists a.e. in $\mathbb{R}^n$ and*

$$\|f * g\|_p \leq \|f\|_q \|g\|_r.$$

PROOF If $p = \infty$, Hölder's inequality implies the conclusion. Assume $p < \infty$. Note $r \leq p$ and $q \leq p$. Define

$$u(x,y) = |f(x-y)|^{q/p} |g(y)|^{r/p}, \quad v(x,y) = |f(x-y)|^{1-q/p} |g(y)|^{1-r/p}.$$

Hölder's inequality implies that

$$\|v(x,\cdot)\|_s \leq \|f\|_q^{1-q/p} \|g\|_r^{1-r/p} \quad \text{where } s \text{ is such that } \frac{1}{s} + \frac{1}{p} = 1.$$

We know that $u(x,\cdot)^p \in L^1(\mathbb{R}^n)$ for almost all x and

$$\int\int u(x,y)^p dy dx = \int \left(\int |f(x-y)|^q |g(y)|^r dx \right) dy = \|f\|_q^q \|g\|_r^r.$$

Therefore

$$\begin{aligned} \int |f(x-y)g(y)| dy &\leq \|v(x,\cdot)\|_s \|u(x,\cdot)\|_p \\ \left(\int |f(x-y)g(y)| dy \right)^p &\leq \|f\|_q^{p-q} \|g\|_r^{p-r} \int u(x,y)^p dy \\ \int \left(\int |f(x-y)g(y)| dy \right)^p dx &\leq \|f\|_q^p \|g\|_r^p. \end{aligned}$$

□

Lemma 3.1.6 *Suppose $K_t \in L^1(\mathbb{R}^n)$, $K_t \geq 0$, $\int_{\mathbb{R}^n} K_t = 1$ for $t \in (0, \infty)$ and*

$$\lim_{t\to 0} \int_{|x|>\delta} K_t(x) dx = 0 \quad \text{for each } \delta > 0.$$

Then $\|K_t * u\|_p \leq \|u\|_p$ and $\lim_{t\to 0} \|K_t * u - u\|_p = 0$ for $u \in L^p(\mathbb{R}^n)$, $1 \leq p < \infty$.

PROOF $\|K_t * u\|_p \le \|u\|_p$ follows from Lemma 3.1.5. Pick $\varepsilon > 0$ and choose $g \in C_0(\mathbb{R}^n)$, see Example 1.3.4, such that $\|u - g\|_p < \varepsilon/3$. Note that

$$\begin{aligned} \|K_t * u - u\|_p &\le \|K_t * u - K_t * g\|_p + \|K_t * g - g\|_p + \|g - u\|_p \\ &< 2\varepsilon/3 + \|K_t * g - g\|_p. \end{aligned}$$

The following calculation shows that $\|K_t * g - g\|_p < \varepsilon/3$ for all sufficiently small t, which completes the proof. $\int K_t = 1$ and Hölder's inequality imply

$$\begin{aligned} (K_t * g)(x) - g(x) &= \int_{\mathbb{R}^n} K_t(s)(g(x-s) - g(x))ds \\ |(K_t * g)(x) - g(x)|^p &\le \int_{\mathbb{R}^n} K_t(s)|g(x-s) - g(x)|^p ds \\ \int_{\mathbb{R}^n} |(K_t * g)(x) - g(x)|^p dx &\le \int_{\mathbb{R}^n} K_t(s) I(s) ds, \end{aligned}$$

where $I(s) = \int |g(x-s) - g(x)|^p dx \le 2^p \|g\|_p^p$. Since $g \in C_0(\mathbb{R}^n)$, there exists $\delta > 0$ such that $I(s) < (\varepsilon/3)^p/2$ for $|s| < \delta$; hence

$$\int_{\mathbb{R}^n} |(K_t * g)(x) - g(x)|^p dx \le (\varepsilon/3)^p/2 + 2^p \|g\|_p^p \int_{|s|>\delta} K_t(s) ds \xrightarrow{t \to 0} (\varepsilon/3)^p/2$$

and $\|K_t * g - g\|_p < \varepsilon/3$ for small enough t. □

Theorem 3.1.7 *If $1 \le p < \infty$, $u \in L^p(\Omega)$ and J_ε is a mollifier, then*

(a) $J_\varepsilon * u \in C^\infty(\Omega)$, $\|J_\varepsilon * u\|_p \le \|u\|_p$

(b) $\lim_{\varepsilon \to 0} \|J_\varepsilon * u - u\|_p = 0$.

Moreover, $C_0^\infty(\Omega)$ is dense in $L^p(\Omega)$.

PROOF (a) and (b) follow from Lemma 3.1.6 by letting $u = 0$ in $\mathbb{R}^n \backslash \Omega$. Pick $\varepsilon > 0$ and $g \in C_0(\Omega)$, see Example 1.3.4, such that $\|u - g\|_p < \varepsilon/2$. Then

$$\|u - J_\delta * g\|_p \le \|u - g\|_p + \|g - J_\delta * g\|_p < \varepsilon/2 + \|g - J_\delta * g\|_p;$$

hence $\|u - J_\delta * g\|_p < \varepsilon$ for all small enough δ and, since Theorem 3.1.4 implies that $J_\delta * g \in C_0^\infty(\Omega)$ for all small enough δ, the proof is complete. □

Corollary 3.1.8 *If $1 \le p \le q < \infty$, $u \in L^p(\Omega) \cap L^q(\Omega)$ and $\varepsilon > 0$, then there exists $v \in C_0^\infty(\Omega)$ such that $\|u - v\|_p < \varepsilon$ and $\|u - v\|_q < \varepsilon$.*

PROOF Choose a nonempty open set ω such that $\overline{\omega}$ is compact, $\overline{\omega} \subset \Omega$ and

$$\int_{\Omega\setminus\omega} |u|^p < \varepsilon^p/2, \quad \int_{\Omega\setminus\omega} |u|^q < \varepsilon^q/2.$$

Theorem 3.1.7 implies that we can choose $v \in C_0^\infty(\omega)$ such that

$$\mu(\omega)^{q/p-1} \int_\omega |u-v|^q < 2^{-q/p}\varepsilon^q, \quad \int_\omega |u-v|^q < \varepsilon^q/2$$

and hence if $v = 0$ in $\Omega\setminus\omega$, then $\|u-v\|_p < \varepsilon$ and $\|u-v\|_q < \varepsilon$. □

3.2 Fourier Transform

Let $\mathcal{S}$ denotes the set of all complex valued $f \in C^\infty(\mathbb{R}^n)$ such that

$$\sup_{x\in\mathbb{R}^n} |x^\beta (D^\alpha f)(x)| < \infty \quad \text{for all multi-indices } \alpha, \beta.$$

$\mathcal{S}$ is called the space of **rapidly decreasing functions** or the **Schwartz space**. Since $C_0^\infty(\mathbb{R}^n) \subset \mathcal{S}$, Theorem 3.1.7 implies that $\mathcal{S}$ is dense in $L^2(\mathbb{R}^n)$.

For $f \in \mathcal{S}$ and $x \in \mathbb{R}^n$, define

$$\hat{f}(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-ix\cdot s} f(s) ds \tag{3.1}$$

where $x \cdot s = \sum_{k=1}^n x_k s_k$. $\hat{f}$ is called the **Fourier transform** of f. The mapping $f \to \hat{f}$ is also called the Fourier transform.

EXAMPLE 3.2.1 Let $\phi(s) = e^{-z|s|^2}$, $\operatorname{Re} z > 0$. Note $\hat{\phi}(x_1, \ldots, x_n) = I(x_1)\cdots I(x_n)$ where

$$I(t) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-itr - zr^2} dr.$$

Differentiation of I and integration by parts of the result gives $I'(t) = -tI(t)/(2z)$. Hence $I(t) = I(0)e^{-t^2/(4z)}$. To find $I(0)$ consider

$$K(b) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-(a+ib)r^2} dr, \quad a > 0, b \in \mathbb{R}$$

and note that differentiation of K and integration by parts of the result gives that $2K'(b) = -iK(b)/(a+ib)$. Hence $K(b) = 1/\sqrt{2(a+ib)}$. Therefore $I(0) = 1/\sqrt{2z}$ and

$$\hat{\phi}(x) = (\sqrt{2z})^{-n} e^{-|x|^2/(4z)}, \quad \operatorname{Re}\sqrt{2z} > 0.$$

Lemma 3.2.2 *If $f \in \mathcal{S}$, then $\hat{f} \in \mathcal{S}$, and for all multi-indices α, β we have*

$$x^\beta (D^\alpha \hat{f})(x) = (-i)^{|\alpha|+|\beta|} \widehat{D^\beta g}(x) \quad \text{where } g(s) = s^\alpha f(s); \tag{3.2}$$

moreover, $(P(D)f)^\wedge(x) = P(ix)\hat{f}(x)$ for every polynomial P and every $x \in \mathbb{R}^n$.

PROOF The DCT implies that $\hat{f} \in C(\mathbb{R}^n)$, from which it is easy to verify that $i^{|\alpha|} D^\alpha \hat{f} = \hat{g}$. Hence, $\hat{f} \in C^\infty(\mathbb{R}^n)$. Integration by parts then implies (3.2) and therefore $\hat{f} \in \mathcal{S}$. □

Lemma 3.2.3 *If $f, g \in \mathcal{S}$, then $\int f\hat{g} = \int \hat{f} g$.*

PROOF Apply the Fubini Theorem on $\int\int f(x)g(y)e^{-ix\cdot y}dxdy$. □

For any function f defined on $\mathbb{R}^n$, let $(Rf)(x) = f(-x)$ for $x \in \mathbb{R}^n$.
For $f \in \mathcal{S}$ define $\check{f} \in \mathcal{S}$ by $\check{f} = R\hat{f}$, i.e.

$$\check{f}(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{ix\cdot s} f(s) ds \quad \text{for} \quad x \in \mathbb{R}^n.$$

$\check{f}$ is called the **inverse Fourier transform** of f because of the following:

Theorem 3.2.4 *If $f \in \mathcal{S}$ and $h = \hat{f}$, then $\check{h} = f$.*

PROOF Let $\phi(x) = e^{-|x|^2/2}$. Using $g(x) = \phi(x\varepsilon)e^{ix\cdot y}$, with $\varepsilon > 0$, $y \in \mathbb{R}^n$, in Lemma 3.2.3 gives

$$\int f(x)\phi((x-y)/\varepsilon)\varepsilon^{-n}dx = \int \hat{f}(x)\phi(x\varepsilon)e^{ix\cdot y}dx$$

$$f(y) \xleftarrow{\varepsilon\to 0} (2\pi)^{-n/2} \int f(y+\varepsilon t)\phi(t)dt = (2\pi)^{-n/2} \int \hat{f}(x)\phi(x\varepsilon)e^{ix\cdot y}dx \xrightarrow{\varepsilon\to 0} \check{h}(y).$$

□

Lemma 3.2.5 *If $f, g \in \mathcal{S}$, then $\int f\overline{g} = \int \hat{f}\overline{\hat{g}}$ and $\|f\|_2 = \|\hat{f}\|_2$.*

PROOF Let $h = \hat{g}$. Theorem 3.2.4 implies $\overline{g} = \check{\overline{h}}$ and therefore Lemma 3.2.3 implies the assertions. □

Define the **Fourier transform** $\mathcal{F}$ on $L^2(\mathbb{R}^n)$ as follows. Pick $f \in L^2(\mathbb{R}^n)$ and choose $f_1, f_2, \ldots$ in $\mathcal{S}$ so that $f_k \to f$ in $L^2(\mathbb{R}^n)$. Lemma 3.2.5 implies that $\hat{f}_k$ converge in $L^2(\mathbb{R}^n)$ and that the limit does not depend on the particular selection of the sequence $\{f_k\}$. Define $\mathcal{F}f = \lim_{k\to\infty} \hat{f}_k$. Note $\mathcal{F}f = \hat{f}$ for $f \in \mathcal{S}$.

Theorem 3.2.6 *The Fourier transform $\mathcal{F}$ is a one-to-one linear map from $L^2(\mathbb{R}^n)$ onto $L^2(\mathbb{R}^n)$, and $\mathcal{F}^{-1} = R\mathcal{F}$. Moreover, $(f,g) = (\mathcal{F}f, \mathcal{F}g)$ for $f, g \in L^2(\mathbb{R}^n)$ and*

$$(\mathcal{F}h)(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-ix\cdot s} h(s) ds \ \text{ a.e. } \ \text{for} \quad h \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n).$$

PROOF The definiton of $\mathcal{F}$ and Lemma 3.2.5 imply that $\mathcal{F}$ is linear and $(f,g) = (\mathcal{F}f, \mathcal{F}g)$ for $f, g \in L^2(\mathbb{R}^n)$. Hence, $\mathcal{F}$ is also one-to-one.

Pick $g \in L^2(\mathbb{R}^n)$ and choose $g_k \in \mathcal{S}$ such that $g_k \to g$ in $L^2(\mathbb{R}^n)$. Define $f_k = R\hat{g}_k \in \mathcal{S}$ and note that f_k converge in $L^2(\mathbb{R}^n)$ to $R\mathcal{F}g$. Since $f_k = \widehat{Rg_k}$, Theorem 3.2.4 implies that $R\hat{f}_k = \check{f}_k = Rg_k$. Hence, $\hat{f}_k = g_k$ and the definition of $\mathcal{F}$ implies $\mathcal{F}(R\mathcal{F}g) = g$. Hence, $\mathcal{F}$ is onto and $\mathcal{F}^{-1} = R\mathcal{F}$.

If $h \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then Corollary 3.1.8 and the definition of $\mathcal{F}$ imply the expression for $\mathcal{F}h$. □

Corollary 3.2.7 *If $f \in L^2(\mathbb{R}^n)$, $g \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then $g * f \in L^2(\mathbb{R}^n)$ and*

$$\mathcal{F}(g * f) = (2\pi)^{n/2}(\mathcal{F}g)\mathcal{F}f.$$

PROOF Choose $f_k \in \mathcal{S}$ such that $\|f_k - f\|_2 \to 0$. Let $h_k = g * f_k$ and $h = g * f$. Lemma 3.1.5 implies $h_k \in L^2 \cap L^1$ and $\|h_k - h\|_2 \to 0$. Theorem 3.2.6 implies

$$(\mathcal{F}h_k)(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-ix\cdot s} g(s-y) f_k(y) dy\, ds$$

and hence $\mathcal{F}h_k = (2\pi)^{n/2}(\mathcal{F}g)\mathcal{F}f_k$ by the Fubini Theorem. Letting $k \to \infty$ completes the proof. □

Theorem 3.2.8 *Suppose f is such that the function $e^{-ax}f(x)$, $x \in (0,\infty)$, belongs to $L^2(0,\infty)$ for some $a \in \mathbb{R}$. Define the* **Laplace transform** *of f to be*

$$F(s) = \int_0^\infty e^{-sx} f(x) dx \quad \textit{for} \quad s \in \mathbb{C},\ \operatorname{Re} s > a.$$

Then F is an analytic function and, if we define

$$g_c(\lambda, x) = \frac{1}{2\pi} \int_{-c}^{c} e^{(\lambda + i\omega)x} F(\lambda + i\omega) d\omega \quad \textit{for} \quad x \in \mathbb{R},\ c > 0,\ \lambda > a,$$

then

$$\lim_{c\to\infty} \int_0^\infty |g_c(\lambda, x) - f(x)|^2 e^{-2\lambda x} dx = 0 \quad \textit{for} \quad \lambda > a.$$

PROOF Since

$$|e^{-(s+h)x} - e^{-sx} + hxe^{-sx}||f(x)| \le |hx|^2 e^{(|h| + a - \operatorname{Re} s)x} e^{-ax} |f(x)|,$$

we have that F is analytic. Fix any $\lambda > a$. For $x > 0$ define $\phi(x) = e^{-\lambda x} f(x)$ and let $\phi(x) = 0$ for $x \le 0$. Note that

$$F(\lambda + i\omega) = \sqrt{2\pi}(\mathcal{F}\phi)(\omega) \quad \text{for almost all } \omega \in \mathbb{R}.$$

Since $\chi_{(-c,c)}\mathcal{F}\phi$ converge in $L^2(\mathbb{R})$ to $\mathcal{F}\phi$ as $c \to \infty$, Theorem 3.2.6 implies that $\mathcal{F}^{-1}\chi_{(-c,c)}\mathcal{F}\phi$ converges to ϕ and, since $(\mathcal{F}^{-1}\chi_{(-c,c)}\mathcal{F}\phi)(x) = e^{-\lambda x}g_c(\lambda, x)$, we are done. □

If D is an open set in $\mathbb{C}^n$ and f is a continuous complex function in D, then f is said to be **analytic** in D if it is analytic in each variable separately. f is said to be an **entire** function in $\mathbb{C}^n$ if it is analytic in $\mathbb{C}^n$.

Lemma 3.2.9 *If $f \in C_0^\infty(\mathbb{R}^n)$ and $\hat{f}(x)$ is defined by (3.1) for every $x \in \mathbb{C}^n$, then $\hat{f}$ is entire and $(P(D)f)\hat{}(x) = P(ix)\hat{f}(x)$ for every polynomial P and every $x \in \mathbb{C}^n$.*

PROOF For $x \in \mathbb{C}^n$, $s \in \mathbb{R}^n$, $k \geq 1$, let

$$g_k(s) = \sum_{|\alpha| \leq k} \frac{x^\alpha (-is)^\alpha}{\alpha!} = \sum_{j=0}^{k} \frac{(-ix \cdot s)^j}{j!} \to e^{-ix \cdot s}$$

and, since $|g_k(s)| \leq e^{|x||s|}$, the DCT implies that

$$\hat{f}(x) = \sum_\alpha c_\alpha x^\alpha \quad \text{where} \quad c_\alpha = \frac{(-i)^{|\alpha|}}{\alpha!(2\pi)^{n/2}} \int_{\mathbb{R}^n} s^\alpha f(s)ds;$$

since the series is absolutely convergent, $\hat{f}$ is entire. Since $(P(D)f)\hat{} - P(i\cdot)\hat{f}$ is entire and, by Lemma 3.2.2, equal to 0 on $\mathbb{R}^n$, it is equal to 0 on the whole $\mathbb{C}^n$. □

EXAMPLE 3.2.10 The following **Heisenberg inequality** will be proven

$$\left(\int_{\mathbb{R}^n} |x - X|^2 |\phi(x)|^2 dx\right)\left(\int_{\mathbb{R}^n} |p - P|^2 |\hat{\phi}(p)|^2 dp\right) \geq n^2/4 \tag{3.3}$$

for $\phi \in \mathcal{S}$ with $\|\phi\|_2 = 1$ and any $X, P \in \mathbb{R}^n$.

Define $\psi(x) = e^{-iP \cdot x}\phi(x + X)$ and note

$$0 = \int_{\mathbb{R}^n} D_k(x_k|\psi(x)|^2)dx = 1 + 2\mathrm{Re}\int_{\mathbb{R}^n} x_k \overline{\psi(x)} D_k\psi(x)dx$$

$$\begin{aligned}
1/2 &\leq \left(\int_{\mathbb{R}^n} x_k^2 |\psi(x)|^2 dx\right)^{1/2}\left(\int_{\mathbb{R}^n} |D_k\psi(x)|^2 dx\right)^{1/2} \\
&= \left(\int_{\mathbb{R}^n} x_k^2 |\psi(x)|^2 dx\right)^{1/2}\left(\int_{\mathbb{R}^n} |p_k\hat{\psi}(p)|^2 dp\right)^{1/2} \\
n/2 &\leq \left(\int_{\mathbb{R}^n} |x|^2 |\psi(x)|^2 dx\right)^{1/2}\left(\int_{\mathbb{R}^n} |p|^2 |\hat{\psi}(p)|^2 dp\right)^{1/2}.
\end{aligned}$$

A change of variables implies (3.3).

If ϕ is taken to be a wave function of a quantum mechanical system and X, P are expected values of the position and momentum, then the uncertainty of the position is determined by

$$\sigma_x = \left(\int_{\mathbb{R}^n} |x - X|^2 |\phi(x)|^2 dx \right)^{1/2}$$

and uncertainty of the momentum is determined by

$$\sigma_p = \left(\int_{\mathbb{R}^n} |p - P|^2 |\hat{\phi}(p)|^2 dp \right)^{1/2}.$$

Thus, $\sigma_x \sigma_p \geq n/2$. So, if you know the position (small σ_x), then (3.3) implies that you do not know momentum (large σ_p). This is known as the **uncertainty principle**.

In the rest of the book we shall often refer to the following:

EXAMPLE 3.2.11 If $f \in \mathcal{S}$, $\lambda \in \mathbb{C} \backslash (-\infty, 0]$, then there exists a unique $u \in \mathcal{S}$ such that

$$\lambda u - \Delta u = f. \tag{3.4}$$

This follows from the observation $\hat{u}(\xi) = (\lambda + |\xi|^2)^{-1} \hat{f}(\xi)$ which also enables us to obtain an explicit formula for u:

$$\begin{aligned} u(x) &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{ix\cdot\xi} (\lambda + |\xi|^2)^{-1} \hat{f}(\xi) d\xi \\ &= (2\pi)^{-n/2} \mu^{-1} \int_{\mathbb{R}^n} e^{ix\cdot\xi} (\mu + |\xi|^2/\mu)^{-1} \hat{f}(\xi) d\xi, \quad \mu \equiv \sqrt{\lambda},\ \operatorname{Re}\mu > 0 \\ &= (2\pi)^{-n/2} \mu^{-1} \int_0^\infty \left(\int_{\mathbb{R}^n} e^{ix\cdot\xi} e^{-r\mu - r|\xi|^2/\mu} \hat{f}(\xi) d\xi \right) dr. \end{aligned}$$

Using Lemma 3.2.3 and Example 3.2.1 gives

$$u(x) = (4\pi)^{-n/2} \sqrt{\mu}^{\,n-2} \int_0^\infty \int_{\mathbb{R}^n} r^{-n/2} e^{-\mu r - \mu|x-\xi|^2/(4r)} f(\xi) d\xi\, dr.$$

Since the integrand is absolutely integrable, the Fubini Theorem implies

$$u(x) = \int_{\mathbb{R}^n} G_\lambda(x - s) f(s) ds \tag{3.5}$$

where

$$G_\lambda(s) = (4\pi)^{-n/2} \sqrt{\mu}^{\,n-2} \int_0^\infty r^{-n/2} e^{-\mu r - \mu|s|^2/(4r)} dr, \quad \operatorname{Re}\sqrt{\mu} > 0 \tag{3.6}$$

and

$$\|G_\lambda\|_1 \leq \frac{1}{|\lambda| (\cos \frac{\arg(\lambda)}{2})^{1+\frac{n}{2}}}, \quad |\arg(\lambda)| < \pi. \tag{3.7}$$

When $\lambda > 0$, the expression (3.6) simplifies to

$$G_\lambda(s) = (4\pi)^{-n/2} \int_0^\infty r^{-n/2} e^{-r\lambda - |s|^2/(4r)} dr \quad \text{and} \quad \|G_\lambda\|_1 = 1/\lambda. \tag{3.8}$$

When $n = 1$, then differentiation under the integral in (3.6) and a change of variables $t = s^2/(4r)$ give $G'_\lambda(s) = -\mu G_\lambda(s)$ for $s > 0$. Hence $G_\lambda(s) = G_\lambda(0)e^{-\mu|s|}$ and therefore, in view of Example 3.2.1, we have

$$G_\lambda(s) = \int_0^\infty e^{-\mu r - \mu s^2/(4r)} \frac{dr}{\sqrt{4\pi\mu r}} = \frac{e^{-\mu|s|}}{2\mu} \quad (\text{when } n = 1). \tag{3.9}$$

If we replace μr in (3.6) by $\mu r t^4$ and differentiate with respect to t, we in effect reduce n by 2. Knowing (3.9) we can thus obtain

$$G_\lambda(s) = \frac{e^{-\mu|s|}}{4\pi|s|} \quad (\text{when } n = 3). \tag{3.10}$$

It is easy to deduce from (3.6) (Exercise 3) that

$$G_\lambda(s) = (2\pi)^{-\frac{n}{2}} (\mu/|s|)^{\frac{n}{2}-1} \int_0^\infty e^{-|s|\mu \cosh t} \cosh(t(\tfrac{n}{2} - 1)) dt \tag{3.11}$$

which can be expressed in terms of the modified Bessel function K as

$$G_\lambda(s) = (2\pi)^{-\frac{n}{2}} (\mu/|s|)^{\frac{n}{2}-1} K_{\frac{n}{2}-1}(|s|\mu). \tag{3.12}$$

When $n = 2$, the series representation of K_0 gives

$$G_\lambda(s) = \frac{1}{4\pi} \sum_{k=0}^\infty \frac{|s|^{2k}\lambda^k}{4^k(k!)^2} \left(2b_k - \ln \frac{|s|^2\lambda}{4}\right) \quad (\text{when } n = 2) \tag{3.13}$$

where $-b_0 = \gamma = .5772\ldots$ is the Euler constant and $b_k = b_{k-1} + 1/k$ for $k \geq 1$.

EXAMPLE 3.2.12 If $1 \leq q \leq p \leq \infty$ and $\alpha = \frac{n}{2}(\frac{1}{q} - \frac{1}{p}) < 1$, then there exists c such that

$$\|u\|_p \leq c\|\Delta u\|_q^\alpha \|u\|_q^{1-\alpha} \quad \text{for all} \quad u \in \mathcal{S}. \tag{3.14}$$

To see this let us first estimate L^t norm of G_λ as given by (3.8), where $\lambda \in (0, \infty)$ and $t \in [1, \infty]$ is such that $1/p + 1 = 1/t + 1/q$. If $t \in (1, \infty)$ Hölder's inequality implies

$$G_\lambda(s)^t \leq (4\pi)^{-tn/2} \int_0^\infty r^{-\alpha - n/2} e^{-r\lambda - t|s|^2/(4r)} dr \left(\int_0^\infty r^{-\alpha} e^{-r\lambda} dr\right)^{t-1};$$

hence, for all t,

$$\|G_\lambda\|_t \leq (4\pi)^{-\alpha} \Gamma(1-\alpha) \lambda^{\alpha-1}.$$

Since $u = G_\lambda * (\lambda u - \Delta u)$, Lemma 3.1.5 implies

$$\|u\|_p \leq (4\pi)^{-\alpha} \Gamma(1-\alpha)(\lambda^\alpha \|u\|_q + \lambda^{\alpha-1} \|\Delta u\|_q)$$

which, together with the fact

$$\inf_{0<\varepsilon<\infty}(A\varepsilon^{-\gamma}+B\varepsilon^{\delta}) \le 2A^{\frac{\delta}{\gamma+\delta}}B^{\frac{\gamma}{\gamma+\delta}} \quad \text{for} \quad A,B,\gamma,\delta \in [0,\infty),\ \gamma+\delta>0, \tag{3.15}$$

implies (3.14).

If $1 \le q \le p \le \infty$ and $\beta = \frac{1}{2} + \frac{n}{2}(\frac{1}{q}-\frac{1}{p}) < 1$, then one can show in exactly the same way (Exercise 4) that there exists c such that

$$\|D_k u\|_p \le c\|\Delta u\|_q^{\beta}\|u\|_q^{1-\beta} \quad \text{for all} \quad u \in \mathcal{S},\ k=1,\cdots,n. \tag{3.16}$$

EXAMPLE 3.2.13 Let $\Omega = \mathbb{R}^n \times (0,\infty)$. We shall prove that if $p \in [1,\infty)$, $f \in L^p(\mathbb{R}^n)$, then there exists u which solves the following initial value parabolic problem:

$$u \in C^{\infty}(\Omega),\ \ \|u(\cdot,t)\|_p \le \|f\|_p \text{ for } t>0, \tag{3.17}$$

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x_1^2}+\cdots+\frac{\partial^2 u}{\partial x_n^2} \qquad \text{in } \Omega, \tag{3.18}$$

$$\lim_{t\to 0+}\|u(\cdot,t)-f\|_p = 0. \tag{3.19}$$

Applying the Fourier transform *formally* gives

$$\begin{aligned}
\hat{u}_t(\xi,t) &= -|\xi|^2\hat{u}(\xi,t)\\
\hat{u}(\xi,t) &= e^{-t|\xi|^2}\hat{f}(\xi)\\
u(x,t) &= (2\pi)^{-n/2}\int_{\mathbb{R}^n} e^{ix\cdot\xi-t|\xi|^2}\hat{f}(\xi)d\xi\\
u(x,t) &= \int_{\mathbb{R}^n} K(x-y,t)f(y)dy
\end{aligned} \tag{3.20}$$

using Lemma 3.2.3 and Example 3.2.1, where

$$K(x,t) = (4\pi t)^{-n/2}e^{-|x|^2/(4t)}. \tag{3.21}$$

Instead of trying to justify this procedure, *define* u by (3.20). It is easy to verify that $u \in C^{\infty}(\Omega)$ and that it satisfies (3.18). Since $K>0$, $\|K(\cdot,t)\|_1 = 1$ and

$$\int_{|x|>\delta} K(x,t)dx = \int_{|s|>\delta/\sqrt{t}} K(s,1)ds \xrightarrow{t\to 0} 0 \quad \text{for each } \delta>0\ ,$$

Lemma 3.1.6 implies (3.19) and the bound in (3.17).

EXAMPLE 3.2.14 Let $\Omega = \mathbb{R}^n \times (0,\infty)$. We shall prove that if $p \in [1,\infty)$, $f \in L^p(\mathbb{R}^n)$, then there exists u which solves the following boundary value problem:

$$u \in C^{\infty}(\Omega),\ \ \|u(\cdot,y)\|_p \le \|f\|_p \text{ for } y>0, \tag{3.22}$$

$$\frac{\partial^2 u}{\partial x_1^2}+\cdots+\frac{\partial^2 u}{\partial x_n^2}+\frac{\partial^2 u}{\partial y^2} = 0 \qquad \text{in } \Omega, \tag{3.23}$$

$$\lim_{y\to 0+} \|u(\cdot,y) - f\|_p = 0. \tag{3.24}$$

Applying the Fourier transform *formally* to (3.23) gives

$$\hat{u}_{yy}(\xi,y) - |\xi|^2\hat{u}(\xi,y) = 0$$

$$\hat{u}(\xi,y) = a(\xi)e^{-|\xi|y} + b(\xi)\mathrm{e}^{|\xi|y};$$

this, (3.22) and the boundary condition (3.24) suggest

$$\begin{aligned}
\hat{u}(\xi,y) &= \mathrm{e}^{-y|\xi|}\hat{f}(\xi) \\
u(x,y) &= (2\pi)^{-n/2}\int_{\mathbb{R}^n} e^{ix\cdot\xi - y|\xi|}\hat{f}(\xi)d\xi \\
u(x,y) &= \int_{\mathbb{R}^n} P(x-s,y)f(s)ds \quad \text{for} \quad x\in\mathbb{R}^n,\ y>0,
\end{aligned} \tag{3.25}$$

where

$$P(x,y) = (2\pi)^{-n}\int_{\mathbb{R}^n} \mathrm{e}^{i\xi\cdot x}e^{-|\xi|y}d\xi.$$

Using (3.9) with $\mu = 1$, $s = |\xi|y$, gives

$$P(x,y) = (2\pi)^{-n}\int_{\mathbb{R}^n} e^{i\xi\cdot x}\left(\int_0^\infty (\pi r)^{-1/2}e^{-r-|\xi|^2y^2/(4r)}dr\right)d\xi.$$

The Fubini Theorem and Example 3.2.1 imply

$$P(x,y) = \pi^{-\frac{n+1}{2}}y^{-n}\int_0^\infty r^{\frac{n+1}{2}-1}e^{-r-r|x|^2/y^2}dr \tag{3.26}$$

$$P(x,y) = \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}}\frac{y}{(|x|^2+y^2)^{\frac{n+1}{2}}}. \tag{3.27}$$

Using (3.26) it is easy to see that $\|P(\cdot,y)\|_1 = 1$; thus we can *define* u by (3.25) and Lemma 3.1.5 implies the bound in (3.22). (3.24) follows from Lemma 3.1.6 since

$$\int_{|x|>\delta} P(x,y)dx = \int_{|s|>\delta/y} P(s,1)ds \overset{y\to 0}{\longrightarrow} 0 \quad \text{for each } \delta > 0.$$

Let $z = (x,y)$. By induction, $D_z^\alpha P(z) = h_{|\alpha|+1}(z)|z|^{-n-1-2|\alpha|}$ for every multi-index α, where h_m is a homogeneous polynomial of degree m in $n+1$ variables. Thus $|D_z^\alpha P(z)| \le c_\alpha|z|^{-n-|\alpha|}$ which implies that $D_z^\alpha P(\cdot,y) * f \in C(\Omega)$. Hence, the Fubini Theorem and induction imply that $u \in C^\infty(\Omega)$. (3.23) follows by verification.

EXAMPLE 3.2.15 Consider the initial value problem for the wave equation:

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x_1^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2} \quad \text{for} \quad x\in\mathbb{R}^n,\ t\ge 0$$

$$u(x,0) = f(x) \quad \text{for} \quad x\in\mathbb{R}^n$$

$$\frac{\partial u}{\partial t}(x,0) = g(x) \quad \text{for} \quad x \in \mathbb{R}^n.$$

Applying the Fourier transform *formally* gives

$$\hat{u}_{tt}(\xi,t) + |\xi|^2 \hat{u}(\xi,t) = 0$$

$$\hat{u}(\xi,0) = \hat{f}(\xi)$$

$$\hat{u}_t(\xi,0) = \hat{g}(\xi)$$

and hence

$$\hat{u}(\xi,t) = \hat{f}(\xi)\cos|\xi|t + \hat{g}(\xi)\frac{\sin|\xi|t}{|\xi|}. \tag{3.28}$$

If $f, g \in \mathcal{S}$, then for each $t \in \mathbb{R}$, the right hand side of the above equation is also in $\mathcal{S}$ which enables us to define $u(\cdot,t) \in \mathcal{S}$ so that (3.28) holds (Exercise 5). It is easy to see that this u satisfies the initial value problem. When $n = 1$, the inversion of (3.28) gives

$$u(x,t) = \frac{1}{2}(f(x+t) + f(x-t)) + \frac{1}{2}\int_{x-t}^{x+t} g(s)ds. \tag{3.29}$$

When $f, g \in L^2(\mathbb{R}^n)$, one can still use (3.28) to define $u(\cdot,t) \in L^2(\mathbb{R}^n)$. In particular, one can show that (3.29) remains valid when $n = 1$. Thus, u may not be differentiable in this case and, since it still seems to be the appropriate solution, we need to generalize the notion of a solution.

3.3 Distributions

A **distribution** f in Ω is a linear map from $C_0^\infty(\Omega)$ to $\mathbb{C}$ such that for every compact set $K \subset \Omega$ there exist $C < \infty$ and an integer $k \geq 0$ such that

$$|f(\phi)| \leq C\|\phi\|_{k,\infty} \quad \text{for all } \phi \in C_0^\infty(\Omega) \text{ with } \operatorname{supp}(\phi) \subset K.$$

If the same k can be chosen for every compact set K, then the smallest of such k is said to be the order of f. The set of all distributions in Ω is denoted by $\mathcal{D}'(\Omega)$ since $C_0^\infty(\Omega)$ is sometimes denoted by $\mathcal{D}(\Omega)$.

For $u \in L^1_{loc}(\Omega)$, let f_u denote the distribution in Ω defined by

$$f_u(\phi) = \int_\Omega u\phi \quad \text{for all} \quad \phi \in C_0^\infty(\Omega). \tag{3.30}$$

If $X \subset L^1_{loc}(\Omega)$ and f is a distribution in Ω, then the notation $f \in X$ is sometimes used in the literature to indicate that $f = f_u$ for some $u \in X$.

For every distribution f in Ω and every multi-index α, define distribution $\mathsf{D}^\alpha f$ in Ω, called the αth **derivative of the distribution** f, by

$$(\mathsf{D}^\alpha f)(\phi) = (-1)^{|\alpha|} f(D^\alpha \phi) \quad \text{for all} \quad \phi \in C_0^\infty(\Omega).$$

Lemma 2.4.4 implies that $\mathsf{D}^\alpha f_u = f_{D^\alpha u}$ when $u \in C^{|\alpha|}(\Omega)$.

If f is a distribution in $\mathbb{R}^n$ and $\varphi \in C_0^\infty(\mathbb{R}^n)$, then the **convolution** $f * \varphi$ is a complex function of n real variables defined by

$$(f * \varphi)(x) = f(\varphi_x) \text{ where } \varphi_x(y) = \varphi(x - y) \text{ for } x, y \in \mathbb{R}^n. \tag{3.31}$$

Note that if $u \in L^1_{loc}(\mathbb{R}^n)$, then $f_u * \varphi = \varphi * u = u * \varphi$.

For $c \in \Omega$, define the **delta function** δ_c to be the distribution in Ω given by

$$\delta_c(\phi) = \phi(c) \quad \text{for all} \quad \phi \in C_0^\infty(\Omega).$$

Let $P \not\equiv 0$ be a polynomial in n variables with complex coefficients. A distribution G in $\mathbb{R}^n$ is said to be a **fundamental solution** of the operator $P(D)$ if

$$P(\mathsf{D})G = \delta_0,$$

i.e., $G(P(-D)\varphi) = \varphi(0)$ for all $\varphi \in C_0^\infty(\mathbb{R}^n)$. The main purpose of this section is to prove the Malgrange-Ehrenpreis Theorem which states that *fundamental solutions always exist.* The following Theorem 3.3.2 shows that if G is a fundamental solution of the operator $P(D)$ and $v \in C_0^\infty(\mathbb{R}^n)$, then $u = G * v \in C^\infty(\mathbb{R}^n)$ and

$$P(D)u = v.$$

EXAMPLE 3.3.1 (3.5) with $x = 0$, together with (3.4) and (3.8), implies

$$\int_{\mathbb{R}^n} G_\lambda(s)\,((\lambda - \Delta)u)(s)\,ds = u(0) \quad \text{for} \quad u \in C_0^\infty(\mathbb{R}^n) \tag{3.32}$$

when $\lambda > 0$ and hence f_{G_λ} is a fundamental solution of $\lambda - \Delta$. Since

$$0 < G_\lambda(x) \le G_0(x) = \tfrac{1}{4}\pi^{-n/2}\Gamma(\tfrac{n}{2} - 1)|x|^{2-n} \text{ for } \lambda > 0,\ x \in \mathbb{R}^n \ \text{ (when } n \ge 3)$$

the DCT implies that (3.32) is true also for $\lambda = 0$ when $n \ge 3$. Using (3.13) while letting $\lambda \to 0$ in (3.32) gives that (3.32) is true also for $\lambda = 0$ provided that

$$G_0(x) = -\tfrac{1}{2\pi} \ln |x| \quad \text{(when } n = 2).$$

When $n = 1$, it is easy to check that if we take $G_0(x) = -x$ for $x \ge 0$ and $G_0(x) = 0$ for $x < 0$, then (3.32) remains true with $\lambda = 0$.

In view of the above discussion we have that if $v \in C_0^\infty(\mathbb{R}^n)$ and $u = G_0 * v$, then $u \in C^\infty(\mathbb{R}^n)$ and $-\Delta u = v$.

Theorem 3.3.2 *If $f \in \mathcal{D}'(\mathbb{R}^n)$ and $v \in C_0^\infty(\mathbb{R}^n)$, then $f * v \in C^\infty(\mathbb{R}^n)$; moreover, $D^\alpha(f * v) = f * (D^\alpha v) = (\mathsf{D}^\alpha f) * v$ for every multi-index α.*

PROOF Suppose that $\operatorname{supp}(v) \subset B(0,r)$ and that $x, h \in B(0,r)$. Pick $C < \infty$ and an integer k such that

$$|f(\phi)| \leq C\|\phi\|_{k,\infty} \quad \text{for all } \phi \in C_0^\infty(\mathbb{R}^n) \text{ with } \operatorname{supp}(\phi) \subset B(0,3r).$$

Let $u(x) = f(v_x)$, see (3.31), and note that $\operatorname{supp}(v_{x+h}) \subset B(0,3r)$; hence

$$\left|u(x+h) - u(x) - \textstyle\sum_{j=1}^n h_j f((D_j v)_x)\right| = |f(\phi)| \leq C\|\phi\|_{k,\infty}$$

where $\phi = v_{x+h} - v_x - \sum_{j=1}^n h_j (D_j v)_x$. Since for every multi-index β,

$$(-1)^{|\beta|}(D^\beta \phi)(y) = (D^\beta v)(x+h-y) - (D^\beta v)(x-y) - \sum_{j=1}^n h_j (D_j D^\beta v)(x-y)$$

$$\|D^\beta \phi\|_\infty \leq n|h|^2 \max_{|\gamma|=2} \|D^{\gamma+\beta} v\|_\infty,$$

we have that

$$\left|u(x+h) - u(x) - \textstyle\sum_{j=1}^n h_j f((D_j v)_x)\right| \leq Cn|h|^2 \|v\|_{k+2,\infty}.$$

Therefore u is continuous, $(D_j u)(x) = f((D_j v)_x)$ for $j = 1, \ldots, n$ and we can apply the above argument with $D_j v$ in place of v. Repeating this argument gives that for every multi-index α we have

$$(D^\alpha u)(x) = f((D^\alpha v)_x) = (-1)^{|\alpha|} f(D^\alpha v_x) = (\mathrm{D}^\alpha f)(v_x).$$

□

Lemma 3.3.3 *Let $w_\tau = (e^{i\tau_1}, \cdots, e^{i\tau_n}) \in \mathbb{C}^n$ for $\tau \in T \equiv (-\pi, \pi)^n$. If P is a polynomial in $\mathbb{C}^n$, $P \not\equiv 0$, then there exists $A < \infty$ such that*

$$|f(z)| \leq A \int_T |(fP)(z + w_\tau)| d\tau$$

for every entire function f in $\mathbb{C}^n$ and for every $z \in \mathbb{C}^n$.

PROOF Suppose first that F is an entire function in $\mathbb{C}$ and that

$$Q(\lambda) = a_m \lambda^m + \cdots + a_0 = a_m(\lambda - z_1) \cdots (\lambda - z_m) \quad \text{for} \quad \lambda \in \mathbb{C}.$$

Let $g(\lambda) = F(\lambda) a_m (1 - \overline{z}_1 \lambda) \cdots (1 - \overline{z}_m \lambda)$ and note that $|g(\lambda)| = |(FQ)(\lambda)|$ when $|\lambda| = 1$. Hence, Cauchy's formula implies

$$|a_m F(0)| = |g(0)| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(e^{it})| dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} |(FQ)(e^{it})| dt. \tag{3.33}$$

P can be written in the form $P = P_0 + \cdots + P_m$, where $P_j(z) = \sum_{|\alpha|=j} c_\alpha z^\alpha$ and $P_m \not\equiv 0$. Since $\int_T |P_m(w_\tau)|^2 d\tau = (2\pi)^n \sum_{|\alpha|=m} |c_\alpha|^2 \neq 0$, define A by

$$\frac{1}{A} = \int_T |P_m(w_\tau)| d\tau. \tag{3.34}$$

For $z \in \mathbb{C}^n$, $\tau \in T$ and $\lambda \in \mathbb{C}$ define

$$F(\lambda) = f(z + w_\tau \lambda), \quad Q(\lambda) = P(z + w_\tau \lambda) = P_m(w_\tau)\lambda^m + \cdots + P(z)$$

and note that F is an entire function in $\mathbb{C}$ and (3.33) implies

$$|P_m(w_\tau) f(z)| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |(fP)(z + w_\tau e^{it})| dt;$$

hence, (3.34), the Fubini Theorem and the periodicity of the integrand imply

$$|f(z)| \leq \frac{A}{2\pi} \int_T \int_{-\pi}^{\pi} |(fP)(z + w_{\tau+t})| dt\, d\tau = A \int_T |(fP)(z + w_\tau)| d\tau.$$

□

Theorem 3.3.4 (Malgrange-Ehrenpreis) *If P is any polynomial in $\mathbb{C}^n$ such that $P \not\equiv 0$, then there exists a distribution G in $\mathbb{R}^n$ such that $P(\mathrm{D})G = \delta_0$.*

PROOF Take $\phi \in C_0^\infty(\mathbb{R}^n)$ and let $\psi = P(-D)\phi$. Lemma 3.2.9 implies $\hat{\psi}(x) = P(-ix)\hat{\phi}(x)$ for $x \in \mathbb{C}^n$; hence Lemma 3.3.3 implies

$$|\hat{\phi}(x)| \leq A \int_T |\hat{\psi}(x + w_\tau)| d\tau \quad \text{for} \quad x \in \mathbb{R}^n,$$

where $w_\tau = (e^{i\tau_1}, \cdots, e^{i\tau_n}) \in \mathbb{C}^n$ for $\tau \in T \equiv (-\pi, \pi)^n$ and $A < \infty$, depends on P only. Since $\phi(0) = (2\pi)^{-n/2} \int \hat{\phi}(x) dx$, we have that

$$|\phi(0)| \leq A(2\pi)^{-n/2} \|P(-D)\phi\|, \tag{3.35}$$

where $\|\cdot\|$ is defined by

$$\|\varphi\| = \int_{\mathbb{R}^n} \int_T |\hat{\varphi}(x + w_\tau)| d\tau\, dx \quad \text{for} \quad \varphi \in C_0^\infty(\mathbb{R}^n).$$

Fix an integer $m > n/2$, $r > 0$ and $\varphi \in C_0^\infty(\mathbb{R}^n)$ with $\mathrm{supp}(\varphi) \subset B(0, r)$. If $w \in \mathbb{C}^n$ and $|w| \geq 1$, then Lemmas 3.2.2, 3.2.5 and (1.7) imply

$$\begin{aligned} \int_{\mathbb{R}^n} (1 + |x|^2)^m |\hat{\varphi}(x + w)|^2 dx &\leq c_1 \sum_{|\alpha| \leq m} \int_{\mathbb{R}^n} |x^\alpha \hat{\varphi}(x + w)|^2 dx \\ &= c_1 \sum_{|\alpha| \leq m} \int_{\mathbb{R}^n} |D^\alpha (e^{-iw\cdot(\cdot)} \varphi)|^2 \\ &\leq (c_2 |w|^m e^{|w| r} \|\varphi\|_{m,\infty})^2; \end{aligned}$$

hence

$$\begin{aligned} \|\varphi\| &= \int_T d\tau \int_{\mathbb{R}^n} (1+|x|^2)^{m/2} |\hat{\varphi}(x+w_\tau)| (1+|x|^2)^{-m/2} dx \\ &\leq c_3 \int_T |w_\tau|^m e^{|w_\tau|r} \|\varphi\|_{m,\infty} d\tau \\ \|\varphi\| &\leq c_4 \|\varphi\|_{m,\infty}, \end{aligned} \tag{3.36}$$

where $c_1, c_2, c_3, c_4 \in (0,\infty)$ depend only on m, n and r. In particular, $\|\varphi\| < \infty$. If $\|\varphi\| = 0$, then $\varphi = 0$ by Lemma 3.3.3. Thus, $\|\cdot\|$ is a norm on $C_0^\infty(\mathbb{R}^n)$.

Let M be the subspace of $C_0^\infty(\mathbb{R}^n)$ that consists of functions $P(-D)\phi$ with $\phi \in C_0^\infty(\mathbb{R}^n)$. (3.35) implies that we can define a bounded linear functional g on M by $g(P(-D)\phi) = \phi(0)$. The Hahn-Banach Theorem 1.5.6 implies that there exists a linear functional G on $C_0^\infty(\mathbb{R}^n)$ such that for all $\phi \in C_0^\infty(\mathbb{R}^n)$ we have

$$G(P(-D)\phi) = \phi(0) \quad \text{and} \quad |G(\phi)| \leq A(2\pi)^{-n/2} \|\phi\|.$$

(3.36) implies that G is a distribution in $\mathbb{R}^n$ of order $\leq m$. □

3.4 Weak Derivatives

The following Theorem is a partial converse of Lemma 2.4.4 and will play a central role in our generalization of derivatives of functions.

Theorem 3.4.1 *Assume $p, q \in [1,\infty)$, $u \in L^p_{loc}(\Omega)$, $v \in L^q_{loc}(\Omega)$, α is a multi-index,*

$$(-1)^{|\alpha|} \int_\Omega u D^\alpha \varphi = \int_\Omega v\varphi \quad \textit{for all} \quad \varphi \in C_0^\infty(\Omega)$$

and that $K \subset \Omega$, K is compact. Pick $d \in (0,\infty)$ such that $K + \overline{B(0,d)} \subset \Omega$ and define

$$u_\varepsilon(x) = \int_{K+B(0,d)} J_\varepsilon(x-y)u(y)dy, \qquad v_\varepsilon(x) = \int_{K+B(0,d)} J_\varepsilon(x-y)v(y)dy,$$

for $x \in \mathbb{R}^n$ where J_ε is a mollifier. Then

(a) $u_\varepsilon, v_\varepsilon \in C_0^\infty(\mathbb{R}^n)$ for every $\varepsilon > 0$

(b) $(D^\alpha u_\varepsilon)(x) = v_\varepsilon(x)$ when $x \in K$ and $0 < \varepsilon < d$

(c) $\lim_{\varepsilon\to 0} \int_K |u_\varepsilon - u|^p = \lim_{\varepsilon\to 0} \int_K |D^\alpha u_\varepsilon - v|^q = 0$

(d) if $u = 0$, then $v = 0$ a.e. in Ω.

PROOF Theorem 3.1.1 implies (a) and, when $x \in K$, $0 < \varepsilon < d$, then

$$(D^\alpha u_\varepsilon)(x) = \int_\Omega (D^\alpha J_\varepsilon)(x-y)u(y)dy = \int_\Omega J_\varepsilon(x-y)v(y)dy = v_\varepsilon(x)$$

since $B(x,\varepsilon) \subset K + B(0,d)$. (c) follows from Theorem 3.1.7. If $u = 0$, then (c) implies that $v = 0$ a.e. on each compact $K \subset \Omega$ and, since Ω is equal to a countable union of compact sets, (d) follows. □

Definition 3.4.2 When $u \in L^1_{loc}(\Omega)$ and α is a multi-index we say that u has the αth **weak derivative** (or that $\mathsf{D}^\alpha u$ exists) if there exists $v \in L^1_{loc}(\Omega)$ such that

$$(-1)^{|\alpha|}\int_\Omega u D^\alpha \varphi = \int_\Omega v\varphi \quad \text{for all} \quad \varphi \in C_0^\infty(\Omega). \tag{3.37}$$

(d) of Theorem 3.4.1 implies that if u has an αth weak derivative, then there is only one $v \in L^1_{loc}(\Omega)$ for which equation (3.37) is true and, in this case, define $\mathsf{D}^\alpha u = v$. When $D^\alpha = D_i$, let $\mathsf{D}^\alpha = \mathsf{D}_i$.

Requirement (3.37) could be strengthened a bit, because Theorem 3.1.4 implies that if $u \in L^1_{loc}(\Omega)$ and $\mathsf{D}^\alpha u$ exists for some multi-index α, then

$$(-1)^{|\alpha|}\int_\Omega u D^\alpha \varphi = \int_\Omega \varphi \mathsf{D}^\alpha u \quad \text{for all} \quad \varphi \in C_0^{|\alpha|}(\Omega).$$

If $u \in L^1_{loc}(\Omega)$ and if α is any multi-index, then the αth derivative of the distribution f_u, see (3.30), exists and is given by the left hand side of (3.37). The above definition says that $\mathsf{D}^\alpha u$ exists provided that $\mathsf{D}^\alpha f_u = f_v$ for some $v \in L^1_{loc}(\Omega)$.

If $u \in C^m(\Omega)$ and α is a multi-index with $|\alpha| \le m$, then by Lemma 2.4.4, the weak derivative $\mathsf{D}^\alpha u$ exists and is equal to the classical derivative $D^\alpha u$ almost everywhere in Ω. In one dimension, a function that has the first weak derivative must be locally equal (a.e.) to an absolutely continuous function (Lemma 2.4.5). In higher dimensions existence of all first order weak derivatives may not imply local boundedness; for example, if $u(x) = |x|^{-\varepsilon}$ for $x \in \mathbb{R}^2$, $\varepsilon < 1$, then u has both first order weak derivatives. Thus, this definition of the weak derivative expands the classical notion of the derivative.

If $u \in L^1_{loc}(\Omega)$ and $\mathsf{D}^\alpha u$ exists, then Theorem 3.4.1 implies that for each compact $K \subset \Omega$ there exist $u_\varepsilon \in C_0^\infty(\mathbb{R}^n)$ such that $u_\varepsilon \to u$ and $D^\alpha u_\varepsilon \to \mathsf{D}^\alpha u$ in $L^1(K)$ as $\varepsilon \to 0$. The converse is also true:

Theorem 3.4.3 *Suppose $u, v \in L^1_{loc}(\Omega)$, α is a multi-index and for each compact $K \subset \Omega$ there exist $f_1, f_2, \ldots$ in $L^1_{loc}(\Omega)$ such that $\mathsf{D}^\alpha f_i$ exists for all $i \ge 1$ and*

$$\lim_{i\to\infty}\int_K |u - f_i| = \lim_{i\to\infty}\int_K |v - \mathsf{D}^\alpha f_i| = 0.$$

Then $\mathsf{D}^\alpha u$ exists and $\mathsf{D}^\alpha u = v$.

PROOF Choose $\varphi \in C_0^\infty(\Omega)$ and let $K = \text{supp}(\varphi)$; then

$$\int_K uD^\alpha\varphi \overset{i\to\infty}{\longleftarrow} \int_K f_i D^\alpha\varphi = (-1)^{|\alpha|}\int_K \varphi \mathsf{D}^\alpha f_i \overset{i\to\infty}{\longrightarrow} (-1)^{|\alpha|}\int_K \varphi v.$$

□

It is easy to see that D^α is a linear operator in $L^1_{loc}(\Omega)$. If $u \in L^1_{loc}(\Omega)$ and both $\mathsf{D}^\alpha u$ and $\mathsf{D}^{\alpha+\beta}u$ exist, then $\mathsf{D}^\beta(\mathsf{D}^\alpha u)$ exists and is equal to $\mathsf{D}^{\alpha+\beta}$u. Also, if u has the weak derivative $\mathsf{D}^\alpha u$ and if $\mathsf{D}^\alpha u$ has the weak derivative $\mathsf{D}^\beta(\mathsf{D}^\alpha u)$, then $\mathsf{D}^{\alpha+\beta}u$ exists and $\mathsf{D}^{\alpha+\beta}u = \mathsf{D}^\beta(\mathsf{D}^\alpha u)$.

Observe that weak derivatives are defined globally on nonempty open sets. Clearly, if $u \in L^1_{loc}(\Omega)$ has a weak derivative on Ω, then u has (the same) weak derivative on any nonempty open set $\Omega' \subset \Omega$. Theorem 3.4.4 below tells us that if $u \in L^1_{loc}(\Omega)$ has an αth weak derivative on an open neighborhood of each point in Ω, then u has the αth weak derivative on Ω; so weak derivatives are really local properties of functions.

Theorem 3.4.4 *If Γ is a collection of nonempty open sets whose union is Ω, and if $u \in L^1_{loc}(\Omega)$ is such that for some multi-index α the αth weak derivative of u exists on each member of Γ, then u has the αth weak derivative on Ω.*

PROOF Let $\psi_1, \psi_2, \ldots$ be a partition of unity as given by Theorem 3.1.3. Choose $V_i \in \Gamma$ so that $\text{supp}(\psi_i) \subset V_i$ and let $g_i \in L^1_{loc}(V_i)$ be the αth weak derivative of u on V_i. Define $g_i = 0$ on $\Omega \backslash V_i$ and let

$$g = \sum_{i=1}^{\infty} g_i\psi_i.$$

If K is any compact set such that $K \subset \Omega$, then all but finitely many ψ_i vanish on K; thus g is well defined on K and

$$\int_K |g| \leq \sum_{i=1}^{m} \int_{K\cap\text{supp}(\psi_i)} |g_i|\psi_i < \infty$$

for some $m < \infty$, and $g \in L^1_{loc}(\Omega)$. If $\varphi \in C_0^\infty(\Omega)$, let $K = \text{supp}(\varphi)$; then

$$\int_\Omega uD^\alpha\varphi = \int_K uD^\alpha \sum_{i=1}^{m} \varphi\psi_i = \sum_{i=1}^{m}\int_{V_i} uD^\alpha(\varphi\psi_i)$$

$$= \sum_{i=1}^{m}(-1)^{|\alpha|}\int_{V_i} g_i\varphi\psi_i = (-1)^{|\alpha|}\int_\Omega g\varphi$$

for some $m < \infty$. □

With the help of the Fourier transform, weak derivatives can be characterized as follows:

Theorem 3.4.5 *Suppose $f \in L^2(\mathbb{R}^n)$, α is a multi-index and $g(x) = i^{|\alpha|}x^\alpha(\mathcal{F}f)(x)$ for $x \in \mathbb{R}^n$. Then $g \in L^2(\mathbb{R}^n)$ if and only if $\mathsf{D}^\alpha f$ exists and belongs to $L^2(\mathbb{R}^n)$. If $g \in L^2(\mathbb{R}^n)$, then $\mathcal{F}\mathsf{D}^\alpha f = g$.*

PROOF Assume first that $\mathsf{D}^\alpha f$ exists and that it belongs to $L^2(\mathbb{R}^n)$. Let $f_\varepsilon = J_\varepsilon * f$ for $\varepsilon > 0$. Since $D^\alpha f_\varepsilon = (D^\alpha J_\varepsilon) * f = J_\varepsilon * \mathsf{D}^\alpha f$, Theorems 3.1.7 and 3.2.6 imply that $\|\mathcal{F}f_\varepsilon - \mathcal{F}f\|_2 \to 0$ and $\|\mathcal{F}D^\alpha f_\varepsilon - \mathcal{F}\mathsf{D}^\alpha f\|_2 \to 0$ as $\varepsilon \to 0$. Corollary 3.2.7 and Lemma 3.2.2 imply

$$\mathcal{F}D^\alpha f_\varepsilon = (2\pi)^{n/2}(\mathcal{F}D^\alpha J_\varepsilon)\mathcal{F}f = (2\pi)^{n/2}(ix)^\alpha(\mathcal{F}J_\varepsilon)\mathcal{F}f = (ix)^\alpha(\mathcal{F}f_\varepsilon)(x),$$

and taking the pointwise limits gives that $g = \mathcal{F}\mathsf{D}^\alpha f \in L^2(\mathbb{R}^n)$.

If $g \in L^2(\mathbb{R}^n)$ and $\varphi \in C_0^\infty(\mathbb{R}^n)$, then

$$(f, D^\alpha\varphi) = (\mathcal{F}f, \mathcal{F}D^\alpha\varphi) = (-1)^{|\alpha|}(g, \mathcal{F}\varphi) = (-1)^{|\alpha|}(\mathcal{F}^{-1}g, \varphi),$$

implying that $\mathsf{D}^\alpha f = \mathcal{F}^{-1}g$. □

Let e_i be the unit vector in $\mathbb{R}^n$ such that $(e_i)_i = 1$ and $(e_i)_j = 0$ for $j \neq i$. Define

$$(\delta_i^h u)(x) = \frac{u(x + he_i) - u(x)}{h}. \tag{3.38}$$

The relation between the finite difference operator δ_i^h and D_i is highlighted next. These results will be used later to prove the regularity of weak solutions.

Theorem 3.4.6 *If $p \in [1, \infty)$, $u \in L^p_{loc}(\Omega)$, $\mathsf{D}_i u$ exists and $\mathsf{D}_i u \in L^p_{loc}(\Omega)$ for some $1 \leq i \leq n$ and C is a compact set such that $K \equiv C + \overline{B(0,H)} \subset \Omega$ for some $H > 0$, then*

$$\int_C |\delta_i^h u|^p \leq \int_K |\mathsf{D}_i u|^p \quad \text{for} \quad 0 < |h| < H \tag{3.39}$$

$$\lim_{h\to 0} \int_C |\delta_i^h u - \mathsf{D}_i u|^p = 0. \tag{3.40}$$

PROOF Let u_ε be as in Theorem 3.4.1. If $x \in C$ and $0 < |h| < H$, then

$$\begin{aligned}
(\delta_i^h u_\varepsilon)(x) &= \int_0^1 (D_i u_\varepsilon)(x + the_i)dt \\
|(\delta_i^h u_\varepsilon)(x)|^p &\leq \int_0^1 |(D_i u_\varepsilon)(x + the_i)|^p dt \\
\int_C |\delta_i^h u_\varepsilon|^p &\leq \int_0^1 dt \int_{C+the_i} |D_i u_\varepsilon|^p \leq \int_K |D_i u_\varepsilon|^p,
\end{aligned}$$

which implies (3.39) since $u_\varepsilon, D_i u_\varepsilon$ converge in $L^p(K)$ to $u, \mathsf{D}_i u$. (3.39) implies

$$\begin{aligned} 3^{1-p}|\delta_i^h u - \mathsf{D}_i u|^p &\le |\delta_i^h(u-u_\varepsilon)|^p + |\delta_i^h u_\varepsilon - D_i u_\varepsilon|^p + |D_i u_\varepsilon - \mathsf{D}_i u|^p \\ 3^{1-p}\int_C |\delta_i^h u - \mathsf{D}_i u|^p &\le \int_C |\delta_i^h u_\varepsilon - D_i u_\varepsilon|^p + 2\int_K |D_i u_\varepsilon - \mathsf{D}_i u|^p; \end{aligned}$$

choose first ε so that $\int_K |D_i u_\varepsilon - \mathsf{D}_i u|^p$ is small enough and then note

$$\begin{aligned} (\delta_i^h u_\varepsilon - D_i u_\varepsilon)(x) &= \int_0^1 ((D_i u_\varepsilon)(x + the_i) - (D_i u_\varepsilon)(x))dt \\ \int_C |\delta_i^h u_\varepsilon - D_i u_\varepsilon|^p &\le \int_0^1 \int_C |(D_i u_\varepsilon)(x+the_i) - (D_i u_\varepsilon)(x)|^p dx\, dt \overset{h\to 0}{\longrightarrow} 0 \end{aligned}$$

because of the uniform continuity of $D_i u_\varepsilon$ on K, which proves (3.40). □

Theorem 3.4.7 *Suppose that $1 < p \le \infty$, $u \in L^p_{loc}(\Omega)$, $1 \le i \le n$ and for each compact subset C of Ω there exist $h_1, h_2, \ldots$ in $\mathbb{R}\backslash\{0\}$ such that*

$$\lim_{k\to\infty} h_k = 0 \quad \textit{and} \quad \sup_k \int_C |\delta_i^{h_k} u|^p < \infty.$$

Then, $\mathsf{D}_i u$ exists and $\mathsf{D}_i u \in L^p_{loc}(\Omega)$.

PROOF Pick any nonempty compact $K \subset \Omega$ and $d > 0$ such that $\overline{C} \subset \Omega$ where $C = K + B(0,d)$. Corollary 1.5.4 implies that there exist $v \in L^p(C)$ and $h_k \in \mathbb{R}\backslash\{0\}$ with $h_k \to 0$ such that $\int_C \phi \delta_i^{h_k} u \to \int_C v\phi$ for all $\phi \in C_0^\infty(C)$. Since $\int_C \phi \delta_i^h u = -\int_C u \delta_i^{-h}\phi$ when h is small enough and $\delta_i^{-h_k}\phi \to D_i\phi$, we have that $-\int_C u D_i\phi = \int_C v\phi$ for all $\phi \in C_0^\infty(C)$ and hence $\mathsf{D}_i u$ exists in C. Theorem 3.4.4 implies that $\mathsf{D}_i u$ exists in Ω. □

Theorem 3.4.8 *Suppose Ω is connected, $u \in L^1_{loc}(\Omega)$, $\mathsf{D}^\alpha u$ exists and $\mathsf{D}^\alpha u = 0$ whenever $|\alpha| = 1$. Then there exists $c \in \mathbb{C}$ such that $u = c$ a.e. in Ω.*

PROOF Choose $x \in \Omega$ and $r \in (0,\infty)$ so that $K = \overline{B(x,r)} \subset \Omega$. Let u_ε be as in Theorem 3.4.1 and note that (b) of the Theorem implies $\nabla u_\varepsilon = 0$ in K. If $y \in K$, then

$$u_\varepsilon(y) - u_\varepsilon(x) = \int_0^1 (\nabla u_\varepsilon)(x + t(y-x)) \cdot (y-x)dt = 0;$$

hence there exists $c_\varepsilon \in \mathbb{C}$ such that $u_\varepsilon = c_\varepsilon$ in K. (c) of Theorem 3.4.1 implies that c_ε converge to some c as $\varepsilon \to 0$ and that $u(y) = c$ for almost all $y \in K$.

Let $V_1, V_2, \ldots$ be nonempty open balls such that $\Omega = \cup\overline{V_i} = \cup V_i$. By the above argument there exist c_i such that $u = c_i$ a.e. in V_i. Let A be the union of all V_i such that $c_i = c_1$ and let B be the union of other V_i. It can be easily seen that $A \cap B$ is empty and, since Ω is connected, B has to be empty. □

Theorem 3.4.9 *Suppose $f \in L^1_{loc}(\Omega)$, α is a multi-index with $|\alpha| = 1$ and $\mathsf{D}^\alpha f$ exists. Then $\mathsf{D}^\alpha|f|$ exists and $\mathsf{D}^\alpha|f| = \mathrm{Re}\,(g\mathsf{D}^\alpha f)$ where g is defined by*

$$g(x) = \begin{cases} \overline{f(x)}/|f(x)| & \text{if } x \in \Omega \text{ and } 0 < |f(x)| < \infty \\ 0 & \text{for other } x \in \Omega \end{cases}$$

PROOF For $\varepsilon > 0$ define $u_\varepsilon = (|f|^2 + \varepsilon)^{1/2}$ and $g_\varepsilon = \overline{f}/u_\varepsilon$. Choose a compact $K \subset \Omega$. By Theorem 3.4.1 there exist $f_1, f_2, \ldots$ in $C_0^\infty(\mathbb{R}^n)$ such that

$$\lim_{i\to\infty} \int_K |f - f_i| = \lim_{i\to\infty} \int_K |\mathsf{D}^\alpha f - D^\alpha f_i| = 0,$$

$$\lim_{i\to\infty} f_i(x) = f(x) \quad \text{for almost all } x \in K.$$

Define $u_{i,\varepsilon} = (|f_i|^2 + \varepsilon)^{1/2}$, $g_{i,\varepsilon} = \overline{f_i}/u_{i,\varepsilon}$. Clearly $D^\alpha u_{i,\varepsilon} = \mathrm{Re}\,(g_{i,\varepsilon} D^\alpha f_i)$,

$$\int_K |u_{i,\varepsilon} - u_\varepsilon| \le \int_K |f_i - f| \to 0 \quad \text{as } i \to \infty$$

$$\int_K |D^\alpha u_{i,\varepsilon} - \mathrm{Re}\,(g_\varepsilon \mathsf{D}^\alpha f)| \le \int_K |D^\alpha f_i - \mathsf{D}^\alpha f| + \int_K |g_{i,\varepsilon} - g_\varepsilon||\mathsf{D}^\alpha f|,$$

which also converges to 0 (by the DCT). By Theorem 3.4.3 we have that $\mathsf{D}^\alpha u_\varepsilon = \mathrm{Re}\,(g_\varepsilon \mathsf{D}^\alpha f)$. Since for every compact $K \subset \Omega$ we have that

$$\int_K |u_\varepsilon - |f|| \to 0 \quad \text{as } \varepsilon \to 0$$

$$\int_K |\mathsf{D}^\alpha u_\varepsilon - \mathrm{Re}\,(g\mathsf{D}^\alpha f)| \to 0 \quad \text{as } \varepsilon \to 0,$$

another application of Theorem 3.4.3 completes the proof. □

The following formula for a weak derivative of the product generalizes (1.7).

Theorem 3.4.10 *Suppose $u \in L^1_{loc}(\Omega)$, $g \in C^{|\alpha|}(\Omega)$ for some multi-index α, and $\mathsf{D}^\beta u$ exist for all multi-indices $\beta \le \alpha$. Then $\mathsf{D}^\alpha(ug)$ exists and*

$$\mathsf{D}^\alpha(ug) = \sum_{\beta\le\alpha} \binom{\alpha}{\beta} (\mathsf{D}^\beta u) D^{\alpha-\beta} g.$$

PROOF Choose any compact $K \subset \Omega$ and let $u_\varepsilon \in C_0^\infty(\mathbb{R}^n)$ be as in Theorem 3.4.1. Observe that

$$\lim_{\varepsilon \to 0} \int_K |D^\beta u_\varepsilon - \mathrm{D}^\beta u| = 0 \quad \text{for all} \quad \beta \leq \alpha.$$

Therefore, using (1.7),

$$\lim_{\varepsilon \to 0} \int_K \left| D^\alpha(u_\varepsilon g) - \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} (\mathrm{D}^\beta u) D^{\alpha-\beta} g \right| = 0$$

and Theorem 3.4.3 implies the assertion. □

In the rest of this section the interpolation inequalities for intermediate derivatives will be studied.

Lemma 3.4.11 *If* $1 \leq p \leq \infty$ *and* $-\infty \leq a < b \leq \infty$, *then*

$$\|f'\|_p \leq \tfrac{\varepsilon}{2} \|f''\|_p + \tfrac{6}{\varepsilon} \|f\|_p \quad \textit{for} \quad f \in C^2(a,b),\ 0 < \varepsilon < b - a.$$

PROOF Let $\delta = \varepsilon/2$ and pick c so that $c \pm \delta \in (a,b)$. If $x, x+h \in (a,b)$, then

$$hf'(x) = f(x+h) - f(x) - \int_0^h (h-s) f''(x+s) ds$$

hence, choosing first $h = \delta$ and then $h = -\delta$, we obtain that

$$\delta f'(x) = f_1(x) + f_2(x) + f_3(x),$$

where

$$f_1(x) = \begin{cases} f(x+\delta) & \text{if } a < x \leq c \\ -f(x-\delta) & \text{if } c < x < b \end{cases}$$

$$f_2(x) = \begin{cases} -f(x) & \text{if } a < x \leq c \\ f(x) & \text{if } c < x < b \end{cases}$$

$$f_3(x) = \begin{cases} \int_0^\delta (s-\delta) f''(x+s) ds & \text{if } a < x \leq c \\ \int_0^\delta (\delta - s) f''(x-s) ds & \text{if } c < x < b. \end{cases}$$

It is easy to see that $\|f_1\|_p \leq 2^{1/p} \|f\|_p$ and $\|f_2\|_p = \|f\|_p$. If $p < \infty$, Hölder's inequality implies

$$|f_3(x)|^p \leq (\delta^2/2)^{p-1} \int_0^\delta (\delta - s) |f''(x+s)|^p ds \quad \text{if} \quad a < x \leq c$$

$$\int_a^c |f_3(x)|^p dx \leq (\delta^2/2)^{p-1} \int_0^\delta (\delta - s) \|f''\|_p^p ds = (\delta^2/2)^p \|f''\|_p^p.$$

This and a similar calculation on (c, b) gives $\|f_3\|_p \leq \|f''\|_p 2^{1/p}\delta^2/2$ - which can be easily verified also when $p = \infty$. Thus

$$\delta\|f'\|_p \leq (2^{1/p} + 1)\|f\|_p + \delta^2 2^{1/p-1}\|f''\|_p,$$

which proves the assertion. □

Lemma 3.4.12 *For every integer $m \geq 0$ there exists $C_m \in (0, \infty)$ such that*

$$\varepsilon^j \|f^{(j)}\|_p \leq \varepsilon^m \|f^{(m)}\|_p + C_m\|f\|_p$$

whenever $p \in [1, \infty]$, $-\infty \leq a < b \leq \infty$, $f \in C^m(a, b)$, $0 < \varepsilon < b - a$ and $0 \leq j \leq m$.

PROOF Let A_m denote the assertion of the Lemma. A_0 and A_1 are obviously true; A_2 follows from Lemma 3.4.11. So assume that $m \geq 3$ and that A_{m-1} is true. Note that A_{m-1} implies

$$(\varepsilon/2)^{m-1}\|f^{(m-1)}\|_p \leq (\varepsilon/2)^m\|f^{(m)}\|_p + C_{m-1}(\varepsilon/2)\|f'\|_p$$

$$\varepsilon\delta\|f'\|_p \leq (\varepsilon\delta)^{m-1}\|f^{(m-1)}\|_p + C_{m-1}\|f\|_p$$

for $\delta \in (0, 1)$. Hence, choosing δ so that $(2\delta)^{m-2}C_{m-1} < 1/2$ implies

$$\varepsilon^{m-1}\|f^{(m-1)}\|_p \leq \varepsilon^m\|f^{(m)}\|_p + C'\|f\|_p$$

for some constant C'. This, together with A_{m-1}, imply A_m. □

Definition 3.4.13 For integers $m \geq 0$, $p \in [1, \infty]$ and those $u \in L^1_{loc}(\Omega)$ for which $\mathrm{D}^\alpha u$ exists whenever $|\alpha| = m$, define

$$|u|_{m,p} = \max_{|\alpha|=m} \|\mathrm{D}^\alpha u\|_p \in [0, \infty].$$

$|\cdot|_{m,p}$ is a seminorm when restricted to those $u \in L^1_{loc}(\Omega)$ for which $\mathrm{D}^\alpha u \in L^p(\Omega)$ whenever $|\alpha| = m$.

Lemma 3.4.14 *Suppose $\Omega = (a_1, b_1) \times \cdots \times (a_n, b_n)$ for some $-\infty \leq a_i < b_i \leq \infty$ and $1 \leq p \leq \infty$. Then, for every integer $m \geq 0$ there exists $c \in (0, \infty)$, depending on m and n only, such that*

$$\varepsilon^j|f|_{j,p} \leq c\varepsilon^m|f|_{m,p} + c\|f\|_p \quad \text{for} \quad f \in C^m(\Omega),\ 0 < \varepsilon < \min_i b_i - a_i,\ 0 \leq j \leq m.$$

PROOF Note first that Lemma 3.4.12 implies that

$$\varepsilon^l \|D_i^l g\|_p \le \varepsilon^k \|D_i^k g\|_p + C_k \|g\|_p \quad \text{for} \quad g \in C^k(\Omega),\ 0 \le l \le k,\ 1 \le i \le n.$$

Choose a multi-index $\alpha = (\alpha_1, \dots, \alpha_n)$ with $|\alpha| = j$ and define $\beta_k = \alpha_k + \cdots + \alpha_n$, $f_k = D_k^{\alpha_k} \cdots D_n^{\alpha_n} f$. Note that for $1 \le k \le n$,

$$\begin{aligned}
\varepsilon^{\alpha_k} \|f_k\|_p &= \varepsilon^{\alpha_k} \|D_k^{\alpha_k} f_{k+1}\|_p \\
&\le \varepsilon^{m-\beta_{k+1}} \|D_k^{m-\beta_{k+1}} f_{k+1}\|_p + C_{m-\beta_{k+1}} \|f_{k+1}\|_p \\
&\le \varepsilon^{m-\beta_{k+1}} |f|_{m,p} + C_{m-\beta_{k+1}} \|f_{k+1}\|_p \\
\varepsilon^{\beta_k} \|f_k\|_p &\le \varepsilon^m |f|_{m,p} + C_{m-\beta_{k+1}} \varepsilon^{\beta_{k+1}} \|f_{k+1}\|_p,
\end{aligned}$$

where $\beta_{n+1} = 0$ and $f_{n+1} = f$. Hence, if $C = \max_{k \le m} C_k$, then

$$\varepsilon^j \|D^\alpha f\|_p = \varepsilon^{\beta_1} \|f_1\|_p \le (1 + C + \cdots + C^n) \varepsilon^m |f|_{m,p} + C^{n+1} \|f\|_p,$$

which implies the conclusion of the Lemma. □

Corollary 3.4.15 *For every integer $m \ge 1$ there exists $c \in (0, \infty)$, depending on m and n only, such that*

$$|f|_{j,p} \le c |f|_{m,p}^{j/m} \|f\|_p^{1-j/m} \quad \text{for} \quad f \in C_0^m(\Omega),\ 0 \le j \le m,\ 1 \le p \le \infty.$$

PROOF If $f \in C_0^m(\Omega)$ and $f = 0$ in $\mathbb{R}^n \backslash \Omega$, then $f \in C_0^m(\mathbb{R}^n)$; hence the conclusion follows from Lemma 3.4.14 and (3.15). □

Theorem 3.4.16 *Suppose $\Omega = (a_1, b_1) \times \cdots \times (a_n, b_n)$ for some $-\infty \le a_i < b_i \le \infty$. Suppose also that $1 \le p < \infty$, $f \in L^p(\Omega)$ and that there exists $\mathrm{D}^\alpha f \in L^p(\Omega)$ whenever $|\alpha| = m$. Then there exists $\mathrm{D}^\beta f \in L^p(\Omega)$ for all $|\beta| \le m$. Moreover, for some $c \in (0, \infty)$, depending on m and n only, we have that*

$$\varepsilon^j |f|_{j,p} \le c\varepsilon^m |f|_{m,p} + c\|f\|_p \quad \text{for} \quad 0 < \varepsilon < \min_i b_i - a_i,\ 0 \le j \le m.$$

PROOF Let $\Omega' = (c_1, d_1) \times \cdots \times (c_n, d_n)$ for some $a_i < c_i < d_i < b_i$. Theorem 3.4.1 implies the existence of $f_1, f_2, \dots$ in $C_0^\infty(\mathbb{R}^n)$ such that

$$\int_{\Omega'} |f_k - f|^p \to 0, \quad \int_{\Omega'} |D^\alpha f_k - \mathrm{D}^\alpha f|^p \to 0 \quad \text{when } |\alpha| = m \text{ as } k \to \infty.$$

Lemma 3.4.14 implies that $\{D^\beta f_k\}$ is a Cauchy sequence in $L^p(\Omega')$ for each multi-index β with $|\beta| \le m$. Thus there exists $\mathrm{D}^\beta f \in L^p(\Omega')$ for all $|\beta| \le m$ and, by Lemma 3.4.14, we have that

$$\varepsilon^j |f|_{j,p,\Omega'} \le c\varepsilon^m |f|_{m,p,\Omega'} + c\|f\|_{p,\Omega'} \quad \text{for} \quad 0 < \varepsilon < \min_i d_i - c_i,\ 0 \le j \le m.$$

Theorem 3.4.4 implies that $\mathsf{D}^\beta f$ exists in Ω for all $|\beta| \leq m$. Letting $\Omega' \to \Omega$ gives the bound. □

The above interpolation inequalities are actually true under much weaker assumptions on Ω. See Adams [1] for more details.

An obvious consequence of Theorems 3.4.4 and 3.4.16 is:

Corollary 3.4.17 *Suppose that $f \in L^1_{loc}(\Omega)$ and that $\mathsf{D}^\alpha f$ exists whenever $|\alpha| = m$. Then $\mathsf{D}^\beta f$ exists whenever $|\beta| \leq m$.*

3.5 Definition and Basic Properties of Sobolev Spaces

Definition 3.5.1 of $W^{m,p}_{loc}(\Omega)$, $W^m_{loc}(\Omega)$, $W^{m,p}(\Omega)$, $W^m(\Omega)$, $(\cdot,\cdot)_m$, $\|\cdot\|_{m,p}$.

For nonnegative integers m and for $p \in [1,\infty]$, define $W^{m,p}_{loc}(\Omega)$ to be the set of all $u \in L^p_{loc}(\Omega)$ which are such that $\mathsf{D}^\alpha u$ exists and $\mathsf{D}^\alpha u \in L^p_{loc}(\Omega)$ for all multi-indices α with $|\alpha| \leq m$. $W^m_{loc}(\Omega) = W^{m,2}_{loc}(\Omega)$.

For nonnegative integers m and for $p \in [1,\infty]$ define $W^{m,p}(\Omega)$ to be the set of all $u \in L^p(\Omega)$ which are such that $\mathsf{D}^\alpha u$ exists and $\mathsf{D}^\alpha u \in L^p(\Omega)$ for all multi-indices α with $|\alpha| \leq m$. $W^m(\Omega) = W^{m,2}(\Omega)$.

For $u, v \in W^m(\Omega)$, define the inner product $(u,v)_m$ by

$$(u,v)_m = \sum_{|\alpha| \leq m} \int_\Omega (\mathsf{D}^\alpha u)\overline{\mathsf{D}^\alpha v}.$$

For $u \in W^{m,p}(\Omega)$, define the norm $\|u\|_{m,p}$ by

$$\|u\|_{m,p} = \left(\sum_{|\alpha| \leq m} \|\mathsf{D}^\alpha u\|_p^p\right)^{1/p} \quad \text{if} \quad 1 \leq p < \infty$$

$$\|u\|_{m,\infty} = \max_{|\alpha| \leq m} \|\mathsf{D}^\alpha u\|_\infty \quad \text{if} \quad p = \infty.$$

Theorem 3.5.2 *$W^{m,p}(\Omega)$ is a Banach space. $W^m(\Omega)$ is a Hilbert space.*

PROOF Clearly, $W^{m,p}(\Omega)$ is a normed linear space. If $u_1, u_2, \ldots$ is a Cauchy sequence in $W^{m,p}(\Omega)$, then $\{\mathsf{D}^\alpha u_i\}$ is a Cauchy sequence in $L^p(\Omega)$ for each multi-index α with $\alpha \leq m$. Let u denote the $L^p(\Omega)$ limit of u_i, and let u_α denote the $L^p(\Omega)$ limits of $\mathsf{D}^\alpha u_i$ for $|\alpha| > 0$. $\mathsf{D}^\alpha u = u_\alpha$ by Theorem 3.4.3. □

Theorem 3.5.3 *If $u \in W^{m,p}(\Omega)$ and $v \in C^m_B(\Omega)$, then $uv \in W^{m,p}(\Omega)$ and*

$$\|uv\|_{m,p} \leq 2^m(1+m)^{n/p}\|u\|_{m,p}\|v\|_{m,\infty}.$$

PROOF See Theorem 3.4.10. □

Lemma 3.5.4 *If $u \in W^{m,p}(\Omega)$, $1 \leq p < \infty$, $m \geq 0$ and J_ε is a mollifier, then*

$$\lim_{\varepsilon \to 0} \int_K |D^\alpha (J_\varepsilon * u) - \mathrm{D}^\alpha u|^p = 0$$

for all compact $K \subset \Omega$ and all multi-indices α for which $|\alpha| \leq m$.

PROOF For all sufficiently small ε and for all $|\alpha| \leq m$, we have that

$$D^\alpha (J_\varepsilon * u) = J_\varepsilon * (\mathrm{D}^\alpha u) \quad \text{in } K.$$

Theorem 3.1.7 implies convergence. □

The following result shows that $W^{m,p}(\Omega)$, for $p < \infty$, is just the completion of the set of $C^\infty(\Omega)$ functions with finite $\| \cdot \|_{m,p}$ norm. This fact is used by some to define $W^{m,p}(\Omega)$ and, in this case, the notation $H^{m,p}(\Omega)$ is often used instead of $W^{m,p}(\Omega)$.

Theorem 3.5.5 *If $p \in [1, \infty)$ and $m \geq 0$, then $C^\infty(\Omega) \cap W^{m,p}(\Omega)$ is dense in $W^{m,p}(\Omega)$.*

PROOF Choose $u \in W^{m,p}(\Omega)$ and $\varepsilon > 0$. Let $\psi_1, \psi_2, \ldots$ be a partition of unity as in Theorem 3.1.3 (take, for example, $\Gamma = \{\Omega\}$).

Choose $s_i \in (0, 1/i)$ so that $K_i \equiv \mathrm{supp}(\psi_i) + \overline{B(0, s_i)} \subset \Omega$. If $t \in (0, s_i)$, then $J_t * (u\psi_i) \in C_0^\infty(\Omega)$, $\mathrm{supp}(J_t * (u\psi_i)) \subset K_i$. By Lemma 3.5.4 (see also Theorem 3.4.10),

$$\| J_{t_i} * (u\psi_i) - u\psi_i \|_{m,p} < \varepsilon 2^{-i} \tag{3.41}$$

for some $t_i \in (0, 1/i)$, $t_i < s_i$.

Pick any $x \in \Omega$ and let $r > 0$ be such that $\overline{B(x, 2r)} \subset \Omega$. There exists an integer $m > 1/r$ such that $\psi_i = 0$ on $\overline{B(x, 2r)}$ for all $i > m$. It is easy to see that $J_{t_i} * (u\psi_i) = 0$ on $B(x, r)$ for $i >$ m. Therefore, if

$$\phi = \sum_{i=1}^{\infty} J_{t_i} * (u\psi_i),$$

then in a neighborhood of each point in Ω, all but finitely many terms of the sum vanish identically. Thus $\phi \in C^\infty(\Omega)$.

The following identity is obviously true pointwise a.e.

$$u - \phi = \sum_{i=1}^{\infty} u\psi_i - J_{t_i} * (u\psi_i).$$

(3.41) implies that $u - \phi \in W^{m,p}(\Omega)$ and $\| u - \phi \|_{m,p} < \varepsilon$. □

$W^m(\mathbb{R}^n)$ can be nicely characterized with the help of the Fourier transform:

Theorem 3.5.6 *Suppose* $f \in L^2(\mathbb{R}^n)$, $m \geq 0$ *and* $g(x) = (1+|x|^2)^{m/2}(\mathcal{F}f)(x)$ *for* $x \in \mathbb{R}^n$. *Then* $f \in W^m(\mathbb{R}^n)$ *if and only if* $g \in L^2(\mathbb{R}^n)$. *Moreover, there exists* $c \in (0,\infty)$, *depending only on* m *and* n, *such that*

$$\|f\|_{m,2} \leq \|g\|_2 \leq c\|f\|_{m,2} \quad \textit{when } f \in W^m(\mathbb{R}^n).$$

PROOF Note that for some $c \in (0,\infty)$

$$\sum_{|\alpha|\leq m}(x^\alpha)^2 \leq (1+|x|^2)^m = \sum_{k=0}^m \binom{m}{k} \sum_{|\alpha|=k} \frac{k!}{\alpha!}(x^\alpha)^2 \leq c^2 \sum_{|\alpha|\leq m}(x^\alpha)^2$$

$$\sum_{|\alpha|\leq m}|x^\alpha(\mathcal{F}f)(x)|^2 \leq |g(x)|^2 \leq c^2 \sum_{|\alpha|\leq m}|x^\alpha(\mathcal{F}f)(x)|^2.$$

This and Theorems 3.4.5 and 3.2.6 imply the conclusion. □

Definition 3.5.7 of $W_0^{m,p}(\Omega), W_0^m(\Omega)$.

$C_0^\infty(\Omega)$ is clearly a subspace of each of the spaces $W^{m,p}(\Omega)$.

For nonnegative integers m and for $p \in [1,\infty]$, define $W_0^{m,p}(\Omega)$ to be the closure of $C_0^\infty(\Omega)$ in $W^{m,p}(\Omega)$. $W_0^m(\Omega) = W_0^{m,2}(\Omega)$.

Clearly, $W_0^{m,p}(\Omega)$ is a Banach space with norm $\|\cdot\|_{m,p}$ and $W_0^m(\Omega)$ is a Hilbert space with inner product $(\cdot,\cdot)_m$.

Theorem 3.1.4 implies that $C_0^m(\Omega) \subset W_0^{m,p}(\Omega)$; thus,

the closure of $C_0^m(\Omega)$ in $W^{m,p}(\Omega)$ is also equal to $W_0^{m,p}(\Omega)$.

This and Theorem 3.5.3 imply that

if $u \in W_0^{m,p}(\Omega)$ and $v \in C_B^m(\Omega)$, then $uv \in W_0^{m,p}(\Omega)$.

Also, this and Theorem 3.5.5 imply that

if $p < \infty$, $u \in W^{m,p}(\Omega)$, $v \in C_0^m(\Omega)$, then $uv \in W_0^{m,p}(\Omega)$.

Except for some special sets Ω, we have that $W^{m,p}(\Omega) \neq W_0^{m,p}(\Omega)$. One of these special Ω is $\mathbb{R}^n$.

Theorem 3.5.8 $W^{m,p}(\mathbb{R}^n) = W_0^{m,p}(\mathbb{R}^n)$ *if* $p \in [1,\infty)$, $m \geq 0$.

PROOF Choose $u \in W^{m,p}(\mathbb{R}^n)$ and let $f \in C_0^\infty(\mathbb{R}^n)$ be such that $f(x) = 1$ for $|x| \leq 1$ and $f(x) = 0$ for $|x| \geq 2$ (Theorem 3.1.2). Define $u_i(x) = u(x)f(x/i)$ and note that u_i are in $W_0^{m,p}(\mathbb{R}^n)$. Theorem 3.4.10 implies that there exists c such that for all $i \geq 1$,

$$\|u_i - u\|_{m,p}^p \leq c \sum_{|\alpha| \leq m} \int_{|x| \geq i} |\mathrm{D}^\alpha u|^p,$$

and therefore $u \in W_0^{m,p}(\mathbb{R}^n)$. □

When Ω lies between two parallel planes, with a normal ν and at a distance d apart, then (3.42) holds. In other words:

Theorem 3.5.9 *Suppose Ω is such that for some $\nu \in \mathbb{R}^n$, with $|\nu| = 1$ and some $d \in (0, \infty)$, we have that $(x - y) \cdot \nu \leq d$ for all x, y in Ω. Then*

$$\|f\|_p \leq d \, \|\nu \cdot \nabla f\|_p \quad \textit{for all} \quad f \in W_0^{1,p}(\Omega), \; 1 \leq p \leq \infty. \tag{3.42}$$

PROOF Choose $f \in C_0^\infty(\Omega)$ and note that for all $x \in \Omega$,

$$f(x) = -\int_0^d \nu \cdot \nabla f(x + \nu t) dt.$$

Hence, if $p < \infty$ and $1/p + 1/q = 1$, then

$$|f(x)|^p \leq d^{p/q} \int_0^d |\nu \cdot \nabla f(x + \nu t)|^p dt$$

$$\int_\Omega |f|^p = \int_{\mathbb{R}^n} |f|^p \leq d^{p/q} \int_0^d \int_{\mathbb{R}^n} |\nu \cdot \nabla f(x + \nu t)|^p dx \, dt = d^{1+p/q} \int_\Omega |\nu \cdot \nabla f|^p,$$

therefore $\|f\|_p \leq d \, \|\nu \cdot \nabla f\|_p$ - which is obviously true also for $p = \infty$. Since $\|\nu \cdot \nabla f\|_p \leq |\nu|_q \|f\|_{1,p}$, the inequality holds also in $W_0^{1,p}(\Omega)$. □

(3.42) is called the **Ponicare inequality** by some (see also Lemma 2.5.1 and the Dirichlet problem in Section 3.7). Others call the inequality (3.43) below the Ponicare inequality. Ω, that has properties (i) and (ii) in Theorem 3.5.10 below, is said to be **star-shaped** with respect to a ball contained in Ω. See also Exercise 14.

Theorem 3.5.10 *Suppose that Ω is such that*

(i) $B(x_0, r_1) \subset \Omega \subset B(x_0, r_2)$ for some $x_0 \in \mathbb{R}^n$, $0 < r_1 \leq r_2 < \infty$

(ii) if $y \in B(x_0, r_1)$ and $x \in \Omega$, then $y + t(x - y) \in \Omega$ for all $t \in [0, 1]$

and $f \in C^1(\Omega)$, $p \in [1,\infty]$, $\partial f/\partial x_i \in L^p(\Omega)$ *for* $1 \le i \le n$. *Then* $f \in L^p(\Omega)$ *and*

$$\left\| f - \frac{1}{\mu(W)} \int_W f \right\|_p \le (r_1 + r_2) \left(1 + \frac{r_2}{r_1}\right)^{\frac{n-1}{p}} \left(1 + \left(\frac{\mu(\Omega)}{\mu(W)}\right)^{\frac{1}{p}}\right) \sum_{i=1}^n \left\| \frac{\partial f}{\partial x_i} \right\|_p \tag{3.43}$$

for all measurable $W \subset \Omega$ *with measure* $\mu(W) > 0$.

PROOF Choose $r \in (0, r_1)$ and let $B = B(x_0, r)$.

If $x \in \Omega$, $y \in B$, then

$$f(x) = f(y) + \sum_{i=1}^n \int_0^1 (x_i - y_i) \frac{\partial f}{\partial x_i}(y + t(x-y))dt;$$

integrating with respect to y gives

$$\mu(B) f(x) = \int_B f(y)dy + \sum_{i=1}^n I_i(x), \tag{3.44}$$

where

$$I_i(x) = \int_B \int_0^1 (x_i - y_i) \frac{\partial f}{\partial x_i}(y + t(x-y))dt\,dy, \quad x \in \Omega,\ 1 \le i \le n.$$

Let us prove the following inequality:

$$\|I_i\|_p \le \mu(B)(r + r_2)(1 + r_2/r)^{(n-1)/p} \left\| \frac{\partial f}{\partial x_i} \right\|_p \quad \text{for} \quad 1 \le i \le n. \tag{3.45}$$

(3.45) is obvious if $p = \infty$; so, assume $p < \infty$ and note

$$\begin{aligned}
&|I_i(x)|^p \\
&\quad \le \left(\int_B \int_0^1 |x-y| dt\,dy \right)^{p-1} \int_B \int_0^1 |x-y| \left| \frac{\partial f}{\partial x_i}(y + t(x-y)) \right|^p dt\,dy \\
&\quad \le (\mu(B)(r+r_2))^{p-1} \int_B \int_0^1 |x-y| \left| \frac{\partial f}{\partial x_i}(y + t(x-y)) \right|^p dt\,dy \\
&\int_\Omega |I_i(x)|^p dx \\
&\quad \le (\mu(B)(r+r_2))^{p-1} \int_B dy \int_0^1 dt \int_\Omega dx\, |x-y| \left| \frac{\partial f}{\partial x_i}(y + t(x-y)) \right|^p \\
&\quad = (\mu(B)(r+r_2))^{p-1} \int_B dy \int_\Omega dx \int_0^1 dt\, t^{-n-1} g(x,y,t) |x-y| \left| \frac{\partial f}{\partial x_i}(x) \right|^p,
\end{aligned}$$

where $g(x,y,t) = 1$ if $x \in t\Omega + (1-t)y$ and otherwise $g(x,y,t) = 0$. If $x \in \Omega$, $y \in B$, $t \geq 0$ and $x \in t\Omega + (1-t)y$, then for some $z \in \Omega$, $|x-y| = t|z-y| \leq t(r_2 + r)$ and

$$\int_0^1 t^{-n-1} g(x,y,t)dt \leq \int_{\frac{|x-y|}{r_2+r}}^1 t^{-n-1}dt \leq \frac{1}{n}\left(\frac{r_2+r}{|x-y|}\right)^n.$$

Therefore

$$\int_\Omega |I_i(x)|^p dx \leq (\mu(B)(r+r_2))^{p-1}\frac{(r+r_2)^n}{n}\int_\Omega dx \int_B dy\,|x-y|^{1-n}\left|\frac{\partial f}{\partial x_i}(x)\right|^p$$

and inequality (3.45) follows from

$$\int_B |x-y|^{1-n}dy \leq \int_{B(0,r)} |y|^{1-n}dy = n\mu(B)r^{1-n}.$$

In view of (3.44) and (3.45), it is clear that $f \in L^p(\Omega)$. Integration of (3.44) over W gives

$$\mu(B)\int_W f = \mu(W)\int_B f + \sum_{i=1}^n \int_W I_i.$$

Dividing this by $\mu(W)$ and subtracting it from (3.44) gives

$$\begin{aligned}
\mu(B)\left(f - \frac{1}{\mu(W)}\int_W f\right) &= \sum_{i=1}^n \left(I_i - \frac{1}{\mu(W)}\int_W I_i\right) \\
\mu(B)\left\|f - \frac{1}{\mu(W)}\int_W f\right\|_p &\leq \sum_{i=1}^n \|I_i\|_p + \frac{1}{\mu(W)}\int_W |I_i|\mu(\Omega)^{1/p} \\
&\leq \left(1 + \left(\frac{\mu(\Omega)}{\mu(W)}\right)^{\frac{1}{p}}\right)\sum_{i=1}^n \|I_i\|_p.
\end{aligned}$$

Using (3.45) and letting $r \to r_1$ proves (3.43). □

3.6 Imbeddings of $W^{m,p}(\Omega)$

In this section we shall study imbeddings of $W^{m,p}(\Omega)$, where m is a positive integer and $1 \leq p \leq \infty$. We shall begin by showing (Theorem 3.6.3) that, under very weak conditions on Ω,

$$\text{if } mp > n \text{ and } j \geq 0, \text{ then } W^{j+m,p}(\Omega) \subset C_B^j(\Omega).$$

We shall then show (Lemma 3.6.8) that

$$\text{if } mp \leq n,\ p \leq q \leq \tfrac{np}{n-mp},\ q < \infty, \text{ then } W^{m,p}(\Omega) \subset L^q(\Omega)$$

when Ω is a box (possibly the whole $\mathbb{R}^n$). This will give us very general local imbeddings (Corollary 3.6.9), imbeddings of $W_0^{m,p}(\Omega)$ (Theorem 3.6.10), as well as Hölder continuity of functions in $W_0^{m,p}(\Omega)$ for $mp > n$ (Theorem 3.6.12). Under much weaker conditions on Ω, it will then be shown (Theorem 3.6.14) that $mp \leq n$ implies $W^{m,p}(\Omega) \subset L^q(\Omega)$ for $p \leq q < \frac{np}{n-mp}$. Sufficient conditions for compactness of the imbedding of $W^{m,p}(\Omega)$ into $L^q(\Omega)$ are presented in Theorem 3.6.16.

Imbeddings of $W^{m,p}(\Omega)$ in general depend on the boundary of Ω. On the other hand, $W_0^{m,p}(\Omega)$ is a closure of a set of functions that vanish on the boundary of Ω and hence, its imbeddings are, in general, independent of the boundary of Ω. The weakest, commonly used assumption on Ω for which most imbeddings still exist is called the cone property of Ω and is defined as follows:

Definition 3.6.1 For $a \in \mathbb{R}^n$, $a \neq 0$, $\theta \in (0, \pi]$, define

$$\text{cone}(a, \theta) = \{x \in \mathbb{R}^n \mid x \cdot a \geq |x||a| \cos\theta \text{ and } |x| \leq |a|\},$$

and observe that if $t > -n$, then

$$\int_{\text{cone}(a,\theta)} |x|^t dx = \mathrm{c}|a|^{t+n}$$

for some $\mathrm{c} \in (0, \infty)$ which depends only on n, t and θ.

A nonempty open set Ω in $\mathbb{R}^n$, $n \geq 1$, is said to have the **cone property** if there exist $h \in (0, \infty)$, $\theta \in (0, \pi]$ such that for every $x \in \Omega$ we can find $a \in \mathbb{R}^n$ with the properties $|a| = h$ and $x + \text{cone}(a, \theta) \subset \Omega$.

Lemma 3.6.2 *Suppose m is a positive integer, $p \in [1, \infty]$ and either $mp > n$ or $m \geq n$. Suppose also that Ω has the cone property and let h and θ be as in the Definition 3.6.1 of the cone property. Then there exists $\mathrm{c} \in (0, \infty)$ such that*

$$\sup_{\Omega} |u| \leq \mathrm{c} \sum_{i=0}^{m} h^{i-n/p} |u|_{i,p}$$

for all $u \in C^m(\Omega) \cap W^{m,p}(\Omega)$. c depends only on m, n, p and θ.

PROOF $\mathrm{c}_1, \mathrm{c}_2, \ldots$ will denote numbers in $(0, \infty)$ that are determined by m, p, n and θ only. Choose $u \in C^m(\Omega) \cap W^{m,p}(\Omega)$, $x \in \Omega$ and $a \in \mathbb{R}^n$ with $|a| = h$ so that $x + \text{cone}(a, \theta) \subset \Omega$. Abbreviate $C = \text{cone}(a, \theta)$ and let $\mathrm{c}_1 h^n$ be the volume of the cone C.

If $y \in C$, then $x + y - ty \in \Omega$ for $t \in [0,1]$. Hence the Taylor formula (1.8) can be written as

$$u(x) = \sum_{|\alpha|<m} \frac{(-1)^{|\alpha|}}{\alpha!} y^\alpha (D^\alpha u)(x+y) + m \sum_{|\alpha|=m} \frac{(-1)^{|\alpha|}}{\alpha!} \int_0^1 t^{m-1} y^\alpha (D^\alpha u)(x+ty) dt.$$

Integrating both sides over C gives

$$u(x) c_1 h^n = \sum_{|\alpha|<m} \frac{(-1)^{|\alpha|}}{\alpha!} I_\alpha + m \sum_{|\alpha|=m} \frac{(-1)^m}{\alpha!} J_\alpha, \tag{3.46}$$

where

$$I_\alpha = \int_C y^\alpha (D^\alpha u)(x+y) dy, \quad |\alpha| < m$$

$$J_\alpha = \int_0^1 \int_C t^{m-1} y^\alpha (D^\alpha u)(x+ty) dy\, dt, \quad |\alpha| = m.$$

Let $q \in [1, \infty]$ be such that $1/p + 1/q = 1$. Since $|y^\alpha| \le |y|^{|\alpha|}$, Hölder's inequality implies that

$$|I_\alpha| \le c_2 h^{i+n/q} \|D^\alpha u\|_p \le c_2 h^{i+n/q} |u|_{i,p} \quad \text{for} \quad |\alpha| = i < m. \tag{3.47}$$

When $|\alpha| = m$ the scaling gives

$$\begin{aligned} J_\alpha &= \int_0^1 \int_{tC} t^{-n-1} y^\alpha (D^\alpha u)(x+y) dy\, dt \\ &= \int_0^1 \int_C t^{-n-1} f(t,y) y^\alpha (D^\alpha u)(x+y) dy\, dt, \end{aligned}$$

where $f(t,y) = 1$ if $|y| \le ht$ and $f(t,y) = 0$ otherwise. Since

$$\int_0^1 t^{-n-1} f(t,y) dt = (h^n |y|^{-n} - 1)/n \quad \text{for} \quad y \in C,$$

we have

$$|J_\alpha| \le h^n \int_C |y|^{m-n} |(D^\alpha u)(x+y)| dy; \tag{3.48}$$

if $p = 1$, then $m \ge n$ and obviously

$$|J_\alpha| \le h^m |u|_{m,p}; \tag{3.49}$$

if $p > 1$, then $(m-n)q + n = q(m - n/p) > 0$ and we can apply Hölder's inequality to (3.48) to obtain

$$|J_\alpha| \le c_3 h^{m+n/q} |u|_{m,p}. \tag{3.50}$$

Inserting (3.47) and either (3.49) or (3.50) into (3.46) completes the proof. □

Theorem 3.6.3 *Suppose $m \geq 1$, $p \in [1,\infty]$ and either $mp > n$ or $m \geq n$. If Ω has the cone property and h and θ are as in Definition 3.6.1, then*

$$W^{j+m,p}(\Omega) \subset C_B^j(\Omega) \quad \text{for} \quad j \geq 0.$$

Moreover, there exists $c \in (0,\infty)$, depending only on m, n, p and θ, such that

$$|u|_{j,\infty} \leq c \sum_{i=0}^{m} h^{i-n/p} |u|_{i+j,p} \quad \text{for} \quad u \in W^{j+m,p}(\Omega),\ j \geq 0.$$

PROOF Lemma 3.6.2 and Theorem 3.5.5 imply the conclusions when $p < \infty$. If $p = \infty$, choose $r \in (1+n/m, \infty)$ and let B be a nonempty open ball contained in Ω. Since B has the cone property, $W^{j+m,p}(B) \subset W^{j+m,r}(B) \subset C_B^j(B)$. Hence, $W^{j+m,p}(\Omega) \subset C^j(\Omega)$. □

If $u \in W_0^{m,p}(\Omega)$ and $u = 0$ on $\mathbb{R}^n \backslash \Omega$, then $u \in W_0^{m,p}(\mathbb{R}^n)$; and since $\mathbb{R}^n$ has the cone property, we have

Corollary 3.6.4 *Suppose $j \geq 0$, $m \geq 1$, $1 \leq p \leq \infty$ and either $mp > n$ or $m \geq n$. Then $W_0^{j+m,p}(\Omega) \subset C_B^j(\Omega)$ and $W_{loc}^{j+m,p}(\Omega) \subset C^j(\Omega)$.*

We shall now begin studying imbeddings of $W^{m,p}$ into L^q for $q < \infty$.

Lemma 3.6.5 *If $\Omega = (a_1,b_1) \times \cdots \times (a_n,b_n)$ for some $-\infty \leq a_i < b_i \leq \infty$, then*

$$\|g\|_{\frac{n}{n-1}} \leq \varepsilon^{-1} \|g\|_1 + |g|_{1,1} \quad \text{for} \quad g \in W^{1,1}(\Omega),\ 0 < \varepsilon < \min_i b_i - a_i.$$

PROOF Suppose $g \in C^1(\Omega) \cap W^{1,1}(\Omega)$, $a_i < c_i < d_i < b_i$, $\varepsilon < d_i - c_i$, $\Omega' = (c_1,d_1) \times \cdots \times (c_n,d_n)$, and let $\int_i$ denote integration of x_i from c_i to d_i.

If $h \in C^1[a,b]$ and $a \leq z \leq b$, then

$$(b-a)h(z) = \int_a^b h(y)dy + \int_a^z (y-a)h'(y)dy - \int_z^b (b-y)h'(y)dy$$

$$|h(z)| \leq \frac{1}{b-a} \int_a^b |h(y)|dy + \int_a^b |h'(y)|dy.$$

Therefore

$$|g(x)| \leq \varepsilon^{-1} \int_i |g| + \int_i |D_i g| \quad \text{for} \quad x \in \Omega',\ i = 1,\ldots n.$$

It is easy to see that this implies the conclusion of the Lemma when $n = 1$; so assume from here on that $n > 1$ and $\alpha = \frac{1}{n-1}$. Note that

$$|g(x)|^{n\alpha} = \left(\int_1 f_1\right)^{\alpha} \cdots \left(\int_n f_n\right)^{\alpha},$$

where $f_i = \varepsilon^{-1}|g| + |D_i g|$. Integration of x_1, using Hölder's inequality (1.10) with $k = p_i = n-1$, gives

$$\int_1 |g|^{n\alpha} \le \left(\int_1 f_1\right)^{\alpha} \prod_{i=2}^{n} \left(\int_1 \int_i f_i\right)^{\alpha}.$$

Likewise, integration of $x_1, \dots, x_k$ gives

$$\int_1 \cdots \int_k |g|^{n\alpha} \le \prod_{i=1}^{k} \left(\int_1 \cdots \int_k f_i\right)^{\alpha} \prod_{i=k+1}^{n} \left(\int_1 \cdots \int_k \int_i f_i\right)^{\alpha}$$

for $k = 1, \dots, n$ and hence

$$\int_{\Omega'} |g|^{n\alpha} \le \prod_{i=1}^{n} \left(\int_{\Omega'} f_i\right)^{\alpha} \le \prod_{i=1}^{n} \left(\int_{\Omega} f_i\right)^{\alpha} \le \left(\varepsilon^{-1}\|g\|_1 + |g|_{1,1}\right)^{n\alpha};$$

thus $\|g\|_{\frac{n}{n-1}} \le \varepsilon^{-1}\|g\|_1 + |g|_{1,1}$ and Theorem 3.5.5 implies the conclusion. □

If $1 \le p \le q \le r \le \infty$, $p \ne r$ and $f \in L^p \cap L^r$, then Hölder's inequality implies

$$\|f\|_q \le \|f\|_p^{1-\theta} \|f\|_r^{\theta}, \quad \text{where} \quad \theta = \frac{\frac{1}{p} - \frac{1}{q}}{\frac{1}{p} - \frac{1}{r}}. \tag{3.51}$$

If this and the following fact,

$$A^{1-\theta} B^{\theta} \le \varepsilon^{-\theta} A + \varepsilon^{1-\theta} B \quad \text{when } A \ge 0,\ B \ge 0,\ \varepsilon > 0,\ \theta \in [0,1], \tag{3.52}$$

are used in Lemma 3.6.5, we obtain

Corollary 3.6.6 *Suppose* $\Omega = (a_1, b_1) \times \cdots \times (a_n, b_n)$ *for some* $-\infty \le a_i < b_i \le \infty$. *If* $1 \le q \le \infty$ *and* $\frac{1}{q} \ge 1 - \frac{1}{n}$, *then*

$$\|g\|_q \le \varepsilon^{-n(1-\frac{1}{q})} (2\|g\|_1 + \varepsilon |g|_{1,1}) \quad \textit{for} \quad g \in W^{1,1}(\Omega),\ 0 < \varepsilon < \min_i b_i - a_i.$$

Lemma 3.6.7 *Suppose* $\Omega = (a_1, b_1) \times \cdots \times (a_n, b_n)$ *for some* $-\infty \le a_i < b_i \le \infty$. *If* $1 \le p \le q < \infty$ *and* $\frac{1}{q} \ge \frac{1}{p} - \frac{1}{n}$, *then*

$$\|h\|_q \le \varepsilon^{-n(\frac{1}{p}-\frac{1}{q})} (2\|h\|_p + \varepsilon q |h|_{1,p}) \quad \textit{for} \quad h \in W^{1,p}(\Omega),\ 0 < \varepsilon < \min_i b_i - a_i.$$

PROOF Assume $p > 1$. Let $s = 1 + q\,\frac{p-1}{p} \in (1,\infty)$ and pick $f \in C^\infty(\Omega) \cap W^{1,p}(\Omega)$. Let $g = |f|^s$ and note that $g \in C^1(\Omega)$, $|D_i g| \le s|f|^{s-1}|D_i f|$. Let $\Omega' = (c_1,d_1)\times\cdots\times(c_n,d_n)$ for some $a_i < c_i < d_i < b_i$ with $\varepsilon < d_i - c_i$. Note

$$\|D_i g\|_{1,\Omega'} \le s\|D_i f\|_{p,\Omega'}\|f\|^{s-1}_{q,\Omega'}$$

$$\|g\|_{1,\Omega'} \le \|f\|_{p,\Omega'}\|f\|^{s-1}_{q,\Omega'}.$$

If $t = q/s$, then $\frac{1}{t} \ge 1 - \frac{1}{n}$. Hence Corollary 3.6.6 implies

$$\|f\|^s_{q,\Omega'} = \|g\|_{t,\Omega'} \le \varepsilon^{-n(1-\frac{1}{t})}(2\|f\|_{p,\Omega'}\|f\|^{s-1}_{q,\Omega'} + \varepsilon s|f|_{1,p,\Omega'}\|f\|^{s-1}_{q,\Omega'})$$

$$\|f\|_{q,\Omega'} \le \varepsilon^{-n(\frac{1}{p}-\frac{1}{q})}(2\|f\|_{p,\Omega'} + \varepsilon s|f|_{1,p,\Omega'}),$$

and letting $\Omega' \to \Omega$ and $f \to h$ (Theorem 3.5.5) implies the conclusion. □

Lemma 3.6.8 *Suppose $\Omega = (a_1,b_1)\times\cdots\times(a_n,b_n)$ for some $-\infty \le a_i < b_i \le \infty$. If $m \ge 1$, $1 \le p \le q < \infty$ and $\frac{1}{q} \ge \frac{1}{p} - \frac{m}{n}$, then*

$$\|h\|_q \le (2+q)^m \sum_{k=0}^{m} \varepsilon^{k-n(\frac{1}{p}-\frac{1}{q})}|h|_{k,p} \quad \text{for} \quad h \in W^{m,p}(\Omega),\ 0 < \varepsilon < \min_i b_i - a_i.$$

PROOF The conclusion of this Lemma is proven in Lemma 3.6.7 when $m = 1$. Assume that $m > 1$ and that the conclusion is true for $m-1$. Choose s so that

$$\max\{\tfrac{1}{q},\ \tfrac{1}{p} - \tfrac{m-1}{n}\} \le \tfrac{1}{s} \le \min\{\tfrac{1}{q} + \tfrac{1}{n},\ \tfrac{1}{p}\}.$$

Since $m - 1 \ge 1$, $1 \le p \le s < \infty$ and $\frac{1}{s} \ge \frac{1}{p} - \frac{m-1}{n}$, we have

$$\|h\|_s \le (2+q)^{m-1} \sum_{k=0}^{m-1} \varepsilon^{k-n(\frac{1}{p}-\frac{1}{s})}|h|_{k,p}$$

$$|h|_{1,s} \le (2+q)^{m-1} \sum_{k=0}^{m-1} \varepsilon^{k-n(\frac{1}{p}-\frac{1}{s})}|h|_{k+1,p}$$

and hence $h \in W^{1,s}(\Omega)$. Since $1 \le s \le q < \infty$, $\frac{1}{q} \ge \frac{1}{s} - \frac{1}{n}$, Lemma 3.6.7 implies

$$\begin{aligned}\|h\|_q &\le e^{-n(\frac{1}{s}-\frac{1}{q})}(2\|h\|_s + \varepsilon q|h|_{1,s}) \\ &\le (2+q)^{m-1}\mathrm{e}^{-n(\frac{1}{p}-\frac{1}{q})}\left(2\sum_{k=0}^{m-1}\varepsilon^k|h|_{k,p} + q\sum_{k=1}^{m}\varepsilon^k|h|_{k,p}\right),\end{aligned}$$

which shows that the conclusion is true also for m. □

Corollary 3.6.9 *Suppose* $j \geq 0$, $m \geq 1$, $1 \leq p \leq q < \infty$ *and* $\frac{1}{q} \geq \frac{1}{p} - \frac{m}{n}$. *Then* $W^{j+m,p}_{loc}(\Omega) \subset W^{j,q}_{loc}(\Omega)$.

Theorem 3.6.10 *Suppose* $j \geq 0$, $m \geq 1$, $1 \leq p \leq q \leq \infty$, $\frac{1}{q} \geq \frac{1}{p} - \frac{m}{n}$ *and, when* $q = \infty$, *it is required also that either* $mp > n$ *or* $m \geq n$. *Then* $W^{j+m,p}_0(\Omega) \subset W^{j,q}_0(\Omega)$ *and*

$$|u|_{j,q} \leq c|u|^{\theta}_{j+m,p}\|u\|^{1-\theta}_p \quad \textit{for} \quad u \in W^{j+m,p}_0(\Omega),\ \theta = \tfrac{n}{j+m}(\tfrac{j}{n} + \tfrac{1}{p} - \tfrac{1}{q}),$$

where $c \in (0, \infty)$ *depends only on* j, m, n, p *and* q.

PROOF Choose $v \in C^\infty_0(\Omega)$ and let $v = 0$ in $\mathbb{R}^n \backslash \Omega$. Lemmas 3.6.8 and 3.6.2, with $\mathbb{R}^n$ in place of Ω, imply

$$|v|_{j,q} \leq c_1 \sum_{k=0}^{m} \varepsilon^{k-n(\frac{1}{p}-\frac{1}{q})}|v|_{k+j,p} \quad \text{for} \quad \varepsilon > 0.$$

Lemma 3.4.14 implies

$$|v|_{j,q} \leq c_2(\varepsilon^{m-n(\frac{1}{p}-\frac{1}{q})}|v|_{j+m,p} + \varepsilon^{-j-n(\frac{1}{p}-\frac{1}{q})}\|v\|_p) \quad \text{for} \quad \varepsilon > 0;$$

hence (3.15) implies the bound stated. Since this is true for every $j \geq 0$ and $C^\infty_0(\Omega)$ is dense in $W^{j+m,p}_0(\Omega)$, the rest follows. □

Before continuing with our study of imbeddings of $W^{m,p}$ into L^q, $q < \infty$, let us apply the above result to show the Hölder continuity of functions in $W^{m,p}_0(\Omega)$ when $mp > n$.

Lemma 3.6.11 *If* $p > n$, *then there exists* $c \in (0, \infty)$ *such that*

$$|u(x) - u(y)| \leq c|x - y|^{1-n/p}|u|_{1,p} \quad \textit{for all} \quad x, y \in \mathbb{R}^n,\ u \in C^1_0(\mathbb{R}^n).$$

PROOF Suppose $u \in C^1_0(\mathbb{R}^n)$, $x, y \in \mathbb{R}^n$ and $d = |x - y|$. Note that

$$u(x) = u(x + z) - \sum_{i=1}^{n} \int_0^1 z_i(D_i u)(x + tz)dt.$$

Subtracting from this expression the corresponding expression with y in place of x, and integrating z over the ball $B = B(0, d)$, gives

$$(u(x) - u(y))c_0 d^n = I - J(x) + J(y), \tag{3.53}$$

where

$$I = \int_B (u(x + z) - u(y + z))dz$$

$$J(x) = \sum_{i=1}^{n} \int_B \int_0^1 z_i (D_i u)(x + tz) dt\, dz$$

and $c_0, c_1, \ldots$ denote numbers in $(0, \infty)$ that depend only on n and p. Using the bound

$$\begin{aligned} J(x) &= \sum_{i=1}^{n} \int_0^1 \int_{tB} t^{-n-1} z_i (D_i u)(x + z) dz\, dt \\ |J(x)| &\leq \int_0^1 t^{-n-1} c_1 (td)^{1+n-n/p} |u|_{1,p} dt \\ &\leq c_2 d^{1+n-n/p} |u|_{1,p}, \end{aligned}$$

as well as

$$\begin{aligned} I &= \int_B \sum_{i=1}^{n} (x_i - y_i) \int_0^1 (D_i u)(y + z + t(x - y)) dt\, dz \\ |I| &\leq c_3 \sum_{i=1}^{n} |x_i - y_i| d^{n-n/p} |u|_{1,p} \\ &\leq c_4 d^{1+n-n/p} |u|_{1,p} \end{aligned}$$

in (3.53) completes the proof. □

Theorem 3.6.12 *Suppose $p \in [1, \infty]$ and m is an integer such that $0 < m - \frac{n}{p} \leq 1$.*

If either $m - \frac{n}{p} < 1$ or $m = 1$ or $m - n = 1$, choose any $\sigma \in \left(0, m - \frac{n}{p}\right]$;

otherwise, choose any $\sigma \in (0, 1)$. If $\theta = \frac{\sigma}{m} + \frac{n}{mp}$, then there exists $c < \infty$ such that

$$\max_{|\alpha|=j} |(D^\alpha u)(x) - (D^\alpha u)(y)| \leq c|x - y|^\sigma |u|^\theta_{j+m,p} |u|^{1-\theta}_{j,p}$$

for all $x, y \in \Omega$, all $j \geq 0$ and all $u \in W_0^{j+m,p}(\Omega)$.

PROOF Since $C_0^\infty(\mathbb{R}^n)$ is dense in $W_0^{j+m,p}(\Omega)$, it is enough to prove the Theorem for $u \in C_0^m(\mathbb{R}^n)$ when $j = 0$.

Assume first that $m = 1$ and let $\alpha = \sigma/(1 - \frac{n}{p})$. Lemma 3.6.11 and Theorem 3.6.10 imply that for all $x, y \in \mathbb{R}^n$,

$$\begin{aligned} |u(x) - u(y)| &\leq |u(x) - u(y)|^\alpha |u(x) - u(y)|^{1-\alpha} \\ &\leq c\left(|x - y|^{1-\frac{n}{p}} |u|_{1,p}\right)^\alpha \left(|u|^{\frac{n}{p}}_{1,p} \|u\|_p^{1-\frac{n}{p}}\right)^{1-\alpha} \\ &= c|x - y|^\sigma |u|^\theta_{1,p} \|u\|_p^{1-\theta}. \end{aligned}$$

Suppose now that $m \geq 2$. Define $q = \frac{n}{1-\sigma} \in (n, \infty]$ and note that if $q = \infty$, then $m - 1 = n$. Since $p \leq \frac{n}{m-1} \leq n < q$ and $\frac{1}{q} \geq \frac{1}{p} - \frac{m-1}{n}$, we can apply Theorem 3.6.10 to obtain that $|u|_{1,q} \leq c_1 |u|_{m,p}^{\theta} \|u\|_p^{1-\theta}$. Lemma 3.6.11 implies

$$|u(x) - u(y)| \leq c_2 |x-y|^{1-n/q} |u|_{1,q} \leq c_2 |x-y|^{\sigma} c_1 |u|_{m,p}^{\theta} \|u\|_p^{1-\theta}.$$

□

In Lemma 3.6.8 there are severe restrictions on Ω; nevertheless, it implies very general local imbeddings. To obtain such results for more general Ω, one can proceed in several different ways. One way is to prove extension theorems first, i.e. one needs to find a map $E : W^{m,p}(\Omega) \to W^{m,p}(\mathbb{R}^n)$ such that for every $u \in W^{m,p}(\Omega)$

(1) $(Eu)(x) = u(x)$ for $x \in \Omega$

(2) $\|Eu\|_{m,p,\mathbb{R}^n} \leq c \|u\|_{m,p,\Omega}$.

Such maps E exist under various conditions and thus the conclusions of Lemma 3.6.8 become applicable to more general Ω. We will not use this approach. Instead we will present a simple direct proof of the conclusions of Lemma 3.6.8 for any Ω that satisfies the cone condition, provided that $\frac{1}{q} > \frac{1}{p} - \frac{m}{n}$. This non-sharpness could be removed by simply using the Hardy-Littlewood-Sobolev inequality instead of Young's inequality to prove (3.55) in the critical case where $\frac{1}{q} = \frac{1}{p} - \frac{m}{n} > 0$, $p > 1$.

Lemma 3.6.13 *Suppose Ω has the cone property and let h and θ be as in Definition 3.6.1 of the cone property. If $m \geq 1$, $1 \leq p \leq q \leq \infty$ and $\frac{1}{q} > \frac{1}{p} - \frac{m}{n}$, then there exists $c < \infty$, depending only on m, n, p, q and θ, such that*

$$\|u\|_q \leq c \sum_{i=0}^{m} h^{i-n(\frac{1}{p}-\frac{1}{q})} |u|_{i,p} \quad \textit{for} \quad u \in C^m(\Omega) \cap W^{m,p}(\Omega).$$

PROOF If $p = q$, the statement is obvious; if $q = \infty$, then the assertion follows from Lemma 3.6.2. So, assume $p < q < \infty$. $c_1, c_2, \ldots$ will denote numbers in $(0, \infty)$ that are determined by m, n, p, q and θ only.

For $j = (j_1, \ldots, j_n) \in \mathbb{Z}^n$, define

$$\text{Box}(j) = \{x \in \mathbb{R}^n \mid j_i \leq x_i/h < j_i + 1 \text{ for all } i = 1, \ldots, n\}$$

$$\text{Box}(j)^* = \{x \in \mathbb{R}^n \mid j_i - 1 \leq x_i/h < j_i + 2 \text{ for all } i = 1, \ldots, n\}$$

and observe that $\text{Box}(j)^*$ is a union of 3^n disjoint boxes, $\text{Box}(j_a)$, that are adjacent to Box(j). Let $\{V_1, V_2, \ldots\}$ denotes the collection (possibly finite) of those sets $\Omega \cap \text{Box}(j)$, $j \in \mathbb{Z}^n$, which are not empty. For each V_i there exists a unique $j \in \mathbb{Z}^n$ such that $V_i = \Omega \cap \text{Box}(j)$, define $V_i^* = \Omega \cap \text{Box}(j)^*$. The following facts will be needed later on:

(F1) $V_i \cap V_j$ is empty if $i \neq j$; $\Omega = \cup V_i$

(F2) if $x \in V_i$, $y \in \Omega$ and $|x - y| \leq h$, then $y \in V_i^*$

(F3) if $f \in L^1(\Omega)$, $f \geq 0$, then

$$\sum_i \int_{V_i^*} f \leq 3^n \int_\Omega f.$$

Choose any $x \in \Omega$. There exists a unique V_i such that $x \in V_i$. Let $a \in \mathbb{R}^n$ be such that $|a| = h$ and $x + \text{cone}(a, \theta) \subset \Omega$. Abbreviate $C = \text{cone}(a, \theta)$ and let $c_1 h^n$ be the volume of C. Observe that $x + C \subset V_i^*$, by (F2). In exactly the same way as in the proof of Lemma 3.6.2 we obtain that

$$u(x) c_1 h^n = \sum_{|\alpha|<m} \frac{(-1)^{|\alpha|}}{\alpha!} I_\alpha + m \sum_{|\alpha|=m} \frac{(-1)^m}{\alpha!} J_\alpha,$$

where

$$|I_\alpha| \leq \int_C |y|^{|\alpha|} |(D^\alpha u)(x + y)| dy, \quad |\alpha| < m$$

$$|J_\alpha| \leq h^n \int_C |y|^{m-n} |(D^\alpha u)(x + y)| dy, \quad |\alpha| = m.$$

The change of variable $x + y \to y$ and the fact that $x + C \subset V_i^*$ give

$$|u(x)| \leq c_2 \sum_{|\alpha|<m} h^{-n} F_\alpha(x) + c_2 \sum_{|\alpha|=m} G_\alpha(x), \tag{3.54}$$

where

$$F_\alpha(x) = \int_{V_i^*} |x - y|^{|\alpha|} |(D^\alpha u)(y)| dy, \quad x \in V_i, \ |\alpha| < m$$

$$G_\alpha(x) = \int_{V_i^*} |x - y|^{m-n} |(D^\alpha u)(y)| dy, \quad x \in V_i, \ |\alpha| = m.$$

Observe that $F_\alpha(x)$, $G_\alpha(x)$ are uniquely defined for each $x \in \Omega$ and that (3.54) holds for every $x \in \Omega$.

Fix any multi-index α such that $|\alpha| \leq m$. Let $t = |\alpha|$ if $|\alpha| < m$ and let $t = m - n$ if $|\alpha| = m$. Define f on all of Ω by

$$f(x) = \int_{V_i^*} |x - y|^t |(D^\alpha u)(y)| dy \quad \text{for} \quad x \in V_i.$$

Hence, $f = F_\alpha$ if $|\alpha| < m$ and $f = G_\alpha$ if $|\alpha| = m$. f equals $k * (\chi_i |D^\alpha u|)$ in V_i, where χ_i is the characteristic function of the set V_i^*, $k(y) = |y|^t$ when $|y| < 2nh$ and 0 in the rest of $\mathbb{R}^n$. Define $\frac{1}{s} = 1 - \frac{1}{p} + \frac{1}{q} \in [\frac{1}{q}, 1)$ and note that

$(m-n)s+n>0$; hence, $ts+n>0$ and therefore $k\in L^s(\mathbb{R}^n)$, $\|k\|_s\le c_3h^{t+n/s}$. Young's inequality (page 112) implies

$$\left(\int_{V_i} f^q\right)^{\frac{1}{q}} \le c_3h^{t+\frac{n}{s}}\left(\int_{V_i^*}|D^\alpha u|^p\right)^{\frac{1}{p}}. \tag{3.55}$$

Hence

$$\int_{V_i} f^q \le c_3^q h^{qt+\frac{nq}{s}}\|D^\alpha u\|_p^{q-p}\int_{V_i^*}|D^\alpha u|^p,$$

and (F1), (F3) imply

$$\int_\Omega f^q \le c_3^q h^{qt+\frac{nq}{s}}\|D^\alpha u\|_p^{q-p}\sum_i\int_{V_i^*}|D^\alpha u|^p \le 3^n c_3^q h^{qt+\frac{nq}{s}}\|D^\alpha u\|_p^{q-p}\int_\Omega |D^\alpha u|^p$$

$$\|f\|_q \le c_4 h^{t+\frac{n}{s}}\|D^\alpha u\|_p.$$

Therefore

$$h^{-n}\|F_\alpha\|_q \le c_4 h^{i+n(\frac{1}{q}-\frac{1}{p})}|u|_{i,p} \quad \text{if } |\alpha|=i<m$$

$$\|G_\alpha\|_q \le c_4 h^{m+n(\frac{1}{q}-\frac{1}{p})}|u|_{m,p} \quad \text{if } |\alpha|=m$$

and if these bounds are used in (3.54), the assertion of the Lemma follows. □

Theorem 3.6.14 *Suppose Ω has the cone property and let h, θ be as in the Definition 3.6.1 of the cone property. If $j\ge 0$, $m\ge 1$, $1\le p\le q\le\infty$ and $\frac{1}{q}>\frac{1}{p}-\frac{m}{n}$, then $W^{j+m,p}(\Omega)\subset W^{j,q}(\Omega)$ and there exists $c<\infty$, depending only on m, n, p, q and θ, such that*

$$|u|_{j,q}\le c\sum_{i=0}^m h^{i-n(\frac{1}{p}-\frac{1}{q})}|u|_{i+j,p} \quad \textit{for} \quad u\in W^{j+m,p}(\Omega).$$

PROOF If $q=\infty$, then the conclusions follow from Theorem 3.6.3 and otherwise from Lemma 3.6.13 and Theorem 3.5.5. □

We shall now present sufficient conditions for compactness of the imbedding of $W^{m,p}$ into L^q. We need first to generalize the Arzela-Ascoli Theorem to L^p spaces. Let T_h denote the translation operator,

$$(T_hu)(x)=u(x+h) \quad \text{for} \quad x,h\in\mathbb{R}^n.$$

Theorem 3.6.15 *Suppose that $\{u_i\}_{i=1}^\infty$ is a bounded sequence in $L^p(\mathbb{R}^n)$, $1\le p<\infty$, such that for each $\varepsilon>0$ there exists $\delta>0$ so that*

$$\|T_hu_i-u_i\|_p<\varepsilon \quad \textit{for all} \quad |h|<\delta,\ i=1,2,\dots.$$

Then there exist integers $n_1 < n_2 < \cdots$ and $v \in L^p(\mathbb{R}^n)$ such that

$$\lim_{i\to\infty} \int_{|x|<r} |u_{n_i}(x) - v(x)|^p dx = 0 \quad \textit{for all} \quad r \in (0,\infty).$$

PROOF For every $\lambda > 0$, $i \in \mathbb{N}$ we have that $J_\lambda * u_i \in C^\infty(\mathbb{R}^n)$ and

$$\|J_\lambda * u_i\|_\infty \le \|J_\lambda\|_q \sup_j \|u_j\|_p$$

$$\|T_h J_\lambda * u_i - J_\lambda * u_i\|_\infty = \|J_\lambda * (T_h u_i - u_i)\|_\infty \le \|J_\lambda\|_q \sup_j \|T_h u_j - u_j\|_p$$

for all $h \in \mathbb{R}^n$, where $1/q = 1 - 1/p$. Thus, we can repeatedly apply the Arzela-Ascoli Theorem 1.1.5 to obtain infinite sets of integers $\mathbb{N} \supset A_1 \supset A_2 \supset \cdots$ and continuous functions $v_1, v_2, \ldots$ such that for each $k \ge 1$,

$$\lim_{i\to\infty, i\in A_k} \sup_{B(0,r)} |J_{1/k} * u_i - v_k| = 0 \quad \text{for all} \quad r \in (0,\infty).$$

Let $n_0 = 1$ and pick $n_k \in A_k$ such that $n_k > n_{k-1}$ for $k \ge 1$. Note that

$$\lim_{i\to\infty} \sup_{B(0,r)} |J_{1/k} * u_{n_i} - v_k| = 0 \quad \text{for all} \quad r \in (0,\infty),\ k \ge 1. \tag{3.56}$$

We claim that $\{u_{n_i}\}$ is a Cauchy sequence in $L^p(B(0,r))$ for any $r \in (0,\infty)$. Fix an $r \in (0,\infty)$ and pick $\varepsilon > 0$. Since for all $u \in L^p(\mathbb{R}^n)$, $\lambda > 0$,

$$\begin{aligned}
(J_\lambda * u)(x) - u(x) &= \int_{B(0,\lambda)} J_\lambda(y)(u(x-y) - u(x))dy \\
|(J_\lambda * u)(x) - u(x)|^p &\le \int_{B(0,\lambda)} J_\lambda(y)|u(x-y) - u(x)|^p dy \\
\|J_\lambda * u - u\|_p^p &\le \int_{B(0,\lambda)} J_\lambda(y)\|T_{-y}u - u\|_p^p dy \\
\|J_\lambda * u - u\|_p &\le \sup_{|y|\le\lambda} \|T_y u - u\|_p,
\end{aligned}$$

we can choose k so that

$$\|J_{1/k} * u_i - u_i\|_p < \varepsilon/6 \quad \text{for all} \quad i \in \mathbb{N}.$$

(3.56) implies that there exists N such that for all $i > N$,

$$\|J_{1/k} * u_{n_i} - v_k\|_{p,B(0,r)} \le \mu(B(0,r))^{1/p} \sup_{B(0,r)} |J_{1/k} * u_{n_i} - v_k| < \varepsilon/3.$$

Therefore, if $i, j > N$, then

$$\begin{aligned}\|u_{n_i} - u_{n_j}\|_{p,B(0,r)} &= \\ \|u_{n_i} - J_{1/k} * u_{n_i} + J_{1/k} * u_{n_i} - v_k + v_k - J_{1/k} * u_{n_j} + J_{1/k} * u_{n_j} - u_{n_j}\|_{p,B(0,r)} &\\ < \varepsilon/6 + \varepsilon/3 + \varepsilon/3 + \varepsilon/6 = \varepsilon,&\end{aligned}$$

which proves the claim.

$L^p(B(0,k))$, $k \in \mathbb{N}$, limits of $\{u_{n_i}\}$ define a function v almost everywhere in $\mathbb{R}^n$ and, since $\|v\|_{p,B(0,k)} \leq \sup_j \|u_j\|_p < \infty$, we have that $v \in L^p(\mathbb{R}^n)$. □

The following Theorem is our main result on compactness of imbeddings. It says, that when an imbedding exists, then it is also compact except possibly at the critical values of parameters.

Theorem 3.6.16 *Suppose Ω is bounded and that it has the cone property. If $j \geq 0$, $m \geq 1$, $1 \leq p < \infty$, $1 \leq q < \infty$, $\frac{1}{q} > \frac{1}{p} - \frac{m}{n}$ and $\{u_i\}_{i=1}^{\infty}$ is a bounded sequence in $W^{j+m,p}(\Omega)$, then there exist integers $n_1 < n_2 < \cdots$ and $v \in W^{j,q}(\Omega)$ such that*

$$\lim_{i\to\infty} \|u_{n_i} - v\|_{j,q} = 0.$$

PROOF Assume first that $j = 0$.

Choose $r \in (q, \infty) \cap (p, \infty)$ such that $\frac{1}{r} > \frac{1}{p} - \frac{m}{n}$. Theorem 3.6.14 implies that $\|u\|_r \leq c_1 \|u\|_{m,p}$ for all $u \in W^{m,p}(\Omega)$. Let $M = \sup_i \|u_i\|_{m,p}$ and define $u_i = 0$ in $\mathbb{R}^n \backslash \Omega$. Note that $\|u_i\|_{1,\mathbb{R}^n} \leq \|u_i\|_r \mu(\Omega)^{1-1/r} \leq c_1 M \mu(\Omega)^{1-1/r}$.

We claim that $\sup_i \|T_h u_i - u_i\|_{1,\mathbb{R}^n} \to 0$ as $|h| \to 0$, where T_h is the translation operator. Pick $\varepsilon > 0$. Choose a compact $K \subset \Omega$ such that

$$4c_1 M \, \mu(\Omega \backslash K)^{1-1/r} < \varepsilon,$$

and pick $d \in (0, \operatorname{dist}(K, \Omega^c)/2)$. Let $\omega = K + B(0,d)$ and note that when $|h| < d$,

$$\int_{\mathbb{R}^n \backslash \omega} |u_i(x+h)| dx \leq \int_{\Omega \backslash K} |u_i(x)| dx \leq \|u_i\|_r \mu(\Omega \backslash K)^{1-1/r} < \varepsilon/4;$$

hence

$$\int_{\mathbb{R}^n \backslash \omega} |u_i(x+h) - u_i(x)| dx < \varepsilon/2 \quad \text{for} \quad i \in \mathbb{N}, \ |h| < d. \tag{3.57}$$

For $f \in C^\infty(\Omega) \cap W^{m,p}(\Omega)$ and $|h| < d$ we have that

$$\int_\omega |f(x+h) - f(x)| dx = \int_\omega \left| \int_0^1 \sum_{i=1}^n h_i (D_i f)(x + ht) dt \right| dx$$

$$\begin{aligned} &\leq \int_0^1 \int_{ht+\omega} \sum_{i=1}^n |h_i(D_i f)(x)| dx\, dt \\ &\leq \int_\Omega \sum_{i=1}^n |h_i(D_i f)(x)| dx \\ &\leq n|h| \|f\|_{1,p} \mu(\Omega)^{1-1/p} \end{aligned}$$

by using Hölder's inequality twice. Hence, Theorem 3.5.5 implies that

$$\int_\omega |u_i(x+h) - u_i(x)| dx \leq n|h| M \mu(\Omega)^{1-1/p} \quad \text{for} \quad i \in \mathbb{N},\ |h| < d.$$

This and (3.57) imply

$$\int_{\mathbb{R}^n} |u_i(x+h) - u_i(x)| dx < \varepsilon/2 + n|h| M \mu(\Omega)^{1-1/p} \quad \text{for} \quad i \in \mathbb{N},\ |h| < d,$$

which proves the claim.

Theorem 3.6.15 implies that there exist $n_1 < n_2 < \cdots$ and $v \in L^1(\Omega)$ such that u_{n_i} converge in $L^1(\Omega)$ to v. (3.51) implies that $\{u_{n_i}\}$ is also a Cauchy sequence in $L^q(\Omega)$. This proves the Theorem in the case $j = 0$.

Since $\{\mathrm{D}^\alpha u_i\}$ is a bounded sequence in $W^{m,p}(\Omega)$ for all $|\alpha| \leq j$, we can extract subsequences $\{\mathrm{D}^\alpha u_{n_i}\}$ which converge in $L^q(\Omega)$ for all $|\alpha| \leq j$ - implying convergence of $\{u_{n_i}\}$ in $W^{j,q}(\Omega)$. □

Corollary 3.6.17 *Suppose that Ω is bounded, $m \geq 1$, $1 \leq p < \infty$, $1 \leq q < \infty$ and $\frac{1}{q} > \frac{1}{p} - \frac{m}{n}$. Then the identity map from $W_0^{m,p}(\Omega)$ to $L^q(\Omega)$ is compact.*

3.7 Elliptic Problems

A version of the **Dirichlet problem** can be stated as follows: for a given f, find u such that

$$\begin{aligned} (\Delta u)(x) &= -f(x) \quad \text{for} \quad x \in \Omega \\ u(x) &= 0 \quad \text{for} \quad x \in \partial\Omega. \end{aligned}$$

Green's identity formally implies

$$\int_\Omega \Delta u\, \overline{v} = -\sum_{i=1}^n \int_\Omega D_i u \overline{D_i v} + \int_{\partial\Omega} \frac{\partial u}{\partial n} \overline{v}.$$

Thus, if v vanishes on $\partial\Omega$ just like u does, then $\sum_{i=1}^n \int_\Omega D_i u \overline{D_i v} = \int_\Omega f \overline{v}$. This suggests the following definition of a sectorial form:

$$\mathfrak{F}(u,v) = \sum_{i=1}^n \int_\Omega \mathrm{D}_i u \overline{\mathrm{D}_i v} \quad \text{for} \quad u, v \in \mathcal{V} \tag{3.58}$$

where $\mathcal{V} = W_0^1(\Omega)$. If $\mathcal{H} = L^2(\Omega)$, then assumptions **H1**, **H2**, **H3** of the section on Sectorial Forms are satisfied. The Dirichlet problem can be reformulated as:

$$\text{for given } f \in \mathcal{H} \text{ find } u \in \mathcal{V} \text{ such that } \mathfrak{F}(u,v) = (f,v) \text{ for all } v \in \mathcal{V}; \tag{3.59}$$

or equivalently, find $u \in \mathcal{D}(A)$ such that $Au = f$, where A is the operator associated with $\mathfrak{F}$. The sectorial form $\mathfrak{F}$ and the operator A associated with it are defined without any restrictions on Ω - any nonempty open set in $\mathbb{R}^n$ will do. Here are some of the properties of A:

(1) A is self-adjoint (Corollary 2.8.5).

(2) Let $\lambda(\Omega)$ be the largest of real numbers λ for which

$$\sum_{i=1}^{n} \int_\Omega |\mathsf{D}_i u|^2 \geq \lambda \int_\Omega |u|^2 \quad \text{for all} \quad u \in W_0^1(\Omega).$$

Theorem 2.8.7 implies that $\lambda(\Omega) \in \sigma(A) \subset [\lambda(\Omega), \infty)$. If $u \in W_0^1(\Omega)$ and $u = 0$ outside Ω, then $u \in W_0^1(\Omega')$ for every open set $\Omega' \supset \Omega$ which implies that $\lambda(\Omega) \geq \lambda(\Omega')$.

(3) If Ω has a finite thickness d in some direction, then the Ponicare inequality (3.42) implies that $\lambda(\Omega) \geq d^{-2}$; hence, the Dirichlet problem (3.59) has a solution for every $f \in \mathcal{H}$ and, by Theorem 2.8.9, the solution is the minimizer of

$$\mathfrak{F}(v,v) - 2\mathrm{Re}\,(f,v), \quad v \in \mathcal{V}.$$

(4) If Ω is bounded, then A has compact resolvent which implies that $\lambda(\Omega)$ is the smallest eigenvalue of A (see Corollary 3.6.17 and Theorems 2.8.2, 2.6.8).

It is easy to see that the operator A can also be represented as follows: $u \in \mathcal{D}(A)$ iff $u \in W_0^1(\Omega)$ and there exists $f \in L^2(\Omega)$ such that $(u, \Delta\phi) = -(f,\phi)$ for all $\phi \in C_0^\infty(\Omega)$; $Au = f$. This representation is usually abbreviated as

$$\mathcal{D}(A) = \{u \in W_0^1(\Omega) \,|\, \Delta u \in L^2(\Omega)\}, \quad Au = -\Delta u \text{ for } u \in \mathcal{D}(A). \tag{3.60}$$

One can show that when the boundary of Ω is nice enough, then $\mathcal{D}(A) = W_0^1(\Omega) \cap W^2(\Omega)$; however, it can happen that $\mathcal{D}(A) \not\subset W^2(\Omega)$, see Exercise 13. Regularity of solutions of elliptic problems is investigated in Section 3.8 where it is shown in particular that if $u \in \mathcal{D}(A)$, then $u \in W^2_{loc}(\Omega)$ (Example 3.8.4). The one dimensional case is discussed in Examples 2.8.1, 2.8.3 and 2.8.8.

One can attempt to solve the problem $\Delta w = -h$ in Ω and $w = g$ on $\partial\Omega$ by first choosing a smooth enough v defined in Ω such that $v = g$ on $\partial\Omega$ and then declaring the solution to be $w = v + u$ where u satisfies (3.59) with $f = h + \Delta v$.

The **Neumann problem** can be stated as follows: for a given f, find u such that

$$(\Delta u)(x) = -f(x) \quad \text{for} \quad x \in \Omega$$

$$\frac{\partial u}{\partial n}(x) = 0 \quad \text{for} \quad x \in \partial\Omega.$$

The variational formulation of the problem can be stated again by (3.59) where $\mathfrak{F}$ is also given by (3.58), the only difference being that $\mathcal{V} = W^1(\Omega)$ in this case. In this formulation, Ω can again be any nonempty open set in $\mathbb{R}^n$; however, $\partial u/\partial n$ does not necessarily make sense on the boundary - more assumptions are needed about the boundary of Ω to show that the the solution of (3.59) actually satisfies $\partial u/\partial n = 0$. Here are some properties of operator A associated with this $\mathfrak{F}$:

(1) A is self-adjoint (Corollary 2.8.5).

(2) $\sigma(A) \subset [0, \infty)$ (Theorem 2.8.7).

(3) if Ω is bounded and has the cone property, then A has compact resolvent (Theorems 3.6.16, 2.8.2), see also Theorems 2.6.8 and 1.7.16.

(4) if Ω is bounded, connected and has the cone property, then $\mathcal{N}(A)$ consists of constant functions only (Theorem 3.4.8). Hence, the Neumann problem (3.59) has a solution iff $\int_\Omega f = 0$ (Theorem 2.2.8).

Other choices of a closed subspace of $W^1(\Omega)$ for $\mathcal{V}$ correspond to other boundary conditions. For some boundary conditions, the sectorial form $\mathfrak{F}$ must also be modified (see, for example, (2.58)).

Let us now turn to the general **2nd order elliptic differential operator**,

$$Lu = -\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left(a_{ij}\frac{\partial u}{\partial x_i}\right) + \sum_{i=1}^{n} a_i \frac{\partial u}{\partial x_i} + a_0 u.$$

Green's identity suggests to consider the form,

$$\mathfrak{F}(u,v) = \sum_{i=1}^{n}\sum_{j=1}^{n} \int_\Omega a_{ij} \mathsf{D}_i u \, \overline{\mathsf{D}_j v} + \sum_{i=1}^{n} \int_\Omega a_i \mathsf{D}_i u \, \overline{v} + \int_\Omega a_0 u \overline{v} \quad \text{for} \quad u, v \in \mathcal{V}, \tag{3.61}$$

where $\mathcal{V}$ is a closed subspace of $W^1(\Omega)$. We assume that a_{ij}, a_i are bounded, complex valued, measurable functions on Ω such that $\operatorname{Im} a_{ij} = \operatorname{Im} a_{ji}$ and that there exists $\delta > 0$ such that

$$\operatorname{Re} \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \delta|\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^n, x \in \Omega. \tag{3.62}$$

(3.62) is known as the **strong ellipticity condition**. (3.62) and $\operatorname{Im} a_{ij} = \operatorname{Im} a_{ji}$ imply

$$\operatorname{Re} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}(x) z_i \overline{z_j} \geq \delta |z|^2 \quad \text{for} \quad z \in \mathbb{C}^n, x \in \Omega,$$

and therefore

$$\operatorname{Re} \sum_{i=1}^{n} \sum_{j=1}^{n} \int_{\Omega} a_{ij} \mathrm{D}_i u \, \overline{\mathrm{D}_j u} \geq \delta \sum_{i=1}^{n} \int_{\Omega} |\mathrm{D}_i u|^2 = \delta(\|u\|_{1,2}^2 - \|u\|_2^2)$$

$$\operatorname{Re} \mathfrak{F}(u,u) \geq \delta(\|u\|_{1,2}^2 - \|u\|_2^2) - \mathrm{c}\|u\|_{1,2}\|u\|_2,$$

where $\mathrm{c} = \| \sum_{i=0}^{n} |a_i|^2 \|_{\infty}^{1/2}$. Hence, for any $M_3 \in (0, \delta)$, there exists $a \in \mathbb{R}$ such that

$$\operatorname{Re} \mathfrak{F}(u,u) \geq M_3 \|u\|_{1,2}^2 + a\|u\|_2^2 \quad \text{for all} \quad u \in \mathcal{V}. \tag{3.63}$$

For elliptic problems, the inequality of form (3.63), with $M_3 > 0$, is called **Gårding's inequality**. If $\mathcal{H} = L^2(\Omega)$, then assumptions **H1**, **H2**, **H3** of the section on Sectorial Forms are again satisfied. In particular, when Ω is bounded and has the cone property, we can again apply the Fredholm alternative: problem (3.59), with $\mathfrak{F}$ given by (3.61), has a solution iff $(f, v) = 0$ for all $v \in \mathcal{V}$ such that $\mathfrak{F}(g, v) = 0$ for all $g \in \mathcal{V}$.

Consider now the **biharmonic operator**,

$$Lu = \Delta^2 u \quad \text{with } u = \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega.$$

Green's identity suggests we define

$$\mathfrak{F}(u,v) = \int_{\Omega} \Delta u \, \overline{\Delta v} \quad \text{for} \quad u, v \in W_0^2(\Omega). \tag{3.64}$$

If $u \in C_0^{\infty}(\Omega)$, let $u = 0$ in $\mathbb{R}^n \backslash \Omega$ and note that

$$\begin{aligned} \mathfrak{F}(u,u) &= \int_{\mathbb{R}^n} |\xi|^4 |\hat{u}(\xi)|^2 d\xi \\ &\geq \frac{1}{2} \int_{\mathbb{R}^n} (1 + |\xi|^2)^2 |\hat{u}(\xi)|^2 d\xi - \int_{\mathbb{R}^n} |\hat{u}(\xi)|^2 d\xi \\ &\geq \frac{1}{2} \|u\|_{2,2}^2 - \|u\|_2^2; \qquad \text{(Theorem 3.5.6)} \end{aligned}$$

and since $C_0^{\infty}(\Omega)$ is dense in $W_0^2(\Omega)$, we obtain Gårding's inequality

$$\mathfrak{F}(u,u) \geq \frac{1}{2} \|u\|_{2,2}^2 - \|u\|_2^2 \quad \text{for} \quad u \in W_0^2(\Omega).$$

Let A be the operator associated with this $\mathfrak{F}$. A is self-adjoint and $\sigma(A) \subset [0, \infty)$ as before. If Ω has finite thickness d in the direction of one coordinate axis, then two

applications of Hörmander's Lemma 2.5.1 give that $\mathfrak{F}(u,u) \geq d^{-4}\|u\|_2^2$ for $u \in C_0^\infty(\Omega)$ and hence for all $u \in W_0^2(\Omega)$; therefore $\sigma(A) \subset [d^{-4}, \infty)$. Corollary 3.6.17 implies that A^{-1} is compact when Ω is bounded.

If the boundary condition $u = \partial u/\partial n = 0$, of the biharmonic operator Δ^2, is replaced by $u = \Delta u = 0$ on $\partial\Omega$, then its natural proper definition in $\mathcal{H} = L^2(\Omega)$ becomes A^2 where A is given by (3.60). Note that if $\mathcal{V} = \mathcal{D}(A)$ and if for $u, v \in \mathcal{V}$

$$[u,v] = (Au, Av) + (u,v), \quad |u| = [u,u]^{1/2}$$

$$\mathfrak{F}(u,v) = (Au, Av),$$

then A^2 is the operator associated with this $\mathfrak{F}$. See also Exercise 19 of Chapter 2. Observe also that it can happen that $\mathcal{V} \not\subset W^2(\Omega)$ in this case (Exercise 13).

The general **strongly elliptic operator** of order $2m$ is given by

$$Lu = \sum_{|\alpha|\leq m} \sum_{|\beta|\leq m} (-1)^{|\beta|} D^\beta (a_{\alpha\beta} D^\alpha u)$$

provided there exists $\delta > 0$ such that

$$\operatorname{Re} \sum_{|\alpha|=m} \sum_{|\beta|=m} a_{\alpha\beta}(x)\xi^\alpha \xi^\beta \geq \delta|\xi|^{2m} \quad \text{for all } \xi \in \mathbb{R}^n, x \in \Omega. \tag{3.65}$$

Define

$$\mathfrak{F}(u,v) = \sum_{|\alpha|\leq m} \sum_{|\beta|\leq m} \int_\Omega a_{\alpha\beta} \mathsf{D}^\alpha u \, \overline{\mathsf{D}^\beta v} \quad \text{for} \quad u, v \in W^m(\Omega). \tag{3.66}$$

It is easy to see that $|\mathfrak{F}(u,v)| \leq M_2 \|u\|_{m,2}\|v\|_{m,2}$ for $u, v \in W^m(\Omega)$ provided only that $a_{\alpha\beta}$ are bounded; however, proof of Gårding's inequality can get rather complicated when $m > 1$. When assumptions of the following Theorem 3.7.1 are satisfied, assumptions **H**1, **H**2, **H**3 of the section on Sectorial Forms are again satisfied, with $\mathcal{H} = L^2(\Omega)$ and $\mathcal{V} = W_0^m(\Omega)$. Moreover, Corollary 3.6.17 implies compactness of the identity map from $\mathcal{V} \to \mathcal{H}$. Hence, the operator associated with $\mathfrak{F}$ has compact resolvent and the Fredholm alternative applies as above. For regularity of solutions see Theorem 3.8.2.

Theorem 3.7.1 *Suppose that Ω is bounded and $\mathfrak{F}$ is given by (3.66), where*

(a) $m \geq 1$ and $a_{\alpha\beta} \in L^\infty(\Omega)$ for all $|\alpha| \leq m$, $|\beta| \leq m$

(b) $a_{\alpha\beta} \in C(\overline{\Omega})$ when $|\alpha| = |\beta| = m$

(c) strong ellipticity condition (3.65) is assumed to hold for some $\delta > 0$.

Then for every $M_3 \in (0, \delta)$, there exists $a \in \mathbb{R}$ such that

$$\operatorname{Re}\mathfrak{F}(u,u) \geq M_3\|u\|_{m,2}^2 + a\|u\|_2^2 \quad \textit{for all} \quad u \in W_0^m(\Omega). \tag{3.67}$$

PROOF Choose $u \in C_0^\infty(\Omega)$. We shall use $R_1, R_2, \ldots$ to denote various terms of $\mathfrak{F}(u,u)$ which can be estimated as

$$|R_i| \le c_i \|u\|_{m-1,2} \|u\|_{m,2},$$

where $c_i < \infty$ do not depend on u. Note that

$$\mathfrak{F}(u,u) = \sum_{|\alpha|=m} \sum_{|\beta|=m} \int_\Omega a_{\alpha\beta} D^\alpha u \, \overline{D^\beta u} + R_1.$$

Define $\theta = (\delta - M_3)/3$, $\varepsilon = \theta(\sum_{|\alpha|=m} 1)^{-1}$ and let $\mu > 0$ be such that

$$|a_{\alpha\beta}(x) - a_{\alpha\beta}(y)| < \varepsilon \quad \text{if } |x-y| < 2\mu,\ x, y \in \overline{\Omega},\ |\alpha| = |\beta| = m.$$

Note that $\overline{\Omega} \subset B(x_1,\mu) \cup \cdots \cup B(x_k,\mu)$ for some $x_j \in \overline{\Omega}$. Choose real valued $\phi_j \in C_0^\infty(\mathbb{R}^n)$ such that $\operatorname{supp}(\phi_j) \subset B(x_j, 2\mu)$ and $\phi_j(x) > 0$ for $x \in B(x_j,\mu)$, $j = 1, \ldots, k$. Note that $\phi \equiv \phi_1^2 + \cdots + \phi_k^2 > 0$ in $\overline{\Omega}$. Hence,

$$\psi_j \equiv \phi_j \phi^{-1/2} \in C_B^m(\Omega) \cap C^\infty(\Omega)$$

and $v_j \equiv u\psi_j \in C_0^\infty(\Omega)$ for $j = 1, \ldots, k$. Since $\psi_1^2 + \cdots + \psi_k^2 = 1$ in $\overline{\Omega}$ and

$$\psi_j D^\alpha u = D^\alpha v_j - \sum_{\beta<\alpha} \binom{\alpha}{\beta} (D^\beta u) D^{\alpha-\beta} \psi_j,$$

it follows that

$$\begin{aligned}
\mathfrak{F}(u,u) - R_1 &= \sum_{j=1}^k \sum_{|\alpha|=m} \sum_{|\beta|=m} \int_\Omega a_{\alpha\beta} \psi_j^2 D^\alpha u \, \overline{D^\beta u} \\
&= \sum_{j=1}^k \sum_{|\alpha|=m} \sum_{|\beta|=m} \int_\Omega a_{\alpha\beta} D^\alpha v_j \, \overline{D^\beta v_j} + R_2 \\
\mathfrak{F}(u,u) - R_1 - R_2 &= \sum_{j=1}^k \sum_{|\alpha|=m} \sum_{|\beta|=m} a_{\alpha\beta}(x_j) \int_\Omega D^\alpha v_j \, \overline{D^\beta v_j} + G,
\end{aligned}$$

where

$$\begin{aligned}
G &= \sum_{j=1}^k \sum_{|\alpha|=m} \sum_{|\beta|=m} \int_{B(x_j,2\mu)\cap\Omega} (a_{\alpha\beta} - a_{\alpha\beta}(x_j)) D^\alpha v_j \, \overline{D^\beta v_j} \\
|G| &\le \varepsilon \sum_{j=1}^k \sum_{|\alpha|=m} \sum_{|\beta|=m} \int_\Omega |D^\alpha v_j| \, |D^\beta v_j|
\end{aligned}$$

$$\leq \varepsilon \sum_{j=1}^{k} \sum_{|\alpha|=m} \sum_{|\beta|=m} \int_{\Omega} |D^{\alpha} v_j|^2$$
$$= \theta \sum_{j=1}^{k} \sum_{|\alpha|=m} \int_{\Omega} |D^{\alpha} v_j|^2.$$

Let $v_j = 0$ on $\mathbb{R}^n \backslash \Omega$ and note that

$$\mathfrak{F}(u,u) - R_1 - R_2 - G = \sum_{j=1}^{k} \int_{\mathbb{R}^n} \sum_{|\alpha|=m} \sum_{|\beta|=m} a_{\alpha\beta}(x_j) \xi^{\alpha} \xi^{\beta} |\hat{v}_j(\xi)|^2 d\xi$$

$$\begin{aligned}
\operatorname{Re}\left(\mathfrak{F}(u,u) - R_1 - R_2 - G\right) &\geq \delta \sum_{j=1}^{k} \int_{\mathbb{R}^n} |\xi|^{2m} |\hat{v}_j(\xi)|^2 d\xi \\
&\geq \delta \sum_{j=1}^{k} \int_{\mathbb{R}^n} \sum_{|\alpha|=m} |\xi^{\alpha} \hat{v}_j(\xi)|^2 d\xi \\
&= \delta \sum_{j=1}^{k} \int_{\Omega} \sum_{|\alpha|=m} |D^{\alpha} v_j|^2 \\
\operatorname{Re}\left(\mathfrak{F}(u,u) - R_1 - R_2\right) &\geq (\delta - \theta) \sum_{j=1}^{k} \int_{\Omega} \sum_{|\alpha|=m} |D^{\alpha} v_j|^2 \\
&= (\delta - \theta) \sum_{j=1}^{k} \int_{\Omega} \sum_{|\alpha|=m} \psi_j^2 |D^{\alpha} u|^2 + R_3 \\
&= (\delta - \theta) \int_{\Omega} \sum_{|\alpha|=m} |D^{\alpha} u|^2 + R_3 \\
&= (\delta - \theta)(\|u\|_{m,2}^2 - \|u\|_{m-1,2}^2) + R_3.
\end{aligned}$$

Therefore

$$\begin{aligned}
\operatorname{Re} \mathfrak{F}(u,u) &\geq (\delta - \theta)(\|u\|_{m,2}^2 - \|u\|_{m-1,2}^2) - c_4 \|u\|_{m-1,2} \|u\|_{m,2} \\
&\geq (\delta - 2\theta) \|u\|_{m,2}^2 - c_5 \|u\|_{m-1,2}^2.
\end{aligned}$$

Since $x^{m-1} \leq \tau x^m + \tau^{1-m}$ for all $x > 0, \tau > 0$, Theorem 3.5.6 implies that

$$\|u\|_{m-1,2}^2 \leq \tau c^2 \|u\|_{m,2}^2 + \tau^{1-m} \|u\|_2^2 \quad \text{for} \quad \tau > 0,$$

and taking τ, such that $\tau c_5 c^2 < \theta$, implies (3.67). □

3.8 Regularity of Weak Solutions

In this section it will be assumed that m is a positive integer, $a_{\alpha\beta} \in C(\Omega)$ for all multi-indices α, β with $|\alpha| \leq m$ and $|\beta| \leq m$, and that

$$\text{Re} \sum_{|\alpha|=m} \sum_{|\beta|=m} a_{\alpha\beta}(x)\xi^{\alpha+\beta} > 0 \quad \text{for} \quad x \in \Omega,\ \xi \in \mathbb{R}^n \backslash \{0\}. \tag{3.68}$$

Observe that this implies that for each compact $K \subset \Omega$ there exists $\delta > 0$ such that

$$\text{Re} \sum_{|\alpha|=m} \sum_{|\beta|=m} a_{\alpha\beta}(x)\xi^{\alpha+\beta} \geq \delta|\xi|^{2m} \quad \text{for} \quad x \in K,\ \xi \in \mathbb{R}^n.$$

For $u \in W^m_{loc}(\Omega)$, $\phi \in C_0^\infty(\Omega)$, define

$$\mathfrak{F}(u, \phi) = \sum_{|\alpha|\leq m} \sum_{|\beta|\leq m} \int_\Omega a_{\alpha\beta} \mathrm{D}^\alpha u\, D^\beta \phi. \tag{3.69}$$

Lemma 3.8.1 *Suppose that $1 \leq j \leq m$, $a_{\alpha\beta} \in C^j(\Omega)$ for $|\alpha| \leq m$, $|\beta| \leq m$ and that $u \in W^m_{loc}(\Omega)$ is such that for each compact set K there exists $C_K < \infty$ such that*

$$|\mathfrak{F}(u, \phi)| \leq C_K \|\phi\|_{m-j,2} \quad \textit{for all} \quad \phi \in C_0^\infty(\Omega) \textit{ with } \text{supp}(\phi) \subset K.$$

Then $u \in W^{m+j}_{loc}(\Omega)$.

PROOF Assume first that $j = 1$. Let K be any nonempty compact subset of Ω and let $d > 0$ be such that $\overline{\mathcal{O}}_4 \subset \Omega$, where $\mathcal{O}_i = K + B(0, id)$, $1 \leq i \leq 4$.

Let $f \in C_0^\infty(\mathbb{R}^n)$ be such that $f(x) = 1$ for $x \in \mathcal{O}_1$ and $f(x) = 0$ for $x \notin \mathcal{O}_2$, see Theorem 3.1.2. Define $v = fu$. Using notation $v^y(x) = v(x + y)$, note that $v^y \in W^m_0(\mathcal{O}_3)$ for all $|y| < d$.

Pick any $1 \leq i \leq n$, $-d < h < d$, $h \neq 0$ and $\phi \in C_0^\infty(\mathcal{O}_3)$. Observe that $\delta_i^h v \in W^m_0(\mathcal{O}_3)$, where δ_i^h is defined by (3.38). Let $R_1, R_2, \ldots$ denote various terms of $\mathfrak{F}(\delta_i^h v, \phi)$ that can be bounded by $c_k \|\phi\|_{m,2}$ for some c_k that do not depend on h and ϕ. Note that

$$\mathfrak{F}(\delta_i^h v, \phi) = \sum_{|\alpha|\leq m} \sum_{|\beta|\leq m} \int_{\mathcal{O}_3} a_{\alpha\beta} (\delta_i^h \mathrm{D}^\alpha (fu))\, D^\beta \phi.$$

Theorems 3.4.10 and 3.4.6 imply that

$$\mathfrak{F}(\delta_i^h v, \phi) = \sum_{|\alpha|\leq m} \sum_{|\beta|\leq m} \int_{\mathcal{O}_3} a_{\alpha\beta} (\delta_i^h (f \mathrm{D}^\alpha u))\, D^\beta \phi + R_1.$$

Since $\delta_i^h(wg) = g\delta_i^h w + w^{he_i}\delta_i^h g$, we have

$$\begin{aligned}
\mathfrak{F}(\delta_i^h v, \phi) &= \sum_{|\alpha|\le m}\sum_{|\beta|\le m}\int_{\mathcal{O}_3}(\delta_i^h(a_{\alpha\beta} f \mathsf{D}^\alpha u) - (f\mathsf{D}^\alpha u)^{he_i}\delta_i^h a_{\alpha\beta})D^\beta\phi + R_1\\
&= \sum_{|\alpha|\le m}\sum_{|\beta|\le m}\int_{\mathcal{O}_3}\delta_i^h(a_{\alpha\beta} f \mathsf{D}^\alpha u)\, D^\beta\phi + R_2\\
&= -\sum_{|\alpha|\le m}\sum_{|\beta|\le m}\int_{\mathcal{O}_2} a_{\alpha\beta} f \mathsf{D}^\alpha u\, \delta_i^{-h} D^\beta\phi + R_2\\
&= -\sum_{|\alpha|\le m}\sum_{|\beta|\le m}\int_{\mathcal{O}_2} a_{\alpha\beta}\mathsf{D}^\alpha u\, f D^\beta\delta_i^{-h}\phi + R_2\\
&= -\sum_{|\alpha|\le m}\sum_{|\beta|\le m}\int_{\mathcal{O}_2} a_{\alpha\beta}\mathsf{D}^\alpha u\, D^\beta(f\delta_i^{-h}\phi) + R_3\\
&= -\mathfrak{F}(u, f\delta_i^{-h}\phi) + R_3\\
|\mathfrak{F}(\delta_i^h v, \phi)| &\le C_{\overline{\mathcal{O}}_2}\|f\delta_i^{-h}\phi\|_{m-1,2} + c_3\|\phi\|_{m,2} \le c_4\|\phi\|_{m,2},
\end{aligned}$$

and since $C_0^\infty(\mathcal{O}_3)$ is dense in $W_0^m(\mathcal{O}_3)$,

$$|\mathfrak{F}(\delta_i^h v, \overline{\delta_i^h v})| \le c_4\|\delta_i^h v\|_{m,2}$$

and Theorem 3.7.1 implies that for some $M_3 > 0$, $a \in \mathbb{R}$ we have

$$a\|\delta_i^h v\|_2^2 + M_3\|\delta_i^h v\|_{m,2}^2 \le c_4\|\delta_i^h v\|_{m,2};$$

hence, by Theorem 3.4.6, $M_3\|\delta_i^h v\|_{m,2}^2 \le c_4\|\delta_i^h v\|_{m,2} + |a|\|\mathsf{D}_i v\|_2^2$, which implies that $\|\delta_i^h v\|_{m,2}$ remains bounded as $h \to 0$ and, since

$$\int_K |\delta_i^h \mathsf{D}^\alpha u|^2 = \int_K |\delta_i^h \mathsf{D}^\alpha v|^2 \le \|\delta_i^h v\|_{m,2}^2 \quad \text{for} \quad |\alpha| \le m,$$

Theorem 3.4.7 implies that $\mathsf{D}_i\mathsf{D}^\alpha u$ exists and $\mathsf{D}_i\mathsf{D}^\alpha u \in L^2_{loc}(\Omega)$ for all $|\alpha| \le m$. This completes the proof in the case $j = 1$.

Assume now that $2 \le j \le m$ and that the assertion of the Lemma is true for $j-1$. Hence we may assume that $u \in W_{loc}^{m+j-1}(\Omega)$. Choose any compact set $K \subset \Omega$, $\phi \in C_0^\infty(\Omega)$ with $\operatorname{supp}(\phi) \subset K$ and $1 \le i \le n$. Let $c_1, c_2, \ldots$ denote constants that do not depend on ϕ. Note that

$$\begin{aligned}
\mathfrak{F}(\mathsf{D}_i u, \phi) &= \sum_{|\alpha|\le m}\sum_{|\beta|\le m}\int_\Omega a_{\alpha\beta}(\mathsf{D}_i\mathsf{D}^\alpha u)D^\beta\phi\\
&= -\mathfrak{F}(u, D_i\phi) - \sum_{|\alpha|\le m}\sum_{|\beta|\le m} E_{\alpha\beta},
\end{aligned}$$

where

$$E_{\alpha\beta} = \int_\Omega D_i a_{\alpha\beta}\, \mathsf{D}^\alpha u\, D^\beta \phi.$$

If $|\beta| \le m - j + 1$, then obviously $|E_{\alpha\beta}| \le c_1 \|\phi\|_{m-j+1,2}$. If $|\beta| > m - j + 1$, then $\beta = \gamma + \mu$ where $|\gamma| = m - j + 1$ and $|\mu| \le j - 1$; hence

$$E_{\alpha\beta} = (-1)^{|\mu|} \int_\Omega \mathsf{D}^\mu (D_i a_{\alpha\beta}\, \mathsf{D}^\alpha u) D^\gamma \phi$$

and $|E_{\alpha\beta}| \le c_2 \|\phi\|_{m-j+1,2}$ also in this case. Therefore

$$|\mathfrak{F}(\mathsf{D}_i u, \phi)| \le C_K \|D_i \phi\|_{m-j,2} + c_3 \|\phi\|_{m-j+1,2} \le c_4 \|\phi\|_{m-j+1,2}$$

and the induction assumption implies that $\mathsf{D}_i u \in W^{m+j-1}_{loc}(\Omega)$. □

Theorem 3.8.2 *Suppose that $m \ge 1$, $k \ge 0$, $a_{\alpha\beta} \in C^{m+k}(\Omega)$ for $|\alpha| \le m$ and $|\beta| \le m$ satisfy (3.68), and that $u \in W^m_{loc}(\Omega)$ and $f \in W^k_{loc}(\Omega)$ are such that*

$$\sum_{|\alpha|\le m} \sum_{|\beta|\le m} \int_\Omega a_{\alpha\beta} \mathsf{D}^\alpha u\, D^\beta \phi = \int_\Omega f\phi \quad \textit{for all} \quad \phi \in C^\infty_0(\Omega).$$

Then $u \in W^{2m+k}_{loc}(\Omega)$ and

$$\sum_{|\alpha|\le m} \sum_{|\beta|\le m} (-1)^{|\beta|} \mathsf{D}^\beta (a_{\alpha\beta} \mathsf{D}^\alpha u) = f.$$

PROOF Lemma 3.8.1 implies the assertion when $k = 0$. Assume that $k \ge 1$ and that the assertion is true for $k - 1$. Note that $u \in W^{2m+k-1}_{loc}(\Omega)$ and that for every $1 \le i \le n$, $\phi \in C^\infty_0(\Omega)$, we have

$$\sum_{|\alpha|\le m} \sum_{|\beta|\le m} \int_\Omega a_{\alpha\beta} \mathsf{D}^\alpha u\, D^\beta D_i \phi = -\int_\Omega \phi \mathsf{D}_i f$$

$$\begin{aligned} \sum_{|\alpha|\le m} \sum_{|\beta|\le m} \int_\Omega a_{\alpha\beta} \mathsf{D}^\alpha \mathsf{D}_i u\, D^\beta \phi &= \int_\Omega \phi \mathsf{D}_i f - \sum_{|\alpha|\le m} \sum_{|\beta|\le m} \int_\Omega D_i a_{\alpha\beta}\, \mathsf{D}^\alpha u\, D^\beta \phi \\ &= \int_\Omega \phi g, \end{aligned}$$

where

$$g = \mathsf{D}_i f - \sum_{|\alpha|\le m} \sum_{|\beta|\le m} (-1)^{|\beta|} \mathsf{D}^\beta (D_i a_{\alpha\beta}\, \mathsf{D}^\alpha u)$$

and, since $g \in W^{k-1}_{loc}(\Omega)$, we have that $\mathsf{D}_i u \in W^{2m+k-1}_{loc}(\Omega)$. □

Theorem 3.8.3 *Suppose that* $m \geq 1$, $k \geq 0$, $b_\alpha \in C^{2m+k}(\Omega)$ *for* $|\alpha| \leq 2m$ *and*

$$\text{Re } \zeta \sum_{|\alpha|=2m} b_\alpha(x)\xi^\alpha > 0 \quad \textit{for all} \quad x \in \Omega,\ \xi \in \mathbb{R}^n$$

for some $\zeta \in \mathbb{C}$. *Suppose also that* $u \in L^2_{loc}(\Omega)$ *and* $f \in W^k_{loc}(\Omega)$ *are such that*

$$\int_\Omega u \sum_{|\alpha|\leq 2m} b_\alpha D^\alpha \phi = \int_\Omega f\phi \quad \textit{for all} \quad \phi \in C_0^\infty(\Omega).$$

Then $u \in W^{2m+k}_{loc}(\Omega)$ *and* $\sum_{|\alpha|\leq 2m}(-1)^{|\alpha|}\mathsf{D}^\alpha(b_\alpha u) = f$.

PROOF Pick any nonempty compact subset K of Ω and let $d > 0$ be such that $\overline{\mathcal{O}} \subset \Omega$ where $\mathcal{O} = K + B(0,d)$.

$w \to (w,u) = \int_{\mathcal{O}} w\overline{u}$ is a bounded linear functional on $W^m(\mathcal{O})$; hence the Riesz Lemma 2.2.4 implies that there exists $v \in W^m(\mathcal{O})$ such that $(w,u) = (w,v)_m$ for all $w \in W^m(\mathcal{O})$ and, in particular,

$$\sum_{|\alpha|\leq m} \int_{\mathcal{O}} \mathsf{D}^\alpha v\, D^\alpha \phi = \int_{\mathcal{O}} u\phi \quad \text{for all} \quad \phi \in C_0^\infty(\mathcal{O}).$$

Theorem 3.8.2 implies that $v \in W^{2m}_{loc}(\mathcal{O})$ and $u = \sum_{|\alpha|\leq m}(-1)^{|\alpha|}\mathsf{D}^{2\alpha}v$. Hence

$$\sum_{|\beta|\leq 2m}\sum_{|\alpha|\leq m} \int_{\mathcal{O}} (-1)^{|\alpha|+m}\zeta b_\beta\, \mathsf{D}^{2\alpha}v\, D^\beta\phi = \int_{\mathcal{O}} (-1)^m \zeta f\,\phi \quad \text{for} \quad \phi \in C_0^\infty(\mathcal{O})$$

and Theorem 3.8.2 implies that $v \in W^{4m+k}_{loc}(\mathcal{O})$; hence $u \in W^{2m+k}_{loc}(\mathcal{O})$. □

EXAMPLE 3.8.4 If $f \in W^k_{loc}(\Omega)$, $k \geq 0$ and $u \in L^2_{loc}(\Omega)$ are such that

$$\int_\Omega u\Delta\varphi = \int_\Omega f\varphi \quad \text{for all} \quad \varphi \in C_0^\infty(\Omega),$$

then Theorem 3.8.3 implies that $u \in W^{2+k}_{loc}(\Omega)$ and $\Delta u = f$. In particular, if $f \in C^\infty(\Omega)$, then $u \in \mathbb{C}^\infty(\Omega)$ by Corollary 3.6.4.

3.9 Exercises

1. Supply the details for the following alternative proof of completeness of normalized Hermite functions in $L^2(\mathbb{R})$. If $(f,\phi_n) = 0$ for all Hermite functions ϕ_n, then the Fourier transform of the function $f(x)e^{-x^2/2}$ is 0 and hence $f = 0$.

2. The electric and magnetic fields $E(x,t), B(x,t) \in \mathbb{C}^3$ satisfy Maxwell's equations,

$$B_t = -\nabla \times E, \quad E_t = \nabla \times B,$$

in $\mathbb{R}^3$ for $t \geq 0$. Solve the equations for given $E(\cdot,0), B(\cdot,0) \in \mathcal{S}^3$.

3. Derive (3.11) from (3.6).

4. Prove (3.16).

5. Show that if $f \in \mathcal{S}$ and $t \in \mathbb{R}$, then $g, h \in \mathcal{S}$, where

$$g(x) = f(x)\cos|x|t, \quad h(x) = f(x)\frac{\sin|x|t}{|x|}.$$

6. Suppose that $1 \leq p < \infty$ and $-\infty < a < b < \infty$. Show that $f \in W_0^{1,p}(a,b)$ iff $f \in AC[a,b]$, $f' \in L^p(a,b)$ and $f(a) = f(b) = 0$.

7. Suppose $u \in L^1_{loc}(\Omega)$, $\{f_1, f_2, \ldots\} \subset L^1_{loc}(\Omega)$, α is a multi-index, $\mathsf{D}^\alpha f_i$ exists for $i \geq 1$, $\mathrm{c} \equiv \sup_i \|\mathsf{D}^\alpha f_i\|_p < \infty$ for some $p \in (1,\infty]$ and

$$\lim_{i\to\infty} \int_\Omega (u - f_i)\varphi = 0 \quad \text{for all} \quad \varphi \in C_0^\infty(\Omega).$$

Show that $\mathsf{D}^\alpha u$ exists and $\|\mathsf{D}^\alpha u\|_p \leq \mathrm{c}$. Show also that the restriction $p > 1$ is necessary.

8. Observe that ϕ, in the proof of Theorem 3.5.5, is a sum of $C_0^\infty(\Omega)$ functions. Is $\phi \in W_0^{m,p}(\Omega)$?

9. Is it true that $W_0^{m,p}(\mathbb{R}^n) = W^{m,p}(\mathbb{R}^n)$ also when $m \geq 0$ and $p = \infty$?

10. Show (3.51).

11. Define $u(x) = \log(\log(4/|x|))$ for $x \in B(0,1) \subset \mathbb{R}^n$. Show that

$$u \in W^{m,p}(B(0,1)) \backslash L^\infty(B(0,1))$$

when $p > 1$ and $mp = n$. (See Lemma 3.6.2.)

12. Suppose $m \in \mathbb{N}$, $p \in [1,\infty]$ and $0 < m - \frac{n}{p} < \mu < \sigma < 1$. Define $u(x) = |x|^\mu$ for $x \in B(0,1) \subset \mathbb{R}^n$. Show that

$$u \in W^{m,p}(B(0,1)) \backslash C^\sigma(B(0,1),\mathbb{C}).$$

(See Theorem 3.6.12.)

13. Let $\Omega \subset \mathbb{R}^2$ be given in polar coordinates (r, ϕ) by $0 < r < 1$, $0 < \phi < \pi/\beta$ where $1/2 < \beta < 1$. Define $u = (1-r)r^\beta \sin\beta\phi$ in Ω. Show that $u \in \mathcal{D}(A)$, where A is given by (3.60). Also show that $u \notin W^2(\Omega)$.

14. Let Ω be a nonempty, bounded and connected set in $\mathbb{R}^n$ with the cone property. Apply Theorem 2.6.8 to the Neumann problem to show that there exists $c < \infty$ such that

$$\left\| f - \frac{1}{\mu(\Omega)} \int_\Omega f \right\|_2^2 \leq c \sum_{i=1}^n \|D_i f\|_2^2 \quad \text{for} \quad f \in W^1(\Omega).$$

What role does this c play in the Neumann problem? Note that this removes the star-shaped condition on Ω in the Ponicare inequality (3.43).

15. Show that Gårding's inequality (3.63) still holds for $\mathfrak{F}$ given by (3.61) when $\mathcal{V} = W_0^1(\Omega)$ and the condition $\operatorname{Im} a_{ij} = \operatorname{Im} a_{ji}$ is replaced by $\operatorname{Im}(a_{ij} - a_{ji}) \in W^{1,\infty}(\Omega)$.

Chapter 4

Semigroups of Linear Operators

4.1 Introduction

We shall study problems related to the abstract differential equation of the form

$$u' + Au = 0,$$

where A is a linear operator in a Banach space X and u' denotes the derivative of $u : [0, \infty) \to X$ in the Banach space, i.e.

$$\lim_{h \to 0} \left\| \frac{1}{h}(u(t+h) - u(t)) - u'(t) \right\| = 0.$$

Assuming that for given $u(0)$ the equation has a unique solution on $[0, \infty)$ implies that there exist linear operators $Q(t)$, for $t \geq 0$, such that

$$u(t) = Q(t)u(0)$$

and $Q(t)Q(s) = Q(t+s)$. Formally, $Q(t) = e^{-At}$. Assuming continuous dependence on initial conditions gives us that $Q(t)$ should be a bounded linear map from its domain into X. In applications it can be usually arranged so that the domain of $Q(t)$ is the whole X. Clearly, the map $t \to Q(t)$ has to have some continuity properties. This, and some technical considerations, suggest the following definitions.

Definition 4.1.1 A family of bounded linear operators $\{Q(t)\}_{t \geq 0}$ on a Banach space X is called a **strongly continuous semigroup** (or C_0 **semigroup**) if

(a) $Q(0) = 1$ (identity map)

(b) $Q(t)Q(s) = Q(t+s)$ for all $t \geq 0, s \geq 0$

(c) $\lim_{t \to 0+} \|Q(t)x - x\| = 0$ for all $x \in X$.

Definition 4.1.2 Let $\{Q(t)\}_{t\geq 0}$ be a strongly continuous semigroup of operators on a Banach space X. Define $\mathcal{D}(A)$ to be the set of all $x \in X$ for which there exists $y \in X$ such that

$$\lim_{t\to 0+} \left\| \frac{1}{t}(x - Q(t)x) - y \right\| = 0;$$

for such x and y define $Ax = y$. The linear operator, $-A$, is called the **generator** (or the **infinitesimal generator**) of the semigroup.

EXAMPLE 4.1.3 Let the Banach space X be $C_u(\mathbb{R})$. For $t \geq 0$ define $Q(t) \in \mathfrak{B}(X)$ by

$$(Q(t)f)(x) = f(x - t) \quad \text{for} \quad f \in X,\ x \in \mathbb{R}.$$

One can easily verify that $\{Q(t)\}_{t\geq 0}$ is a strongly continuous (but not continuous, Exercise 1) semigroup of operators on X. Let $-A$ be its generator. If $f \in \mathcal{D}(A)$ and

$$e(t) = \frac{f - Q(t)f}{t} - Af \quad \text{for} \quad t > 0,$$

then $\lim_{t\to 0+} \|e(t)\| = 0$ and, since for all $x \in \mathbb{R}$,

$$\left| \frac{f(x-t) - f(x)}{-t} - (Af)(x) \right| \leq \|e(t)\|, \quad \left| \frac{f(x+t) - f(x)}{t} - (Af)(x+t) \right| \leq \|e(t)\|,$$

we have that $Af = f'$ and $f \in C_u^1(\mathbb{R})$. If $f \in C_u^1(\mathbb{R})$, then

$$\left(\frac{f - Q(t)f}{t} - f' \right)(x) = \int_0^1 (f'(x - st) - f'(x))ds \xrightarrow{t\to 0} 0$$

uniformly in x, which implies that $f \in \mathcal{D}(A)$. Thus

$$Af = f' \quad \text{for} \quad f \in \mathcal{D}(A) = C_u^1(\mathbb{R}).$$

If $f \in \mathcal{D}(A)$ and $u(t) = Q(t)f$, then $u(t) \in \mathcal{D}(A)$ and

$$\left\| \frac{u(t+h) - u(t)}{h} + Au(t) \right\| = \sup_{x\in\mathbb{R}} \left| \frac{f(x-h) - f(x)}{h} + f'(x) \right| \xrightarrow{h\to 0} 0;$$

hence

$$u' + Au = 0 \quad \text{on} \quad [0, \infty),$$

where u' denotes the derivative of u in the Banach space X. Note also that if $v(x,t) = (u(t))(x) = f(x - t)$, then

$$v_t(x,t) + v_x(x,t) = 0 \quad \text{for} \quad t \geq 0,\ x \in \mathbb{R},$$

where v_t, v_x denote the classical partial derivatives.

EXAMPLE 4.1.4 Suppose X is a complex Banach space and $B \in \mathfrak{B}(X)$. Define

$$Q(\zeta) = \sum_{k=0}^{\infty} \frac{(-\zeta)^k}{k!} B^k \quad \text{for} \quad \zeta \in \mathbb{C}.$$

It can be easily verified that

(a) $Q(\zeta) \in \mathfrak{B}(X)$, $\|Q(\zeta)\| \leq \exp(|\zeta|\|B\|)$ for $\zeta \in \mathbb{C}$

(b) $Q(t)Q(s) = Q(t+s)$ for $t, s \in \mathbb{C}$

(c) $\|\zeta^{-1}(1 - Q(\zeta)) - B\| \leq |\zeta|\|B\|^2 \exp(|\zeta|\|B\|)$ for $\zeta \in \mathbb{C}, \zeta \neq 0$.

This implies that if $x \in X$ and $u(\zeta) = Q(\zeta)x$ for $\zeta \in \mathbb{C}$, then u is analytic and $u' + Bu = 0$. In particular, $\{Q(t)\}_{t\geq 0}$ is a strongly continuous semigroup of operators on X and its generator is $-B$.

EXAMPLE 4.1.5 Suppose $X = L^p(\mathbb{R}^n)$ with $p \in [1, \infty)$, $n \geq 1$. For $t > 0$ define

$$(Q(t)f)(x) = (4\pi t)^{-n/2} \int_{\mathbb{R}^n} \mathrm{e}^{-|x-y|^2/(4t)} f(y)dy \quad \text{for} \quad f \in X,\ x \in \mathbb{R}^n$$

and let $Q(0)$ be the identity map on X. In Example 3.2.13 it is shown that $\|Q(t)\| \leq 1$ for $t \geq 0$ and that (c) of Definition 4.1.1 holds. The Fubini Theorem implies that $Q(t+s) = Q(t)Q(s)$ for $t, s \geq 0$. Therefore $\{Q(t)\}_{t\geq 0}$ is a strongly continuous semigroup of operators on X.

EXAMPLE 4.1.6 Let $\mathcal{H}$ be a Hilbert space, let $\{\varphi_1, \varphi_2, \ldots\}$ be a complete orthonormal set in $\mathcal{H}$ and let $\lambda_1, \lambda_2, \ldots$ be complex numbers such that $\inf_k \operatorname{Re} \lambda_k > -\infty$. For $f \in \mathcal{H}$, $t \geq 0$, define

$$Q(t)f = \sum_{k=1}^{\infty} \mathrm{e}^{-\lambda_k t}(f, \varphi_k)\varphi_k. \tag{4.1}$$

Lemma 2.1.4 and Theorem 2.1.6 imply that $\{Q(t)\}_{t\geq 0}$ is a strongly continuous semigroup of operators on $\mathcal{H}$ and that its generator is given by

$$Af = \sum_{k=1}^{\infty} \lambda_k(f, \varphi_k)\varphi_k \quad \text{for} \quad f \in \mathcal{D}(A) = \{f \in \mathcal{H} \mid \sum_{k=1}^{\infty} |\lambda_k(f, \varphi_k)|^2 < \infty\}. \tag{4.2}$$

One can easily verify that if $f \in \mathcal{D}(A)$, then $u(t) \equiv Q(t)f \in \mathcal{D}(A)$, $Au(t) = Q(t)Af$ and $u' + Au = 0$ for $t \geq 0$.

Observe that the point spectrum of A is equal to $\{\lambda_1, \lambda_2, \ldots\}$ and that $\sigma(A) = \overline{\sigma_p(A)}$ (Exercise 2). Note that in order for (4.1) to make sense for every $f \in \mathcal{H}$ and $t \geq 0$, we need to have that the real part of the spectrum of A is bounded from below.

Theorem 2.6.8 implies that symmetric operators with compact resolvent can be represented by (4.2). A could be $-\Delta$ in a bounded domain with various boundary conditions, see (3.60), for a proper formulation of A. If $\mathcal{H} = L^2(0,1)$, $Af = -f''$ for $f \equiv \mathcal{D}(A) = W^2(0,1) \cap W_0^1(0,1)$ we obtain the one dimensional version studied in Examples 2.8.1, 2.8.3 and 2.8.8. In this case, the eigenvectors of A are $\varphi_k(x) = \sqrt{2}\sin(k\pi x)$, the corresponding eigenvalues are $\lambda_k = (k\pi)^2$, (4.2) holds and, if $g \in \mathcal{H}$, $v(x,t) = (Q(t)g)(x)$, then it is easy to verify directly that

$$v_t(x,t) - v_{xx}(x,t) = 0 \quad \text{for} \quad t > 0, x \in [0,1]$$

$$v(0,t) = v(1,t) = 0 \quad \text{for} \quad t > 0$$

$$\lim_{t\to 0} \int_0^1 |v(x,t) - g(x)|^2 dx = 0.$$

In this example, $Q(t)g \in \cap_{n=1}^{\infty} \mathcal{D}(A^n) \subset C^\infty(0,1)$ for every $t > 0$ and every $g \in L^2(0,1)$ and hence $Q(t)$ is a smoothing operator in this case.

Let $\mathcal{H} = L^2(0,1)$ and $Af = f'$ with domain

$$\mathcal{D}(A) = \{f \in AC[0,1] \mid f' \in L^2(0,1),\ f(0) = f(1)\}.$$

This operator was studied in Examples 1.6.4 and 2.6.9. In particular, the eigenvectors of A are $\varphi_n(x) = e^{2in\pi x}$, the corresponding eigenvalues are $2in\pi$, $n = 0, \pm 1, \ldots$ and (4.2) holds. It is easy to see that, in this case, $(Q(t)f)(x) = f_p(x-t)$ where f_p is the periodic extension of f with period 1. There is no smoothing here.

EXAMPLE 4.1.7 Let $\mathcal{H}$ be a Hilbert space, let $\{\varphi_1, \varphi_2, \ldots\}$ be a complete orthonormal set in $\mathcal{H}$ and let $\lambda_1, \lambda_2, \ldots$ be complex numbers such that, for some $a \in (0, \infty)$, we have that $(\operatorname{Im} \lambda_k)^2 \leq 4a(a + \operatorname{Re} \lambda_k)$ for all $k \geq 1$. This condition implies that we can choose $\mu_k \in \mathbb{C}$ so that $\mu_k^2 = \lambda_k$ and $|\operatorname{Im} \mu_k| \leq a^{1/2}$. Define A as in (4.2).

Define a Hilbert space $\mathcal{V}$ with an inner product $[\cdot, \cdot]$ by

$$\mathcal{V} = \{f \in \mathcal{H} \mid \sum_{k=1}^{\infty} |\mu_k (f, \varphi_k)|^2 < \infty\}$$

$$[f, g] = \sum_{k=1}^{\infty} (1 + |\mu_k|^2)(f, \varphi_k)(\varphi_k, g) \quad \text{for} \quad f, g \in \mathcal{V}.$$

Define $X = \mathcal{V} \times \mathcal{H}$. X is a Hilbert space with an inner product

$$(\{f, g\}, \{u, v\}) = [f, u] + (g, v) \quad \text{for} \quad \{f, g\}, \{u, v\} \in X.$$

For $t \in [0, \infty)$, define $Q(t) \in \mathfrak{B}(X)$ as follows: if $\{f, g\} \in X$, then

$$Q(t)\{f, g\} = \{u(t), v(t)\} \quad \text{where}$$

$$u(t) = \sum_{k=1}^{\infty} (\,(f, \varphi_k) \cos(\mu_k t) + (g, \varphi_k) \sin(\mu_k t)/\mu_k)\varphi_k,$$

$$v(t) = \sum_{k=1}^{\infty} (\,(g, \varphi_k) \cos(\mu_k t) - (f, \varphi_k) \sin(\mu_k t)\mu_k)\varphi_k$$

(set $\sin(\mu_k t)/\mu_k = t$ if $\mu_k = 0$). $\{Q(t)\}_{t \geq 0}$ is a strongly continuous semigroup of operators on X. Its generator, denoted by $-S$, is given by

$$\{f, g\} \in \mathcal{D}(S) \text{ iff } f \in \mathcal{D}(A), g \in \mathcal{V}; \quad S\{f, g\} = \{-g, Af\}.$$

If $\{f, g\} \in \mathcal{D}(S)$ and $\{u(t), v(t)\} = Q(t)\{f, g\}$, then

$$u' = v,\ v' = -Au \text{ on } [0, \infty),\ u(0) = f,\ v(0) = g. \tag{4.3}$$

Backtracking, we see that the expressions for u and v make, in general, sense only when the sequence $|\operatorname{Im} \mu_1|, |\operatorname{Im} \mu_2|, \ldots$ is bounded, which implies that the spectrum of A has to lie within a parabola $(\operatorname{Im} \lambda)^2 \leq 4a(a + \operatorname{Re} \lambda)$ for some $a \in (0, \infty)$.

Concretely, if $\mathcal{H} = L^2(0,1)$, $Af = f''''$ for $f \in \mathcal{D}(A) = W^4(0,1) \cap W_0^2(0,1)$ and φ_k, λ_k are the eigenvectors and the eigenvalues of A, then system (4.3) is an abstract formulation of the PDE:

$$u_{tt}(x,t) + u_{xxxx}(x,t) = 0 \quad \text{for} \quad t \geq 0, x \in [0,1]$$

$$u(0,t) = u_x(0,t) = u(1,t) = u_x(1,t) = 0 \quad \text{for} \quad t \geq 0$$

$$u(x,0) = f(x),\ u_t(x,0) = g(x) \quad \text{for} \quad x \in [0,1].$$

In this case, it is not obvious what (reasonable) conditions on f and g in (4.3) would guarantee that the PDE is satisfied in the classical sense.

4.2 Bochner Integral

A Banach space setting of evolution equations requires taking the derivative in the Banach space. Hence, integration of Banach space valued functions is an important tool in this setting. We shall define the Bochner integral of such functions and derive its basic properties.

In the following, a subset of $\mathbb{R}^n$ is said to be measurable iff it is Lebesgue measurable. μ will denote the Lebesgue measure on $\mathbb{R}^n$. The functions will be defined on a nonempty measurable set $S \subset \mathbb{R}^n$, with ranges in a Banach space X.

$x : S \to X$ is called **weakly measurable** if $s \to \ell(x(s))$ is a Lebesgue measurable function for each $\ell \in X^\star$.

$x : S \to X$ is called **almost separably-valued** if there exists $\{y_1, y_2, \ldots\} \subset X$ such that $\inf_i \|x(s) - y_i\| = 0$ for almost all $s \in S$. When x is almost separably-valued, it is easy to see that $\{y_1, y_2, \ldots\}$ can be chosen to belong to the range of x.

$x : S \to X$ is called **strongly measurable** if it is weakly measurable and almost separably-valued. In the literature, a different definition is usually given which is then followed by the Pettis Theorem that proves their equivalence. Our approach is a bit simpler.

Theorem 4.2.1 *If $x : S \to X$ is continuous at almost every point in S, then x is strongly measurable.*

PROOF x is obviously weakly measurable. Let C be a measurable subset of S such that $\mu(C) = 0$ and x is continuous at each point of $S \backslash C$. If $S \backslash C$ is empty, then we are done. Otherwise, let $\{t_1, t_2, \ldots\}$ be a dense subset of $S \backslash C$ and observe that, by continuity, $\inf_i \|x(s) - x(t_i)\| = 0$ for each s in $S \backslash C$. □

Theorem 4.2.2 *If $x : S \to X$ is strongly measurable, then $s \to \|x(s)\|$ is measurable.*

PROOF Choose $\{y_1, y_2, \ldots\} \subset X$ and a measurable $C \subset S$ such that $\mu(C) = 0$ and $\inf_i \|x(s) - y_i\| = 0$ for all $s \in S \backslash C$. It is enough to prove that the set $V = \{t \in S \backslash C \mid \|x(t)\| \leq a\}$ is measurable for each $a \in (0, \infty)$.

Corollary 1.5.8 implies that we can choose $\ell_1, \ell_2, \ldots$ in $X^\star$ so that

$$\|\ell_i\| \leq 1, \ \ell_i(y_i) = \|y_i\|.$$

Let $V_i = \{t \in S\backslash C \mid |\ell_i(x(t))| \leq a\}$. Note that V_i are measurable sets and that $V \subset \cap_{i=1}^{\infty} V_i$. If $t \in \cap_{i=1}^{\infty} V_i$ and $\varepsilon > 0$, then $\|x(t) - y_j\| < \varepsilon/2$ for some j. Hence,

$$\begin{aligned} \|x(t)\| &< \varepsilon/2 + \|y_j\| = \varepsilon/2 + \ell_j(y_j) \\ &\leq \varepsilon/2 + |\ell_j(y_j - x(t))| + |\ell_j(x(t))| \\ &< \varepsilon + a; \end{aligned}$$

therefore $t \in V$ and $V = \cap_{i=1}^{\infty} V_i$ is measurable. □

$x : S \to X$ is said to be **Bochner integrable** if x is strongly measurable and the function $s \to \|x(s)\|$ is Lebesgue integrable. The set of all such functions x is a vector space and will be denoted by $L(S, X)$. The following Theorem 4.2.3 enables us to define the **Bochner integral** $\int_S x$ of $x \in L(S, X)$ to be $y \in X$ which satisfies (4.4).

Theorem 4.2.3 *If $x \in L(S, X)$, then there exists a unique $y \in X$ such that*

$$\ell(y) = \int_S \ell(x(s))\, ds \quad \textit{for all} \quad \ell \in X^\star. \tag{4.4}$$

Moreover, $\|y\| \leq \int_S \|x(s)\| ds$.

PROOF Assume first that there exists $y \in X$ such that (4.4) holds. Let $\ell \in X^\star$ be a normalized tangent functional to y, see Corollary 1.5.8, and note

$$\|y\|^2 = \ell(y) = \int_S \ell(x(s)) ds \leq \int_S \|y\| \|x(s)\| ds.$$

This shows the 'moreover' part as well as the uniqueness of y.

Choose $\{y_1, y_2, \ldots\} \subset X$ and a measurable $C \subset S$ such that $\mu(C) = 0$ and $\inf_i \|x(s) - y_i\| = 0$ for all $s \in S\backslash C$. For $i, m \geq 1$ define

$$A_{im} = \{t \in S\backslash C \mid \|x(t) - y_i\| \leq 1/m \text{ and } \|x(t) - y_i\| - \|x(t)\| < 0\}.$$

Theorem 4.2.2 implies that the sets A_{im} are measurable. Define

$$B_{1m} = A_{1m}, \ B_{im} = A_{im} \backslash \cup_{j=1}^{i-1} A_{jm} \text{ for } i \geq 2$$

$$x_m = \sum_{i=1}^{\infty} \chi_{B_{im}} y_i$$

and note that x_m is strongly measurable and $\|x(t)-x_m(t)\| \le \min\{1/m, \|x(t)\|\}$ for all $m \ge 1, t \in S\backslash C$. Since $\|x_m\| = \sum_{i=1}^{\infty} \chi_{B_{im}}\|y_i\| \le 2\|x\|$, we have that x_m is Bochner integrable and that $\sum_{i=1}^{\infty} \mu(B_{im})\|y_i\| \le 2\int_S \|x\|$, which enables us to define $z_m = \sum_{i=1}^{\infty} \mu(B_{im})y_i$ that satisfy

$$\ell(z_m) = \sum_{i=1}^{\infty} \mu(B_{im})\ell(y_i) = \int_S \ell(x_m(s))ds \quad \text{for all} \quad \ell \in X^\star,\ m \ge 1.$$

Since $\lim_{m\to\infty} \int_S \|x_m - x\| = 0$ by the DCT, the 'moreover' part implies

$$\|z_m - z_k\| \le \int_S \|x_m - x_k\| \le \int_S \|x_m - x\| + \int_S \|x - x_k\| \overset{m,k\to\infty}{\longrightarrow} 0,$$

which implies that z_m converge to some y. Therefore the DCT implies

$$\ell(y) = \lim_{m\to\infty} \ell(z_m) = \lim_{m\to\infty} \int_S \ell(x_m(s))ds = \int_S \ell(x(s))ds \quad \text{for all} \quad \ell \in X^\star.$$

□

Suppose $y_1, \ldots, y_k \in X$ and let $V_1, \ldots, V_k$ be disjoint measurable subsets of S. Define $x : S \to X$ by

$$x(t) = \sum_{i=1}^{k} \chi_{V_i}(t)\, y_i \quad \text{for} \quad t \in S.$$

Such functions x are called **simple functions**. Note that the range, $\mathcal{R}(x)$, is contained in $\{0, y_1, \ldots, y_k\}$. The proof of the above Theorem 4.2.3 also yields:

Corollary 4.2.4 *For each $x \in L(S,X)$ there exist simple functions $u_k \in L(S,X)$, $k \ge 1$, such that $\mathcal{R}(u_k) \subset \mathcal{R}(x) \cup \{0\}$, $\|u_k(t)\| \le 2\|x(t)\|$ for $k \ge 1$, $t \in S$,*

$$\lim_{k\to\infty} u_k(t) = x(t) \quad \textit{for almost all } t \in S$$

and $\lim_{k\to\infty} \int_S u_k = \int_S x$.

PROOF Let x_m, B_{im} be as in the proof of the above Theorem, with $y_i \in \mathcal{R}(x)$. For $j \ge 1$ define $v_j = \sum_{i=1}^{l} \chi_{B_{im}} y_i$, where m is such that $\int_S \|x - x_m\| < 1/j$ and l is such that $\sum_{i=l+1}^{\infty} \mu(B_{im})\|y_i\| < 1/j$. Note that $\int_S \|x - v_j\| < 2/j$; hence a subsequence $\{v_{j_k}\}$ satisfies $\lim_{k\to\infty} \|x(t) - v_{j_k}(t)\| = 0$ for almost all $t \in S$. Let $u_k = v_{j_k}$. □

This implies

Corollary 4.2.5 *If $x \in L(S,X)$ and $x(s) \in M$ for all $s \in S$, where M is a closed real subspace of X, then $\int_S x \in M$.*

The above Corollary and Theorem 1.7.4 imply

Corollary 4.2.6 *If Y is a Banach space and $Q \in L(S, \mathfrak{B}(X,Y))$ is such that $Q(t)$ is a compact linear operator for each $t \in S$, then $\int_S Q$ is a compact linear operator.*

The Dominated Convergence Theorem generalizes to the Bochner integral:

Theorem 4.2.7 *If $\{x_1, x_2, \ldots\} \subset L(S,X)$, $x : S \to X$, $f \in L^1(S)$ are such that*

(i) $\lim_{i\to\infty} \|x_i(s) - x(s)\| = 0$ for almost all $s \in S$ and

(ii) for each $i \geq 1$ we have that $\|x_i(s)\| \leq f(s)$ for almost all $s \in S$,

then $x \in L(S,X)$ and $\int_S x = \lim_{i\to\infty} \int_S x_i$.

PROOF It is easy to see that x is strongly measurable; hence $x \in L(S,X)$ by (ii). The Lebesgue DCT and Theorem 4.2.3 imply

$$\left\| \int_S x - \int_S x_i \right\| \leq \int_S \|x - x_i\| \to 0 \text{ as } i \to \infty.$$

□

Theorem 4.2.8 *If $x \in L(S,X)$, Y is a Banach space and $T \in \mathfrak{B}(X,Y)$, then*

$$Tx \in L(S,Y) \quad \text{and} \quad T \int_S x = \int_S Tx.$$

PROOF Tx is almost separably-valued and if $\ell \in Y^\star$, then $\ell(Tx) = (T^\star \ell)(x)$ is Lebesgue measurable; hence $Tx \in L(S,Y)$ and

$$\int_S \ell(Tx) = \int_S (T^\star \ell)(x) = (T^\star \ell)\left(\int_S x\right) = \ell\left(T \int_S x\right).$$

□

Corollary 4.2.9 *If X, Y are Banach spaces, $Q \in L(S, \mathfrak{B}(X,Y))$ and $c \in X$, then*

$$Qc \in L(S,Y) \quad \text{and} \quad \left(\int_S Q\right) c = \int_S Qc.$$

PROOF Let $T\Lambda = \Lambda c$ for $\Lambda \in \mathfrak{B}(X,Y)$ and note that $T \in \mathfrak{B}(\mathfrak{B}(X,Y), Y)$.

□

Theorem 4.2.10 *Suppose that Y is a Banach space and $A : \mathcal{D}(A) \subset X \to Y$ is a closed linear operator. If $x : S \to \mathcal{D}(A)$, $x \in L(S,X)$ and $Ax \in L(S,Y)$, then*

$$\int_S x \in \mathcal{D}(A) \quad \textit{and} \quad A\int_S x = \int_S Ax.$$

PROOF Let $\|(u,v)\| = \|u\| + \|v\|$ for $(u,v) \in Z \equiv X \times Y$. If $\ell \in Z^\star$, then $\ell(u,v) = \ell_1 u + \ell_2 v$ for some $\ell_1 \in X^\star$ and $\ell_2 \in Y^\star$. Hence, the function $s \to f(s) \equiv (x(s), Ax(s))$ is Bochner integrable in Z and

$$\int_S \ell f = \int_S \ell_1 x + \int_S \ell_2 Ax = \ell_1 \int_S x + \ell_2 \int_S Ax = \ell\left(\int_S x, \int_S Ax\right).$$

Thus $\int_S f = (\int_S x, \int_S Ax)$. Let $M = \{(u, Au) \mid u \in \mathcal{D}(A)\}$. Since A is closed, M is a closed subspace of Z. Corollary 4.2.5 implies that $\int_S f \in M$. □

Theorem 4.2.11 *Suppose that*

(1) $-\infty < a < b < \infty$, $x : [a,b] \to X$ is continuous

(2) x is differentiable at each point of (a,b)

(3) there exists $f \in L^1(a,b)$ such that $\|x'(t)\| \le f(t)$ for all $t \in (a,b)$.

Then $x' \in L((a,b),X)$ and

$$x(b) = x(a) + \int_a^b x'(t)dt.$$

PROOF Let $x(t) = 0$ for $t \in \mathbb{R}\backslash[a,b]$ and define

$$x_n(t) = n(x(t+1/n) - x(t)) \quad \text{for} \quad t \in \mathbb{R},\ n \ge 1.$$

Note that x_n are strongly measurable and converge to x' as $n \to \infty$ at each point of (a,b). Hence $x' \in L((a,b),X)$.

Choose any $\ell \in X^\star$ and let $g(t) = \ell(x(t))$ for $t \in [a,b]$. Since $g' = \ell(x')$, the classical result for complex valued functions implies

$$\int_a^b \ell(x'(t))dt = \int_a^b g'(t)dt = g(b) - g(a) = \ell(x(b) - x(a)).$$

□

4.3 Basic Properties of Semigroups

Theorem 4.3.1 *Suppose $\{Q(t)\}_{t\geq 0}$ is a strongly continuous semigroup on a Banach space X and let $-A$ be the generator of the semigroup. Then*

(1) There exist $M \in [0,\infty)$, $a \in \mathbb{R}$ such that $\|Q(t)\| \leq Me^{-at}$ for all $t \geq 0$.

(2) $t \to Q(t)x$ is a continuous mapping of $[0,\infty)$ into X for every $x \in X$.

(3) If $x \in X$, $t \geq 0$, then $\int_0^t Q(s)x\,ds \in \mathcal{D}(A)$ and $x - Q(t)x = A\int_0^t Q(s)x\,ds$.

(4) If $x \in \mathcal{D}(A)$, $u(t) = Q(t)x$ for $t \geq 0$, then for each $t \geq 0$ we have that

$$\frac{du}{dt}(t) \text{ exists, } u(t) \in \mathcal{D}(A), \frac{du}{dt}(t) = -Au(t) = -Q(t)Ax.$$

(5) $(A-\lambda)^{-1}Q(t) = Q(t)(A-\lambda)^{-1}$ for every $\lambda \in \rho(A)$, $t \geq 0$.

(6) A is a closed linear operator.

(7) $\cap_{n=1}^{\infty}\mathcal{D}(A^n)$ is dense in X.

(8) If $\tau \in (0,\infty]$, $u : [0,\tau) \to X$ is continuous and such that

$$\frac{du}{dt}(t) \text{ exists, } u(t) \in \mathcal{D}(A) \text{ and } \frac{du}{dt}(t) + Au(t) = 0 \quad \text{for } t \in (0,\tau),$$

then $u(t) = Q(t)u(0)$ for all $t \in [0,\tau)$.

(9) If $\{T(t)\}_{t\geq 0}$ is a strongly continuous semigroup on X and the generator of this semigroup is equals $-A$, then $T(t) = Q(t)$ for all $t \geq 0$.

(10) If λ is any scalar, then $\{e^{\lambda t}Q(t)\}_{t\geq 0}$ is a strongly continuous semigroup on X and the generator of this semigroup is $\lambda - A$.

PROOF Let us show first that $\sup_{0\leq t\leq \varepsilon}\|Q(t)\| < \infty$ for some $\varepsilon > 0$. If there would be no such ε, then one could choose $t_n \in (0, 1/n)$ such that $\|Q(t_n)\| > n$; the uniform boundedness principle (Theorem 1.4.5) would then imply that $\sup_n \|Q(t_n)x\| = \infty$ for some x in X. However, this is not possible because, see Definition 4.1.1, $\lim_{n\to\infty} Q(t_n)x = x$ for every $x \in X$.

Therefore there exist $M \in (1,\infty)$ and $\varepsilon \in (0,\infty)$ such that $\|Q(t)\| \leq M$ for all $t \in [0,\varepsilon]$. Let $M = e^{-a\varepsilon}$. If $n \geq 0$, $n\varepsilon \leq t < (n+1)\varepsilon$, then

$$\|Q(t)\| = \|Q(t-n\varepsilon)Q(\varepsilon)^n\| \leq MM^n \leq Me^{-at}.$$

To show (2), it is enough to observe that for each $x \in X$,

$$\|Q(t+h)x - Q(t)x\| \leq \|Q(t)\|\|Q(h)x - x\| \quad \text{if} \quad t \geq 0, h \geq 0$$

$$\|Q(t-h)x - Q(t)x\| \leq Me^{-a(t-h)}\|Q(h)x - x\| \quad \text{if} \quad 0 \leq h \leq t.$$

(3) is obvious if $t = 0$. Assume $t > 0$ and $h > 0$. Then

$$\begin{aligned}\frac{1}{h}(1 - Q(h))\int_0^t Q(s)x\,ds &= \frac{1}{h}\int_0^t Q(s)x\,ds - \frac{1}{h}\int_0^t Q(s+h)x\,ds \\ &= \frac{1}{h}\int_0^t Q(s)x\,ds - \frac{1}{h}\int_h^{t+h} Q(s)x\,ds\end{aligned}$$

$$= x - Q(t)x + \frac{1}{h}\int_0^h (Q(s)x - x)ds - \frac{1}{h}\int_t^{t+h}(Q(s)x - Q(t)x)ds$$

and this implies (3).

To show (4), choose $h > 0$ and note that for all $t \geq 0$,

$$\frac{1}{h}(1 - Q(h))Q(t)x = Q(t)\frac{1}{h}(x - Q(h)x) \to Q(t)Ax \quad \text{as} \quad h \to 0.$$

This implies that $u(t) \in \mathcal{D}(A)$, $Au(t) = Q(t)Ax$ and $D^+u = -Au$, where D^+ denotes the right derivative. If $0 < h < t$, then

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

$$-Au(t) + Q(t-h)(Ax - \frac{1}{h}(x - Q(h)x)) + Q(t)Ax - Q(t-h)Ax$$

and (4) follows.

To prove (5), choose $y \in X$ and let $x = (A - \lambda)^{-1}y$. (4) implies that $(A - \lambda)Q(t)x = Q(t)(A - \lambda)x = Q(t)y$. Hence, $Q(t)x = (A - \lambda)^{-1}Q(t)y$.

To prove (6), let $x_n \in \mathcal{D}(A)$, $n \geq 1$, be such that for some $x, y \in X$ we have that $\lim_{n\to\infty} x_n = x$, $\lim_{n\to\infty} Ax_n = y$. (4) and Theorem 4.2.11 imply that

$$x_n - Q(t)x_n = \int_0^t Q(s)Ax_n ds \quad \text{for} \quad t \geq 0, n \geq 1.$$

Theorem 4.2.7 enables us take the limit $n \to \infty$. Hence,

$$\frac{1}{t}(x - Q(t)x) = \frac{1}{t}\int_0^t Q(s)y\,ds \to y \quad \text{as} \quad t \to 0;$$

thus $x \in \mathcal{D}(A)$, $Ax = y$ and (6) follows.

Let us prove (7) now. Choose $x \in X$, $\varphi \in C_0^\infty(0, \infty)$ and define

$$y(x, \varphi) = \int_0^\infty \varphi(s)Q(s)x\,ds.$$

Using Theorem 4.2.3, it is easy to see that

$$y(x,\varphi) = -\int_0^\infty \varphi'(s)u(s)\,ds, \quad \text{where} \quad u(t) = \int_0^t Q(s)x\,ds.$$

Since $Au(t) = x - Q(t)x$ by (3) we have that $\varphi' Au \in L((0,\infty), X)$; since A is closed, Theorem 4.2.10 implies that $y(x,\varphi) \in \mathcal{D}(A)$ and

$$Ay(x,\varphi) = \int_0^\infty \varphi'(s)(Q(s)x - x)\,ds = y(x,\varphi').$$

Since φ' is also in $C_0^\infty(0,\infty)$, we have that $y(x,\varphi) \in \cap_{n=1}^\infty \mathcal{D}(A^n)$.

Fix any $x \in X$, $\varepsilon > 0$. Choose $t \in (0,\infty)$, $\mu \in (0,t)$ so that

$$\|Q(t)x - x\| < \varepsilon/2, \qquad \sup_{|s-t|<\mu} \|Q(s)x - Q(t)x\| < \varepsilon/2.$$

Pick $\varphi \in C_0^\infty(0,\infty)$ such that $\varphi \geq 0$, $\varphi(s) = 0$ if $|s-t| \geq \mu$, $\int_0^\infty \varphi = 1$. Then

$$y(x,\varphi) - x = Q(t)x - x + \int_0^\infty \varphi(s)(Q(s)x - Q(t)x)ds$$

$$\|y(x,\varphi) - x\| \leq \|Q(t)x - x\| + \sup_{|s-t|<\mu} \|Q(s)x - Q(t)x\| < \varepsilon,$$

completing (7)'s proof.

To show (8), choose $t \in (0,\tau)$ and let $v(s) = Q(t-s)u(s)$ for $s \in [0,t]$. If $s \in (0,t]$, $s + h \in (0,t]$, then

$$\begin{aligned} v(s+h) - v(s) \;=\; & Q(t-s-h)(u(s+h) - u(s) - hu'(s)) + \\ & Q(t-s-h)u(s) - Q(t-s)u(s) - hQ(t-s)Au(s) + \\ & hQ(t-s)Au(s) - hQ(t-s-h)Au(s). \end{aligned}$$

Therefore $v'(s) = 0$ for $s \in (0,t]$. Since

$$v(s) - v(0) = Q(t-s)(u(s) - u(0)) + Q(t-s)u(0) - Q(t)u(0),$$

we see that v is continuous on $[0,t]$ and therefore Theorem 4.2.11 gives that $v(t) = v(0)$, implying assertion (8).

To obtain (9), choose $x \in \mathcal{D}(A)$ and let $u(t) = T(t)\mathrm{x}$. From (4) we see that u satisfies assumptions of (8), with $\tau = \infty$; therefore $T(t)x = Q(t)x$ for $t \geq 0$. This and the fact that $\mathcal{D}(A)$ is dense in X imply (9).

(10) follows immediately from Definitions 4.1.1 and 4.1.2. □

Theorem 4.3.2 *Suppose that $-A$ is the generator of a strongly continuous semigroup $\{Q(t)\}_{t\geq 0}$ on a Banach space X. Let $M \in [0,\infty)$, $a \in \mathbb{R}$ be such that $\|Q(t)\| \leq Me^{-at}$ for all $t > 0$. Then every scalar λ, with $\operatorname{Re}\lambda < a$, belongs to the resolvent set of A and, moreover,*

$$\|(A-\lambda)^{-n}\| \leq M(a-\operatorname{Re}\lambda)^{-n} \quad \textit{for} \quad n \geq 1,$$

$$(A-\lambda)^{-n}x = \frac{1}{(n-1)!}\int_0^\infty s^{n-1}e^{\lambda s}Q(s)x\,ds \quad \textit{for} \quad n \geq 1, x \in X.$$

PROOF For $n \geq 1$, $x \in X$ define

$$R_n x = \frac{1}{(n-1)!}\int_0^\infty s^{n-1}e^{\lambda s}Q(s)x\,ds.$$

Clearly, R_n is a linear operator on X and, since

$$\|R_n x\| \leq \frac{1}{(n-1)!}\int_0^\infty s^{n-1}Me^{(\operatorname{Re}\lambda - a)s}\|x\|ds = M(a-\operatorname{Re}\lambda)^{-n}\|x\|,$$

we have that $R_n \in \mathfrak{B}(X)$. We will show, by induction, that $R_n = (A-\lambda)^{-n}$.

If $x \in \mathcal{D}(A)$, then

$$\frac{d}{dt}e^{\lambda t}Q(t)x = e^{\lambda t}Q(t)(\lambda - A)x$$

$$x - e^{\lambda t}Q(t)x = \int_0^\infty \chi_{[0,t]}(s)e^{\lambda s}Q(s)(A-\lambda)x\,ds \quad \text{for} \quad t \geq 0.$$

If we let $t \to \infty$, then by the Bochner DCT, Theorem 4.2.7, we obtain that $x = R_1(A-\lambda)x$. Since A is closed, Theorem 4.2.10 implies that $R_1 x \in \mathcal{D}(A)$, $(A-\lambda)R_1 x = x$. This and the facts that $\mathcal{D}(A)$ is dense in X, R_1 is bounded and A is closed give $R_1 y \in \mathcal{D}(A)$, $(A-\lambda)R_1 y = y$ for all $y \in X$. Therefore $(A-\lambda)^{-1} = R_1$.

Suppose now that $R_n = (A-\lambda)^{-n}$ for some $n \geq 1$. If $x \in X$, then

$$\frac{d}{dt}t^n e^{\lambda t}Q(t)R_1 x = nt^{n-1}e^{\lambda t}Q(t)R_1 x - t^n e^{\lambda t}Q(t)x$$

$$t^n e^{\lambda t}Q(t)R_1 x = \int_0^t (ns^{n-1}e^{\lambda s}Q(s)R_1 x - s^n e^{\lambda s}Q(s)x)ds$$

and, by the DCT, we obtain that $R_{n+1}x = R_n R_1 x = (A-\lambda)^{-n-1}x$. □

Corollary 4.3.3 *Suppose $-A$ is the generator of a strongly continuous semigroup $\{Q(t)\}_{t\geq 0}$ on a Banach space X. If λ is in the spectrum of A and if $b \in \mathbb{R}$, $b > \operatorname{Re}\lambda$, then $\sup_{t>0}\|Q(t)x\|e^{bt} = \infty$ for some x in X.*

PROOF If there would be no such x, then the uniform boundedness principle would imply that $\sup_{t>0} \|Q(t)\|e^{bt} < \infty$. This and Theorem 4.3.2 would then imply that λ is not in the spectrum of A. □

For example, if the spectrum of A contains any part of the half-plane $\operatorname{Re}\lambda < 0$, then $\{Q(t)x\}_{t\geq 0}$ is unbounded for some $x \in X$. One can show that x can be chosen so that $x \in \mathcal{D}(A)$, see Exercise 5. This and Examples 4.1.3 and 4.1.6 may lead one to expect that if $a < \inf \operatorname{Re}\sigma(A)$, then the semigroup is bounded by e^{-at}. This expectation is generally false, as Example 4.3.4 shows. However, it will be shown in the next section (Corollary 4.5.11) that it is true in many important cases.

EXAMPLE 4.3.4 Let X denote the set of all $f \in C_u(0,\infty)$ such that

$$\|f\| \equiv \sup_x |f(x)| + \int_0^\infty e^x |f(x)|dx < \infty.$$

X is a complex Banach space. Define $Q(t) : X \to X$ by

$$(Q(t)f)(x) = f(x+t) \quad \text{for} \quad t, x \geq 0.$$

Note that $\|Q(t)f\| \leq \|f\|$ for every $f \in X$ and $t \geq 0$. If $f_{nt}(x) = e^{-n(x-t)^2}$, then

$$\lim_{n\to\infty} \|Q(t)f_{nt}\| = 1 = \lim_{n\to\infty} \|f_{nt}\|,$$

which shows that $\|Q(t)\| = 1$ for $t \geq 0$. It is easy to verify (c) of Definition 4.1.1 and thus $\{Q(t)\}_{t\geq 0}$ is a strongly continuous semigroup on X.

Let $-A$ denote the generator of the semigroup. Since for each $x \in (0,\infty)$ the mapping $f \in X \to f(x) \in \mathbb{C}$ belongs to $X^\star$, we have, by Theorem 4.3.2, that

$$((A-\lambda)^{-1}f)(x) = \int_0^\infty e^{\lambda t} f(x+t)dt \quad \text{for} \quad f \in X,\ x > 0,\ \lambda \in \mathbb{C},\ \operatorname{Re}\lambda < 0.$$

If $g = (A-\lambda)^{-1}f$, then $|g(x)| \leq \int_0^\infty |f(x+t)|dt \leq \int_0^\infty |f(t)|dt$ and

$$\int_0^\infty e^x |g(x)|dx \leq \int_0^\infty \int_0^\infty e^x |f(x+t)|dt\,dx = \int_0^\infty e^x |f(x)|dx - \int_0^\infty |f(x)|dx;$$

thus

$$\|(A-\lambda)^{-1}f\| \leq \|f\| \quad \text{for} \quad f \in X,\ \lambda \in \mathbb{C},\ \operatorname{Re}\lambda < 0$$

and hence Theorem 1.6.11 implies

$$\sigma(A) \subset \{\xi \in \mathbb{C} \mid \operatorname{Re}\xi \geq 1\} \text{ even though } \|Q(t)\| = 1 \text{ for } t \geq 0.$$

The following Theorem 4.3.5, when combined with Theorem 4.3.2, is called the Hille-Yosida Theorem. It tells which operators are generators of C_0 semigroups. Its proof is rather lengthy and technical, but in essence, it is merely a justification of

$$\lim_{n\to\infty} (1 + tA/n)^{-n} = e^{-tA},$$

which is obviously true when t and A are scalars. There are many other ways of constructing the semigroup from a given generator - one of which will be used in Definition 4.5.9 for a bit smaller class of operators.

Theorem 4.3.5 *Suppose that A is a densely defined linear operator in a Banach space X and that for some $M \in [0,\infty)$, $a \in \mathbb{R}$, we have that $(-\infty,a) \subset \rho(A)$ and*

$$\|(A-\lambda)^{-n}\| \leq M(a-\lambda)^{-n} \quad \textit{for} \quad \lambda \in (-\infty,a),\ n=1,2,\ldots \tag{4.5}$$

Then there exists a strongly continuous semigroup $\{Q(t)\}_{t\geq 0}$ on X whose generator is equal to $-A$. Moreover,

$$\|Q(t)\| \leq Me^{-at} \quad \textit{for} \quad t \geq 0, \tag{4.6}$$

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \|Q(t)x - (1+(t/n)A)^{-n}x\| = 0 \quad \textit{for} \quad T \in (0,\infty),\ x \in X. \tag{4.7}$$

PROOF Choose any $T \in (0,\infty)$ and let $K \geq 1$ be such that $K + at \geq 1$ for all $t \in [0,T]$. Note that

$$(1+at/n)^{-1} \leq K,\ (1+at/n)^{-n} \leq K^K \quad \text{for} \quad t \in [0,T],\ n \geq K. \tag{4.8}$$

Since $-n/t < a$ for $t \in (0,T]$, $n \geq K$, we may define

$$R_n(t) = (n/t)(n/t + A)^{-1} = (1+(t/n)A)^{-1} \quad \text{for} \quad t \in (0,T],\ n \geq K, \tag{4.9}$$

and let $R_n(0) = 1$ for $n \geq K$. Equations (4.5) and (4.8) imply that

$$\|R_n(t)\| \leq MK,\ \|R_n^n(t)\| \leq MK^K \quad \text{for} \quad t \in (0,T], n \geq K. \tag{4.10}$$

If $x \in X$, $y \in \mathcal{D}(A)$, $t \in (0,T]$, $n \geq K$, then equations (4.9) and (4.10) imply that

$$R_n(t)x - x = (R_n(t)-1)(x-y) - (t/n)R_n(t)Ay$$

$$\|R_n(t)x - x\| \leq (MK+1)\|x-y\| + tMK\|Ay\|/n. \tag{4.11}$$

Since $\mathcal{D}(A)$ is assumed to be dense in X, equation (4.11) implies

$$\lim_{t\to 0} \|R_n(t)x - x\| = 0 \quad \text{for} \quad n \geq K,\ x \in X.$$

By induction,

$$\lim_{t\to 0} \|R_n^m(t)x - x\| = 0 \quad \text{for} \quad n \geq K,\ x \in X,\ m \geq 1. \tag{4.12}$$

Equation (4.11) and the fact that $\mathcal{D}(A)$ is dense in X also imply

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \|R_n(t)x - x\| = 0 \quad \text{for} \quad x \in X. \tag{4.13}$$

If $t \in (0,T]$, $t+h \in (0,T]$, $n \geq K$, then

$$R_n(t+h) - R_n(t) + (h/n)AR_n^2(t) = (h/t)^2(1 - R_n(t))^2 R_n(t+h)$$

and (4.10) implies that $R_n'(t) = -(1/n)AR_n^2(t)$. By induction,

$$(R_n^m)'(t) = -(m/n)AR_n^{m+1}(t) \quad \text{for} \quad t \in (0,T],\ n \geq K,\ m \geq 1. \tag{4.14}$$

This, (4.12) and (4.10) imply

$$x - R_n^n(t)x = \int_0^t R_n^{n+1}(s)Ax\,ds \quad \text{for} \quad t \in [0,T],\ n \geq K,\ x \in \mathcal{D}(A). \tag{4.15}$$

We are now ready to prove (4.17).

If $y \in \mathcal{D}(A^2)$, $t \in (0,T]$, $n \geq K$, $m \geq K$, then (4.14) implies

$$\frac{d}{ds}R_n^n(t-s)R_m^m(s)y = R_n^{n+1}(t-s)R_m^{m+1}(s)\left(\frac{s}{m} - \frac{t-s}{n}\right)A^2y \quad \text{for} \quad s \in (0,t). \tag{4.16}$$

This, (4.12) and (4.10) imply that

$$R_m^m(t)y - R_n^n(t)y = \int_0^t R_n^{n+1}(t-s)R_m^{m+1}(s)\left(\frac{s}{m} - \frac{t-s}{n}\right)A^2y\,ds.$$

Hence, if $x \in X$, then (4.10) implies

$$R_m^m(t)x - R_n^n(t)x = (R_m^m(t) - R_n^n(t))(x-y) + R_m^m(t)y - R_n^n(t)y$$

$$\|R_m^m(t)x - R_n^n(t)x\| \leq 2MK^K\|x-y\| + T^2M^4K^{2K+2}(1/n + 1/m)\|A^2y\|$$

and, since (4.12) implies that $\mathcal{D}(A^2)$ is dense in X, we have that

$$\lim_{m,n\to\infty} \sup_{t\in[0,T]} \|R_m^m(t)x - R_n^n(t)x\| = 0 \quad \text{for} \quad x \in X. \tag{4.17}$$

Since $T \in (0,\infty)$ is arbitrary, (4.17) enables us to define $Q(t)$ by

$$Q(t)x = \lim_{n\to\infty} R_n^n(t)x \quad \text{for} \quad t \geq 0,\ x \in X. \tag{4.18}$$

This and (4.17) imply (4.7). Clearly, $Q(0) = 1$, and, by (4.10), we have

$$\|Q(t)\| \leq Me^{-at} \quad \text{for} \quad t > 0. \tag{4.19}$$

Since $R_n^n(\cdot)x : [0,T] \to X$ is continuous for each $x \in X$, $n \geq K$, (4.17) also implies that $Q(\cdot)x : [0,\infty) \to X$ is continuous for every $x \in X$. This and (4.19) imply

$$\|x\| \leftarrow \|Q(t)x\| \leq Me^{-at}\|x\| \to M\|x\| \quad \text{as} \quad t \to 0;$$

hence $\|x\| \leq M\|x\|$ for each $x \in X$. Thus, if $X \neq \{0\}$, then $M \geq 1$ and (4.6) holds; if $X = \{0\}$, then (4.6) holds automatically.

(4.16) together with (4.12) and (4.10) also imply that for $y \in \mathcal{D}(A^2)$, $n \geq K$,

$$\|R_n^n(t-s)R_n^n(s)y - R_n^n(t)y\| \leq \frac{t^2M^4K^{2K+2}}{n}\|A^2y\| \quad \text{when} \quad 0 \leq s \leq t \leq T.$$

Therefore, $Q(t-s)Q(s)y = Q(t)y$ for $0 \leq s \leq t$, $y \in \mathcal{D}(A^2)$. Since $\mathcal{D}(A^2)$ is dense in X and $Q(t)$ are bounded, we have that $Q(t)Q(s) = Q(t+s)$ for $s \geq 0$, $t \geq 0$. Therefore $\{Q(t)\}_{t\geq 0}$ is a strongly continuous semigroup on X. Letting $-B$ denotes its generator, (4.10), (4.13) and (4.15) imply

$$x - Q(t)x = \int_0^t Q(s)Ax\,ds \quad \text{for} \quad t \in [0,\infty),\ x \in \mathcal{D}(A).$$

This and strong continuity of Q imply that B is an extension of A. (4.5) and Theorem 4.3.2 imply $a - 1 \in \rho(A) \cap \rho(B)$ and hence $A = B$ by Lemma 1.6.14. □

The implicit **Euler method** for approximating the solution of

$$u' + Au = 0, \quad u(0) = u_0$$

is

$$\frac{u(t+h) - u(t)}{h} + Au(t+h) \approx 0,$$

hence

$$\begin{aligned} u(t+h) &\approx (1+hA)^{-1}u(t) \\ u(nh) &\approx (1+hA)^{-n}u_0 \\ u(t) &\approx (1+(t/n)A)^{-n}u_0. \end{aligned}$$

Note that Theorem 4.3.5 implies convergence of the approximations whenever $-A$ is the generator of a strongly continuous semigroup.

Theorem 4.3.6 *Suppose that $-A$ is the generator of a strongly continuous semigroup $\{Q(t)\}_{t\geq 0}$ on a reflexive Banach space X. Then $\{Q(t)^\star\}_{t\geq 0}$ is a strongly continuous semigroup on $X^\star$ whose generator is $-A^\star$.*

PROOF Since A is closed and densely defined, so is $A^\star$ (Theorems 4.3.1, 1.5.15). Let $M \in [0,\infty)$, $a \in \mathbb{R}$ be such that $\|Q(t)\| \leq Me^{-at}$ for $t \geq 0$. Then, by Theorem 4.3.2,

$$\|(A-\lambda)^{-n}\| \leq M(a-\lambda)^{-n} \quad \text{for} \quad n \geq 1,\ \lambda < a.$$

Hence Theorems 1.5.14 and 1.6.13 imply

$$\|(A^\star - \lambda)^{-n}\| \leq M(a-\lambda)^{-n} \quad \text{for} \quad n \geq 1,\ \lambda < a$$

and therefore, by Theorem 4.3.5, $-A^\star$ is a generator of a strongly continuous semigroup $\{R(t)\}_{t\geq 0}$ on $X^\star$. Theorems 4.3.5 and 1.6.13 also imply

$$\begin{aligned}(R(t)\ell)(x) &= \lim_{n\to\infty}((1+(t/n)A^\star)^{-n}\ell)(x)\\ &= \lim_{n\to\infty}\ell((1+(t/n)A)^{-n}x)\\ &= \ell(Q(t)x) = (Q(t)^\star\ell)(x)\end{aligned}$$

for all $x \in X$, $\ell \in X^\star$, $t \geq 0$ and therefore $R(t) = Q(t)^\star$ for $t \geq 0$. □

This, the definition of the Hilbert space adjoint and Theorem 2.2.5 imply

Corollary 4.3.7 *Suppose $-A$ is the generator of a strongly continuous semigroup $\{Q(t)\}_{t\geq 0}$ on a Hilbert space H. Then $\{Q(t)^*\}_{t\geq 0}$ is a strongly continuous semigroup on H whose generator is $-A^*$.*

A strongly continuous semigroup $\{Q(t)\}_{t\geq 0}$ is said to be a **contraction semigroup** if $\|Q(t)\| \leq 1$ for $t \geq 0$. See Examples 4.1.3, 4.1.5 and 4.3.4 for some contraction semigroups, and Exercise 19 of Chapter 1 for analysis of the generator of the contraction semigroup given in Example 4.1.3.

Theorem 4.3.8 *Suppose that A is a linear operator in a Banach space X. Then $-A$ is the generator of a contraction semigroup if and only if A is a densely defined accretive linear operator and $\mathcal{R}(A+\lambda) = X$ for some $\lambda > 0$.*

PROOF If A is a densely defined accretive linear operator and $\mathcal{R}(A+\lambda) = X$ for some $\lambda > 0$, then Corollary 1.6.17 implies that the assumptions of Theorem 4.3.5 are satisfied with $M = 1$, $a = 0$.

Suppose that $-A$ is the generator of a contraction semigroup $\{Q(t)\}_{t\geq 0}$ and that $x \in \mathcal{D}(A)$. By Corollary 1.5.8 there exists $\ell \in X^\star$ such that $\ell(x) = \|\ell\|^2 = \|x\|^2$ - pick any such ℓ. Define $f(t) = \mathrm{Re}\,\ell(Q(t)x)$ and observe that

$$f(0) = \|x\|^2,\ f(t) \leq \|x\|^2,\ f'(t) = -\mathrm{Re}\,\ell(Q(t)Ax) \quad \text{for} \quad t \geq 0.$$

Therefore, $-\mathrm{Re}\,\ell(Ax) = f'(0) \leq 0$. Theorem 4.3.2 implies that A + 1 is onto. □

EXAMPLE 4.3.9 Let $X = \{f \in C[0,1] \mid f(0) = f(1) = 0\}$ and let $\|\cdot\|$ be the sup norm on X. Assume $a \in C[0,1]$ and $a(x) \in (0,\infty)$ for $x \in [0,1]$. Define

$$Tf = -af'' \quad \text{for} \quad f \in \mathcal{D}(T) \equiv \{f \in C^2[0,1] \mid f, f'' \in X\}.$$

In Example 1.6.18 it is shown that T is an accretive linear operator and that $\mathcal{R}(T+\lambda) = X$ for some $\lambda > 0$. It is easy to show, as in Example 1.3.4, that T is densely defined. Hence, Theorem 4.3.8 implies that $-T$ is the generator of a contraction semigroup $\{Q(t)\}_{t\geq 0}$ on X. Thus, if $f \in \mathcal{D}(T)$ and $u(x,t) = (Q(t)f)(x)$, then (4) of Theorem 4.3.1 implies the existence of a classical solution of the parabolic equation

$$u_t(x,t) = a(x)u_{xx}(x,t) \quad \text{for} \quad t \geq 0,\ 0 \leq x \leq 1$$

$$u(0,t) = u(1,t) = 0 \quad \text{for} \quad t \geq 0$$

$$\lim_{t\to 0+} \sup_{0\leq x\leq 1} |u(x,t) - f(x)| = 0$$

$$|u(x,t)| \leq \|f\| \quad \text{for} \quad t \geq 0,\ 0 \leq x \leq 1.$$

(8) of Theorem 4.3.1 gives uniqueness for the abstract equation. Uniqueness in a more classical sense can be shown by using a maximum principle (see Theorem 6.7.2 and Corollary 6.7.3).

EXAMPLE 4.3.10 We shall now show that $\overline{\Delta}$ is the generator of a contraction semigroup on $X \equiv W_0^{0,p}(\mathbb{R}^n)$, $1 \leq p \leq \infty$. Note that when $p < \infty$, Theorem 3.1.7 implies that $X = L^p(\mathbb{R}^n)$ and, when $p = \infty$, Theorem 3.1.4 implies that $X = C_{t0}$.

Suppose that $f_1, f_2, \ldots$ belong to the Schwartz space $\mathcal{S}$ and are such that $\Delta f_n \overset{X}{\to} g$ and $f_n \overset{X}{\to} 0$ as $n \to \infty$. Then

$$\int_{\mathbb{R}^n} g\phi \leftarrow \int_{\mathbb{R}^n} \Delta f_n\, \phi = \int_{\mathbb{R}^n} f_n \Delta\phi \to 0 \quad \text{for} \quad \phi \in C_0^\infty(\mathbb{R}^n)$$

and hence Theorem 3.4.1 implies that $g = 0$. Therefore Δ, as an operator in X with domain $\mathcal{D}(\Delta) = \mathcal{S}$, is closable; let $\overline{\Delta}$ denote its closure.

Let us show now that $\sigma(\overline{\Delta}) = (-\infty, 0]$ and that

$$(\lambda - \overline{\Delta})^{-1} f = G_\lambda * f \quad \text{for} \quad f \in X,\ \lambda \in \mathbb{C}\backslash(-\infty, 0] \tag{4.20}$$

where G_λ is given by (3.6). This, (3.8) and Lemma 3.1.5 imply that $\|(\lambda - \overline{\Delta})^{-1}\| \leq 1/\lambda$ for $\lambda > 0$ and hence Theorem 4.3.5 implies that $\overline{\Delta}$ is the generator of a contraction semigroup.

Suppose $\lambda \in \mathbb{C}\backslash(-\infty, 0]$ and $f \in X$. There exist $f_n \in \mathcal{S}$ such that $f_n \overset{X}{\to} f$. Let $u_n = G_\lambda * f_n \in \mathcal{S}$. Lemma 3.1.5 and (3.7) imply that $u_n \overset{X}{\to} G_\lambda * f$ and, since it is shown in Example 3.2.11 that $\Delta u_n = \lambda u_n - f_n \overset{X}{\to} \lambda G_\lambda * f - f$, we have that

$$G_\lambda * f \in \mathcal{D}(\overline{\Delta}), \quad (\lambda - \overline{\Delta}) G_\lambda * f = f \quad \text{for} \quad f \in X,\ \lambda \in \mathbb{C}\backslash(-\infty, 0].$$

If $u \in \mathcal{D}(\overline{\Delta})$, then there exist $u_n \in \mathcal{S}$ such that $u_n \overset{X}{\to} u$ and $\Delta u_n \overset{X}{\to} \overline{\Delta}u$. As above, $u_n = G_\lambda * (\lambda u_n - \Delta u_n) \to G_\lambda * (\lambda u - \overline{\Delta}u)$ which implies that

$$u = G_\lambda * (\lambda u - \overline{\Delta}u) \quad \text{for} \quad u \in \mathcal{D}(\overline{\Delta}),\ \lambda \in \mathbb{C}\backslash(-\infty, 0].$$

This completes the proof of (4.20) and also shows that $\sigma(\overline{\Delta}) \subset (-\infty, 0]$.

Pick any $\xi \in \mathbb{R}^n$, $\varepsilon > 0$ and let $f(x) = e^{i\xi\cdot x - \varepsilon^2 x\cdot x}$ for $x \in \mathbb{R}^n$. Note that $f \in \mathcal{S}$,

$$|\xi|^2 f(x) + (\Delta f)(x) = (-4i\varepsilon^2 \xi \cdot x + 4\varepsilon^4 x \cdot x - 2n\varepsilon^2) f(x),$$

$$0 < \|(|\xi|^2 + \Delta) f\|_p \le c(|\xi|\varepsilon + \varepsilon^2)\|f\|_p$$

where c depends only on n and p. This shows that $(\overline{\Delta} + |\xi|^2)^{-1}$ cannot be bounded and hence $(-\infty, 0] \subset \sigma(\overline{\Delta})$. Therefore $(-\infty, 0] = \sigma(\overline{\Delta})$.

Note also that if B is any extension of Δ with a *nonempty* resolvent set, then B has to be a closed operator and hence must be an extension of $\overline{\Delta}$; therefore Lemma 1.6.14 implies that $B = \overline{\Delta}$. When $1 \le p < \infty$ and $n = 1$, this implies that $-\overline{\Delta}$ is equal to the operator S that is given in Example 1.6.7. Similarly, if A is the self-adjoint linear operator in $L^2(\mathbb{R}^n)$ associated with the sectorial form

$$\mathfrak{F}(u, v) = \int_{\mathbb{R}^n} \nabla u \cdot \overline{\nabla v} \quad \text{for} \quad u, v \in W_0^1(\mathbb{R}^n),$$

then $A = -\overline{\Delta}$ when $p = 2$.

Theorem 4.3.8 in a Hilbert space setting is as follows:

Corollary 4.3.11 *If A is a linear operator in a Hilbert space, then $-A$ is the generator of a contraction semigroup iff A is m-accretive.*

In applications, one often encounters the following perturbation problem. You know that A is the generator of a strongly continuous semigroup and that the perturbation B does not effect the 'principle' part of $A + B$. Is $A + B$ the generator of a strongly continuous semigroup? Corollary 4.3.11 and Theorem 4.3.12 provide a partial answer. For other partial answers, see Corollary 5.1.3 and Theorem 4.5.7.

Theorem 4.3.12 *If A is an m-accretive operator and B is an accretive operator in a Hilbert space with $\mathcal{D}(A) \subset \mathcal{D}(B)$, and if there exist $a \in [0, 1)$, $b < \infty$ such that*

$$\|Bu\| \le a\|Au\| + b\|u\| \quad \textit{for all} \quad u \in \mathcal{D}(A),$$

then $A + B$ is m-accretive.

PROOF $A + B$ is obviously accretive. If $\lambda > 0$, then

$$\|(A + \lambda)u\|^2 = \|Au\|^2 + 2\lambda \mathrm{Re}\,(Au, u) + \lambda^2\|u\|^2 \ge \|Au\|^2 \quad \text{for} \quad u \in \mathcal{D}(A);$$

hence $\|A(A+\lambda)^{-1}\| \leq 1$, which implies

$$\|B(A+\lambda)^{-1}\| \leq a\|A(A+\lambda)^{-1}\| + b\|(A+\lambda)^{-1}\| \leq a + b/\lambda.$$

Therefore $\|B(A+\lambda)^{-1}\| \leq 1$ for λ large enough and hence $A+B$ is m-accretive because

$$(A+B+\lambda)^{-1} = (A+\lambda)^{-1}(1+B(A+\lambda)^{-1})^{-1}.$$

□

EXAMPLE 4.3.13 The Schrödinger equation is an evolution equation in a complex Hilbert space,

$$\frac{d\psi}{dt} = iH\psi,$$

where H is a self-adjoint operator called a Hamiltonian. $\pm iH$ are obviously accretive and, by Theorem 2.6.2, m-accretive. Thus, $\pm iH$ generate contraction semigroups.

A Hamiltonian for a particle in a potential field V is defined in $L^2(\mathbb{R}^3)$ by

$$H\psi = -\overline{\Delta}\psi + V\psi \quad \text{for} \quad \psi \in \mathcal{D}(\overline{\Delta}).$$

In order for this H to be well defined and self-adjoint, we need to impose some restrictions on V. We shall now show that the following restrictions are sufficient: assume that V is a real valued function and that $V = V_1 + V_2$ for some $V_1 \in L^2(\mathbb{R}^3)$ and $V_2 \in L^\infty(\mathbb{R}^3)$. Thus, V could, for example, be the Coulomb potential $-e^2(x^2+y^2+z^2)^{-1/2}$ and in this case H would represent the hydrogen atom Hamiltonian. $\overline{\Delta}$ is as defined in Example 4.3.10 with $p = 2$.

If $u \in \mathcal{S}$, then $\|V_1 u\|_2 \leq \|V_1\|_2 \|u\|_\infty$. Hence, (3.14) implies that

$$\|V_1 u\|_2 \leq c\|V_1\|_2 \|\Delta u\|_2^{3/4} \|u\|_2^{1/4}.$$

(3.52) and $\|V_2 u\|_2 \leq \|V_2\|_\infty \|u\|_2$ imply that for any $a \in (0,1)$, we can find $b < \infty$ such that

$$\|Vu\|_2 \leq a\|\overline{\Delta}u\|_2 + b\|u\|_2 \quad \text{for all} \quad u \in \mathcal{D}(\overline{\Delta}).$$

Since $\pm i\overline{\Delta}$ are m-accretive and the multiplication operators $\pm iV$, with domain $\mathcal{D}(\overline{\Delta})$, are accretive, we can apply Theorem 4.3.12 to conclude that $\pm iH$ are m-accretive. Hence $\mathcal{R}(H \pm i) = L^2(\mathbb{R}^3)$ and therefore Theorem 2.6.3 implies that H is self-adjoint.

4.4 Example: Wave Equation

We shall first examine the **generalized wave equation** (4.23) in the same Hilbert space setting as in the section on Sectorial Forms. Hence we assume that $\mathcal{H}$ is a Hilbert space, $[\cdot,\cdot]$ is an inner product on $\mathcal{V} \subset \mathcal{H}$ and $\mathfrak{F}$ is a sectorial sesquilinear form on $\mathcal{V}$ such that hypotheses **H1**, **H2**, **H3** of the Sectorial Forms section are satisfied. Here we assume, in addition, that there exists $b < \infty$ such that

$$|\mathfrak{F}(x,y) - \overline{\mathfrak{F}(y,x)}| \leq 2b|x|\|y\| \quad \text{for all} \quad x, y \in \mathcal{V}, \tag{4.21}$$

where $|x| = [x,x]^{1/2}$. Let A be the operator associated with $\mathfrak{F}$.

The need for an additional assumption is due to the fact that the point spectrum of A has to lie within a parabola, as demonstrated in Example 4.1.7. Assumption (4.21) guarantees that (Exercise 4). Some other consequences of (4.21) are discussed in Theorem 2.8.6 and Corollary 6.1.14.

Define $X = \mathcal{V} \times \mathcal{H}$. X is a Hilbert space with an inner product

$$(\{x,y\},\{z,w\}) = [x,z] + (y,w) \quad \text{for} \quad \{x,y\},\{z,w\} \in X.$$

Define $S : \mathcal{D}(S) \to X$ by $\mathcal{D}(S) = \{\{x,y\} \mid x \in \mathcal{D}(A), y \in \mathcal{V}\}$ and

$$S\{x,y\} = \{-y, Ax\} \quad \text{for} \quad \{x,y\} \in \mathcal{D}(S).$$

Theorem 4.4.1 *S and $-S$ are generators of strongly continuous semigroups on X. Moreover, there exist $c, M \in (0,\infty)$ which depend only on b, a, M_1, M_2, M_3 such that*

$$\|(S-\lambda)^{-n}\| \le M(|\lambda| - c)^{-n} \quad \textit{for} \quad |\lambda| > c,\ \lambda \in \mathbb{R},\ n \ge 1.$$

PROOF Note first that $\mathcal{D}(S)$ is dense in X by Theorem 2.8.2. Next, define a new inner product in X by

$$(\{x,y\},\{z,w\})_{new} = \mathfrak{F}(x,z) + \overline{\mathfrak{F}(z,x)} - 2a(x,z) + 2(y,w)$$

and observe that the corresponding new norm satisfies

$$2\min\{1, M_3\}\|\{x,y\}\|^2 \le \|\{x,y\}\|_{new}^2 \le 2\max\{1, M_2 + |a|M_1^2\}\|\{x,y\}\|^2.$$

If $\{x,y\} \in \mathcal{D}(S)$, then

$$\mathrm{Re}\,(S\{x,y\},\{x,y\})_{new} = \mathrm{Re}\,(\mathfrak{F}(x,y) - \overline{\mathfrak{F}(y,x)} + 2a(y,x))$$

$$|\mathrm{Re}\,(S\{x,y\},\{x,y\})_{new}| \le 2(b + |a|M_1)|x|\|y\| \le c\|\{x,y\}\|_{new}^2,$$

where $c = (b + |a|M_1)/(2M_3^{1/2})$. Thus, if $\lambda \in \mathbb{R}$ and $|\lambda| > c$, then

$$|\mathrm{Re}\,((S-\lambda)\{x,y\},\{x,y\})_{new}| \ge (|\lambda| - c)\|\{x,y\}\|_{new}^2$$

$$\|(S-\lambda)\{x,y\}\|_{new} \ge (|\lambda| - c)\|\{x,y\}\|_{new}. \tag{4.22}$$

If $\{f,g\} \in X$, $\lambda \in \mathbb{R}$ and $|\lambda| > c$, then $-\lambda^2 < a + M_3 M_1^{-2}$; hence Theorem 2.8.2 enables us to define

$$x = (A + \lambda^2)^{-1}(g - \lambda f), \quad y = -\lambda x - f.$$

Note that $\{x,y\} \in \mathcal{D}(S)$ and $(S-\lambda)\{x,y\} = \{f,g\}$. This and (4.22) imply

$$\|(S-\lambda)^{-1}\|_{new} \le (|\lambda| - c)^{-1} \quad \text{for} \quad |\lambda| > c,\ \lambda \in \mathbb{R},$$

which implies $\|(S-\lambda)^{-n}\|_{new} \leq (|\lambda|-c)^{-n}$. Hence, the equivalence of the norms gives

$$\|(S-\lambda)^{-n}\| \leq M(|\lambda|-c)^{-n} \quad \text{for} \quad |\lambda| > c,\ \lambda \in \mathbb{R},\ n \geq 1$$

and some constant $M < \infty$. Therefore, Theorem 4.3.5 implies that both S and $-S$ are generators of strongly continuous semigroups. □

Theorem 4.4.2 *For each $x \in \mathcal{D}(A)$, $y \in \mathcal{V}$ there exists a unique $u \in C^2([0,\infty),\mathcal{H}) \cap C^1([0,\infty),\mathcal{V})$ such that $u(0) = x$, $u'(0) = y$, $u(t) \in \mathcal{D}(A)$ for $t \geq 0$ and*

$$u''(t) = -Au(t) \quad \textit{for all} \quad t \geq 0. \tag{4.23}$$

Let $\{Q(t)\}_{t\geq 0}$ be the strongly continuous semigroup whose generator is $-S$. (4) of Theorem 4.3.1 implies that if $x \in \mathcal{D}(A)$, $y \in \mathcal{V}$ and $\{u(t), v(t)\} = Q(t)\{x, y\}$, then for all $t \geq 0$ we have that $u(t) \in \mathcal{D}(A)$, $v(t) \in \mathcal{V}$, u is differentiable in $\mathcal{V}$ (and hence also in $\mathcal{H}$), $u' = v$, v is differentiable in $\mathcal{H}$ and $v' = -Au$. Hence we have u of Theorem 4.4.2. To show its uniqueness, define $v = u'$ and apply (8) of Theorem 4.3.1.

EXAMPLE 4.4.3 Consider now the initial value problem for the **wave equation**,

$$\frac{\partial^2 u}{\partial t^2} = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left(a_{ij}(x)\frac{\partial u}{\partial x_i}\right) - \sum_{i=1}^{n} a_i(x)\frac{\partial u}{\partial x_i} - a_0(x)u \quad \text{for} \quad t \geq 0,\ x \in \Omega$$

$$u(x,0) = f(x), \quad \frac{\partial u}{\partial t}(x,0) = g(x) \quad \text{for} \quad x \in \Omega$$

$$u(x,t) = 0 \quad \text{for} \quad t \geq 0,\ x \in \partial\Omega.$$

We assume that Ω is an arbitrary nonempty open subset of $\mathbb{R}^n$. $a_{ij} = a_{ji}$ are assumed to be real valued, bounded, measurable functions on Ω such that for some $\delta > 0$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \delta|\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^n, x \in \Omega.$$

We assume that $a_i \in C_B^1(\Omega)$ for $i \geq 1$ and $a_0 \in L^\infty(\Omega)$ are complex valued.

Let $\mathcal{H} = L^2(\Omega)$, $\mathcal{V} = W_0^1(\Omega)$ and define

$$\mathfrak{F}(w,v) = \sum_{i=1}^{n}\sum_{j=1}^{n}\int_\Omega a_{ij}\mathsf{D}_i w\,\overline{\mathsf{D}_j v} + \sum_{i=1}^{n}\int_\Omega a_i \mathsf{D}_i w\,\overline{v} + \int_\Omega a_0 w\overline{v} \quad \text{for} \quad w, v \in \mathcal{V}.$$

In Section 3.7, it is shown that hypotheses **H**1, **H**2, **H**3 of the section on Sectorial Forms are satisfied. Since $\mathfrak{F}(w,v) - \overline{\mathfrak{F}(v,w)} = \int_\Omega z\overline{v}$, where

$$z = \sum_{i=1}^{n}(a_i + \overline{a}_i)\mathsf{D}_i w + \sum_{i=1}^{n} w D_i\overline{a}_i + (a_0 - \overline{a}_0)w,$$

the additional assumption (4.21) is also satisfied and hence Theorem 4.4.2 implies existence of the unique generalized solution u for every $f \in \mathcal{D}(A)$ and $g \in \mathcal{V}$.

The semigroup $\{Q(t)\}_{t\geq 0}$ actually provides a generalized solution u for every $f \in \mathcal{V}$ and $g \in \mathcal{H}$; a detailed characterization of the solution in this case is given below by Theorem 4.4.6. We shall also show that a different semigroup, $\{R(t)\}_{t\geq 0}$, can be constructed which will enable us to define a generalized solution for every $f \in \mathcal{H}$ and $g \in \mathcal{H}$. See also Example 3.2.15.

Let us now examine the operator A associated with $\mathfrak{F}$. When $a_{ij}, a_0 \in C^1(\Omega)$, Theorem 3.8.2 implies that $w \in \mathcal{D}(A)$ iff $w \in W_0^1(\Omega) \cap W_{loc}^2(\Omega)$ and

$$Aw \equiv -\sum_{i=1}^{n}\sum_{j=1}^{n} \mathsf{D}_j(a_{ij}\mathsf{D}_i w) + \sum_{i=1}^{n} a_i \mathsf{D}_i w + a_0 w \in L^2(\Omega).$$

It is easy to apply the induction argument to show that $\mathcal{D}(A^m) \subset W_{loc}^{2m}(\Omega)$ when $m \geq 1$ and $a_{ij}, a_i \in C^{2m-1}(\Omega)$.

If $2m - n/2 > j \geq 0$, then $W_{loc}^{2m}(\Omega) \subset C^j(\Omega)$ by Corollary 3.6.4. This enables us to prove regularity of the generalized solution as follows. Suppose that $f \in \mathcal{D}(A^k)$, $k \geq 2$, $g \in \mathcal{D}(A^{k-1})$ and $A^{k-1}g \in \mathcal{V}$. Let $\{u(t), v(t)\} = Q(t)\{f, g\}$. Note that $\{f, g\} \in \mathcal{D}(S^{2k-1})$. Hence, Theorem 4.3.1 implies that $\{u, v\}^{(2k-1)} = -S^{2k-1}\{u, v\}$ and since $u' = v$, we have that $u^{(2k)}$ exists and is continuous in $\mathcal{H}$, $u(t) \in \mathcal{D}(A^k)$ for $t \geq 0$ and $u^{(2k)} = (-A)^k u$. In particular, if $k > 1 + n/4$ and the coefficients are smooth enough, then $u(t) \in C^2(\Omega)$ for $t \geq 0$.

We shall now approach the generalized wave equation a bit differently. This approach is based on the representation of the sectorial form as given by Theorem 2.8.12; that is,

$$\mathfrak{F}(x, y) = (BGx, Gy) \quad \text{for} \quad x, y \in \mathcal{V},$$

which will enable us to define the semigroup on the Hilbert space $Y \equiv \mathcal{H} \times \mathcal{H}$ with the usual inner product $(\{x, y\}, \{z, w\}) = (x, z) + (y, w)$ for $\{x, y\}, \{z, w\} \in Y$. Define

$$T\{x, y\} = \{-Gy, BGx\} \quad \text{for} \quad \{x, y\} \in \mathcal{D}(T) \equiv \mathcal{V} \times \mathcal{V}.$$

Theorem 4.4.4 *T and $-T$ are generators of strongly continuous semigroups on Y.*

PROOF Note first that Theorem 2.8.12 implies

$$\|G^{-1}x\| \leq M_1|G^{-1}x| \leq M_1 M_3^{-1/2}\|x\| \quad \text{for} \quad x \in \mathcal{H}. \tag{4.24}$$

If $x, y \in \mathcal{V}$, then $\mathfrak{F}(x, y) - \overline{\mathfrak{F}(y, x)} = (Gx, (B^* - B)Gy)$. Hence (4.21) implies

$$|(Gx, (B^* - B)Gy)| \leq 2b|x|\|y\| \leq 2bM_3^{-1/2}\|Gx\|\|y\|$$

and, since $\mathcal{R}(G) = \mathcal{H}$, we have that

$$\|(B^* - B)Gy\| \leq 2bM_3^{-1/2}\|y\| \quad \text{for} \quad y \in \mathcal{V}. \tag{4.25}$$

Note that for $x, y \in \mathcal{V}$,

$$\begin{aligned} 2\mathrm{Re}\,(T\{x,y\},\{x,y\}) &= 2\mathrm{Re}\,((B-1)Gx,y) \\ &= \mathrm{Re}\,((B-B^*)Gx,y) + \mathrm{Re}\,((B+B^*-2)Gx,y) \\ &= \mathrm{Re}\,((B-B^*)Gx,y) + 2a\mathrm{Re}\,(G^{-1}x,y) \end{aligned}$$

since $B + B^* = 2 + 2aG^{-2}$ by Theorem 2.8.12. (4.24) and (4.25) imply

$$\begin{aligned} |\mathrm{Re}\,(T\{x,y\},\{x,y\})| &\le bM_3^{-1/2}\|x\|\|y\| + |a|M_1M_3^{-1/2}\|x\|\|y\| \\ &= 2c\|x\|\|y\| \qquad \text{where } c = (b+|a|M_1)/(2M_3^{1/2}) \\ &\le c\|\{x,y\}\|^2 \\ |\mathrm{Re}\,((T-\lambda)\{x,y\},\{x,y\})| &\ge (|\lambda|-c)\|\{x,y\}\|^2 \end{aligned}$$

$$\|(T-\lambda)\{x,y\}\| \ge (|\lambda|-c)\|\{x,y\}\| \quad \text{for} \quad \{x,y\} \in \mathcal{D}(T),\ \lambda \in \mathbb{R}, |\lambda| > c. \tag{4.26}$$

If $\lambda \in \mathbb{R}$ and $x \in \mathcal{V}$, then Theorems 2.3.5 and 2.8.12 imply

$$0 \le \|G^{1/2}x - |\lambda|G^{-1/2}x\|^2 = ((G+\lambda^2G^{-1})x,x) - 2|\lambda|\|x\|^2. \tag{4.27}$$

For $\lambda \in \mathbb{R}$ define $P_\lambda = BG + \lambda^2G^{-1}$ with $\mathcal{D}(P_\lambda) = \mathcal{V}$ and note that

$$2P_\lambda = 2G + 2\lambda^2G^{-1} + (B+B^*-2)G + (B-B^*)G.$$

$B + B^* = 2 + 2aG^{-2}$ implies

$$P_\lambda = G + \lambda^2G^{-1} + aG^{-1} + \tfrac{1}{2}(B-B^*)G; \tag{4.28}$$

hence (4.27), (4.24) and (4.25) imply

$$\mathrm{Re}\,(P_\lambda x,x) \ge 2(|\lambda|-c)\|x\|^2 \quad \text{for} \quad \lambda \in \mathbb{R},\ x \in \mathcal{V}.$$

Using (4.28) and the fact that G is m-accretive gives that

$$P_\lambda + \mu = (1 + ((\lambda^2+a)G^{-1} + \tfrac{1}{2}(B-B^*)G)(G+\mu)^{-1})(G+\mu) \quad \text{for} \quad \mu > 0.$$

Hence, (4.25) and Theorems 2.3.2, 1.6.8 imply that $\mathcal{R}(P_\lambda+\mu) = \mathcal{H}$ for μ large enough and therefore $P_\lambda - 2(|\lambda|-c)$ is m-accretive and, in particular, $0 \in \rho(P_\lambda)$ when $\lambda \in \mathbb{R}$, $|\lambda| > c$. Therefore, if $\{f,g\} \in Y$, $\lambda \in \mathbb{R}$, $|\lambda| > c$, we can define

$$x = P_\lambda^{-1}(g - \lambda G^{-1}f), \quad y = -G^{-1}(\lambda x + f)$$

and note that $\{x,y\} \in \mathcal{D}(T)$ and $(T-\lambda)\{x,y\} = \{f,g\}$. This and (4.26) imply that

$$\|(T-\lambda)^{-1}\| \le (|\lambda|-c)^{-1} \quad \text{for} \quad |\lambda| > c,\ \lambda \in \mathbb{R}$$

and therefore Theorem 4.3.5 implies that both T and $-T$ are generators of strongly continuous semigroups on Y. □

Let $\{R(t)\}_{t\geq 0}$ be the strongly continuous semigroup on Y whose generator is $-T$.

Theorem 4.4.5 *If $x \in \mathcal{V}$, $y \in \mathcal{H}$ and $\{u(t), v(t)\} = Q(t)\{x, y\}$ for $t \geq 0$, then $R(t)\{x, G^{-1}y\} = \{u(t), G^{-1}v(t)\}$ for $t \geq 0$.*

PROOF (3) of Theorem 4.3.1 and Theorem 4.2.8 imply that $x - u = -\int_0^t v$, $y - v = A\int_0^t u$. Hence $u' = Gw$, where $w \equiv G^{-1}v$ and $G^{-1}y - w = BG\int_0^t u$. Since u is continuous in $\mathcal{V}$, Theorems 2.8.12 and 4.2.10 imply $G^{-1}y - w = \int_0^t BGu$ and thus (8) of Theorem 4.3.1 implies the assertion of the Theorem. □

Theorem 4.4.6 *If $x \in \mathcal{V}$, $y \in \mathcal{H}$ and $\{u(t), w(t)\} = R(t)\{x, G^{-1}y\}$ for $t \geq 0$, then this u is the unique $u \in C([0,\infty), \mathcal{V}) \cap C^1([0,\infty), \mathcal{H})$ which satisfies $u(0) = x$, $u'(0) = y$ and*

$$\frac{d^2}{dt^2}(u(t), z) = -\mathfrak{F}(u(t), z) \quad \text{for} \quad t \geq 0,\ z \in \mathcal{V}.$$

PROOF Theorem 4.4.5 implies that $u \in C([0,\infty), \mathcal{V})$. (4) of Theorem 4.3.1 implies that $u' = Gw$, $w' = -BGu$ and hence for every $z \in \mathcal{V}$, we have that $(u, z)' = (u', z) = (Gw, z) = (w, Gz)$. Thus, (u, z) is actually twice differentiable and

$$(u, z)'' = (w', Gz) = -(BGu, Gz) = -\mathfrak{F}(u, z).$$

To show the uniqueness, let $w = G^{-1}u'$ and note that $w \in C([0,\infty), \mathcal{V})$, $u' = Gw$ and that for every $z \in \mathcal{V}$ we have $(Gw, z)' = (w, Gz)' = -(BGu, Gz)$. Hence

$$(w(t), Gz) - (w(0), Gz) = -\int_0^t (BGu, Gz)$$

and, since the range of G is $\mathcal{H}$, we have that $w(t) - w(0) = -\int_0^t BGu$. Therefore, $w' = -BGu$ and hence (8) of Theorem 4.3.1 implies the uniqueness. □

4.5 Sectorial Operators and Analytic Semigroups

Theorem 2.8.2 implies that *any* operator A associated with a sectorial sesquilinear form satisfies the assumptions of the Hille-Yosida Theorem 4.3.5 and hence defines a semigroup that can be denoted by e^{-At}. Since such operators A generalize elliptic operators (Section 3.7), we can say that

$$u_t + Au = 0$$

generalizes **parabolic equations** and $e^{-At}u(0)$ is the solution of the initial value problem. However, more can be said in this case. In view of Figure 2.1 and Theorem 1.6.16, it is easy to see that $-\zeta A$ is also a generator of a strongly continuous

semigroup when $\zeta \in \mathbb{C}$ and $|\arg(\zeta)|$ is small enough. Hence, it is natural to ask: Is $e^{-A\zeta}$ an analytic function of ζ? Note that semigroups need not be differentiable in general (Example 4.1.3). We shall prove that the answer is yes in this case. Furthermore, we shall show that the analyticity also implies smoothing of solutions, which was observed in Example 4.1.6 and is a characteristic of parabolic equations.

To avoid being constrained to Hilbert spaces, we shall first define a bit larger class of operators for which this analysis can be done. We shall then show some properties of these operators followed by a direct construction of $e^{-A\zeta}$ and its analysis. Complex Banach spaces will be used in this section. In the next section it is shown, Corollary 4.6.2, how to proceed when real function spaces have to be used.

Let $\arg(0) = 0$ and, if $\zeta \in \mathbb{C}\backslash\{0\}$, let $\arg(\zeta) \in [-\pi, \pi)$ be such that $\zeta = |\zeta| e^{i \arg(\zeta)}$.

Definition 4.5.1 For $a \in \mathbb{R}$, $M \in (0, \infty)$, $\theta \in (0, \pi/2)$ and a complex Banach space X, define $\mathfrak{A}(a, M, \theta, X)$ to be the collection of all densely defined linear operators A in X which have the property that

$$\lambda \in \rho(A) \quad \text{and} \quad \|(A - \lambda)^{-1}\| \leq M/|\lambda - a|$$

whenever $\lambda \in \mathbb{C}$ and $|\arg(\lambda - a)| \geq \theta$. A is said to be a **sectorial operator** if A belongs to some $\mathfrak{A}(a, M, \theta, X)$.

EXAMPLE 4.5.2 If A is the linear operator associated with a sectorial sesquilinear form on a complex Hilbert space, then Theorem 2.8.2 implies that

$$\|(A - \lambda)^{-1}\| \leq \begin{cases} \frac{1}{|\lambda - a|} & \text{if } |\arg(\lambda - a)| \geq \theta + \pi/2 \\ \frac{1}{|\lambda - a| \sin(|\arg(\lambda - a)| - \theta)} & \text{if } \theta < |\arg(\lambda - a)| < \theta + \pi/2 \end{cases}$$

where $\theta = \tan^{-1}(M_2/M_3)$. Hence, A is a sectorial operator.

EXAMPLE 4.5.3 Consider $-\overline{\Delta}$ in $W_0^{0,p}(\mathbb{R}^n)$, $1 \leq p \leq \infty$, see Example 4.3.10. Observe that (4.20), (3.7) and Lemma 3.1.5 imply that $-\overline{\Delta}$ is a sectorial operator.

EXAMPLE 4.5.4 For $f \in C_\ell$, define $\ell(f) = \lim_{|x| \to \infty} f(x)$. Define

$$\Delta_\ell f = \overline{\Delta}(f - \ell(f)) \quad \text{for} \quad f \in \mathcal{D}(\Delta_\ell) \equiv \{f \in C_\ell \mid f - \ell(f) \in \mathcal{D}(\overline{\Delta})\},$$

where $\overline{\Delta}$ is defined as in Example 4.3.10 when $p = \infty$.

Observe (Exercise 7) that Δ_ℓ is densely defined, $\sigma(\Delta_\ell) = (-\infty, 0]$ and

$$(\lambda - \Delta_\ell)^{-1} f = G_\lambda * f \quad \text{for} \quad f \in C_\ell, \ \lambda \in \mathbb{C}\backslash(-\infty, 0]$$

where G_λ is given by (3.6). This implies, as in Example 4.5.3, that $-\Delta_\ell$ is a sectorial operator in C_ℓ and that Δ_ℓ is the generator of a contraction semigroup on C_ℓ.

Lemma 4.5.5 *If A is a sectorial operator and if for some $\lambda \in \mathbb{R}$ the line $\mathrm{Re}\,\zeta = \lambda$, $\zeta \in \mathbb{C}$, lies in the resolvent set of A, then for some $\delta > 0$, the strip $|\lambda - \mathrm{Re}\,\zeta| < \delta$, $\zeta \in \mathbb{C}$, also lies in the resolvent set of A.*

PROOF Suppose $A \in \mathfrak{A}(a, M, \theta, X)$. If there would be no such $\delta > 0$, then there would exist a sequence $\{\zeta_k\}$ in $\sigma(A)$ such that $\operatorname{Re}\zeta_k \to \lambda$ and, since $|\operatorname{Im}\zeta_k| \leq (\operatorname{Re}\zeta_k - a)\tan\theta$, we see that a subsequence of $\{\zeta_k\}$ should converge to some ζ with $\operatorname{Re}\zeta = \lambda$ - this is not possible since $\sigma(A)$ is closed. □

Theorem 4.5.6 *Suppose that A is a sectorial operator in X and that for some $a \in \mathbb{R}$,*

$$a < \operatorname{Re}\lambda \quad \textit{for all} \quad \lambda \in \sigma(A).$$

Then $A \in \mathfrak{A}(a, M, \theta, X)$ for some $M \in (0, \infty)$, $\theta \in (0, \pi/2)$.

PROOF Suppose $A \in \mathfrak{A}(a_1, M_1, \theta_1, X)$. If $a \leq a_1$, then

$$|\lambda - a_1| \geq |\lambda - a| \sin\theta_1 \quad \text{when} \quad |\arg(\lambda - a)| \geq \theta_1;$$

hence $A \in \mathfrak{A}(a, M_1/\sin\theta_1, \theta_1, X)$.

If $a > a_1$, then, by Lemma 4.5.5, there exists $b > a$ such that $b < \operatorname{Re}\lambda$ for all $\lambda \in \sigma(A)$. Define θ by $(b-a)\tan\theta = (b-a_1)\tan\theta_1$ and note that

$$S \equiv \{\zeta \in \mathbb{C} \mid |\zeta - a_1| \leq (b-a_1)/\cos\theta_1, \operatorname{Re}\zeta \leq b\} \subset \rho(A).$$

Hence, $|\zeta - a|\|(A-\zeta)^{-1}\|$ is a continuous function of $\zeta \in S$ and therefore has a finite maximum c_1 in S. Thus

$$\|(A-\lambda)^{-1}\| \leq c_1/|\lambda - a| \quad \text{for} \quad \lambda \in S.$$

If $\lambda \notin S$ and $|\arg(\lambda - a)| \geq \theta$, then $|\lambda - a| \leq |\lambda - a_1| + a - a_1 \leq 2|\lambda - a_1|$. Hence choosing $M = \max\{c_1, 2M_1\}$ will do it. □

Theorem 4.5.7 *Suppose $A \in \mathfrak{A}(a, M, \theta, X)$, $B : \mathcal{D}(B) \supset \mathcal{D}(A) \to X$ is linear and*

$$\|Bx\| \leq \varepsilon\|Ax\| + c\|x\| \quad \textit{for} \quad x \in \mathcal{D}(A),$$

where $0 \leq \varepsilon < 1/(1+M)$ and $c < \infty$. Then, $A + B$ is a sectorial operator in X.

PROOF Choose r so that $\mu \equiv \varepsilon(1+M) + (\varepsilon|a| + c)M/r < 1$. If $|\lambda - a| \geq r$ and $|\arg(\lambda - a)| \geq \theta$, then

$$\begin{aligned}
\|B(A-\lambda)^{-1}\| &\leq \varepsilon\|A(A-\lambda)^{-1}\| + c\|(A-\lambda)^{-1}\| \\
&\leq \varepsilon(1 + \|\lambda(A-\lambda)^{-1}\|) + c\|(A-\lambda)^{-1}\| \\
&\leq \varepsilon(1+M) + (\varepsilon|a| + c)M/|\lambda - a| \leq \mu \\
(A+B-\lambda)^{-1} &= (A-\lambda)^{-1}(1 + B(A-\lambda)^{-1})^{-1} \\
\|(A+B-\lambda)^{-1}\| &\leq \frac{M}{(1-\mu)|\lambda - a|};
\end{aligned}$$

hence $A + B \in \mathfrak{A}(a - r/\sin\theta, M/(1-\mu), \theta, X)$. □

EXAMPLE 4.5.8 Let $\overline{\Delta}$ be defined as in Example 4.3.10 in $X \equiv W_0^{0,p}(\mathbb{R}^n)$, $1 \le p \le \infty$. Note that (3.16) implies that for some $c < \infty$,

$$\|\mathsf{D}_k u\|_p \le c\|\overline{\Delta}u\|_p^{1/2}\|u\|_p^{1/2} \le \varepsilon\|\overline{\Delta}u\|_p + c^2/(4\varepsilon)\|u\|_p \quad \text{for} \quad k = 1, \dots, n,$$

for every $\varepsilon > 0$ and every $u \in \mathcal{S}$ - and hence, by Theorem 3.4.3, also for every $u \in \mathcal{D}(\overline{\Delta})$. Theorem 4.5.7 implies that

$$Tu = -\overline{\Delta}u + \sum_{i=1}^{n} a_i \mathsf{D}_i u + a_0 u \quad \text{for} \quad u \in \mathcal{D}(T) \equiv \mathcal{D}(\overline{\Delta})$$

is a sectorial operator in X, provided that $a_i \in L^\infty(\mathbb{R}^n)$. When $p = \infty$, we require also that a_i are continuous.

For $b \in \mathbb{R}$, $\varphi \in (0, \pi/2)$, define a path γ in $\mathbb{C}$ by

$$\gamma(t) = b + |t|\cos\varphi - i\,t\,\sin\varphi \quad \text{for} \quad t \in \mathbb{R}. \tag{4.29}$$

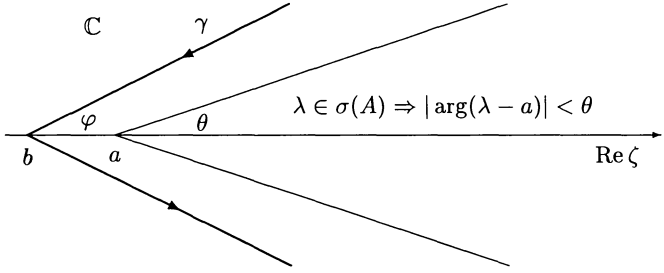


Figure 4.1: Integration path γ.

Let $\lhd_T$ denote the closed path consisting of straight lines from $\gamma(0)$ to $\gamma(T)$ to $\gamma(-T)$ to $\gamma(0)$. The Cauchy formula implies that for any polynomial P and any $\zeta \in \mathbb{C}$,

$$\frac{1}{2\pi i}\int_{\lhd_T} \frac{P(\lambda)e^{-\lambda\zeta}}{\lambda - z} d\lambda = \begin{cases} 0 & \text{if } z \text{ is outside } \lhd_T \\ P(z)e^{-\zeta z} & \text{if } z \text{ is inside } \lhd_T \end{cases}.$$

If $\zeta \neq 0$ and $\varphi + |\arg(\zeta)| < \pi/2$, then

$$\mathrm{Re}\,(\gamma(t)\zeta) \ge b\mathrm{Re}\,\zeta + |t\zeta|\cos(\varphi + |\arg(\zeta)|) \quad \text{for} \quad t \in \mathbb{R}, \tag{4.30}$$

which implies that we can take the limit $T \to \infty$ to obtain for $z \in \mathbb{C}\backslash\{b\}$,

$$\frac{1}{2\pi i}\int_{-\infty}^{\infty} \frac{P(\gamma(t))e^{-\gamma(t)\zeta}}{\gamma(t)-z}\gamma'(t)dt = \begin{cases} 0 & \text{if } |\arg(z-b)| > \varphi \\ P(z)e^{-\zeta z} & \text{if } |\arg(z-b)| < \varphi \end{cases}. \tag{4.31}$$

This motivates the following:

Definition 4.5.9 For $A \in \mathfrak{A}(a, M, \theta, X)$ and $\zeta \in \mathbb{C}$ with $|\arg(\zeta)| < \pi/2 - \theta$ define $e^{-A\zeta} \in \mathfrak{B}(X)$ as follows: $e^{-A\zeta} = 1$ if $\zeta = 0$ and if $\zeta \neq 0$, then

$$e^{-A\zeta} = \frac{1}{2\pi i}\int_{-\infty}^{\infty} (\gamma(t) - A)^{-1} e^{-\gamma(t)\zeta}\gamma'(t)dt \tag{4.32}$$

where γ is given by (4.29), see Figure 4.1, with

$$b \in (-\infty, a) \text{ and } \varphi \in (\theta, \pi/2 - |\arg(\zeta)|). \tag{4.33}$$

To justify the definition, observe first that

$$|\gamma(t) - a| \geq (a-b)\sin\varphi, \ |\arg(\gamma(t)-a)| > \varphi > \theta; \tag{4.34}$$

hence $\gamma(t) \in \rho(A)$ for $t \in \mathbb{R}$ and the integrand in (4.32) is continuous on $\mathbb{R}\backslash\{0\}$ and is therefore strongly measurable. This and (4.30) imply

$$\int_{-\infty}^{\infty} \|(\gamma(t) - A)^{-1} e^{-\gamma(t)\zeta}\gamma'(t)\|dt \leq \frac{2Me^{-b\mathrm{Re}\,\zeta}}{(a-b)|\zeta|\sin\varphi\cos(\varphi + |\arg(\zeta)|)} \tag{4.35}$$

and therefore the Bochner integral in (4.32) exists.

Let us show now that the value of the integral in (4.32) does not depend on the choice of b, φ in (4.33). Let $I(b, \varphi)$ denote the integral. Choose $\ell \in \mathfrak{B}(X)^*$ and note that

$$\ell(I(b,\varphi)) = \int_{\gamma} f(\lambda)d\lambda,$$

where $f(\lambda) = \ell((\lambda - A)^{-1})e^{-\lambda\zeta}$ for $|\arg(\lambda - a)| > \theta$. For $T > 0$ define $E(T)$ by

$$\ell(I(b,\varphi)) = \int_{-T}^{T} f(\gamma(t))\gamma'(t)dt + E(T)$$

and note that $E(T) \to 0$ as $T \to \infty$. Choose $b_1 \in (-\infty, a)$ and $\varphi_1 \in (\theta, \pi/2 - |\arg(\zeta)|)$ and let γ_1, E_1 denote the corresponding γ, E. Theorem 1.6.11 implies that f is analytic. Hence, the Cauchy Theorem gives that

$$\ell(I(b,\varphi)) - \ell(I(b_1,\varphi_1)) = E(T) - E_1(T) - F(-T) + F(T),$$

where

$$F(t) = \int_{\gamma_1(t)\to\gamma(t)} f(\lambda)d\lambda = \int_0^1 f(s\gamma(t) + (1-s)\gamma_1(t))(\gamma(t) - \gamma_1(t))ds.$$

Observe that for all $|t|$ large enough, we have that

$$|s\gamma(t) + (1-s)\gamma_1(t) - a| \geq 1$$

and hence (4.30) implies that

$$|f(s\gamma(t) + (1-s)\gamma_1(t))| \leq \|\ell\| M \exp(-c\mathrm{Re}\,\zeta - |t\zeta|\cos(\psi + |\arg(\zeta)|)),$$

where $c = \min\{b, b_1\}$, $\psi = \max\{\varphi, \varphi_1\}$. Therefore $F(t)$ converges to 0 as $|t| \to \infty$ which implies $\ell(I(b_1, \varphi_1)) = \ell(I(b, \varphi))$. Since this is true for all ℓ in $\mathfrak{B}(X)^\star$, Corollary 1.5.10 implies that $I(b, \varphi)$ does not depend on b, φ. This implies that $e^{-A\zeta}$ does not depend on the particular choice of a, M or θ either. Hence, $e^{-A\zeta}$ depends only on ζ and on the sectorial operator A.

Note that the above argument could be applied to show that many other integration paths could be used in Definition 4.5.9 without changing $e^{-A\zeta}$.

Minimization of (4.35) with respect to $b \in (-\infty, a)$ gives

Theorem 4.5.10 *If $A \in \mathfrak{A}(a, M, \theta, X)$, $\zeta \in \mathbb{C}\backslash\{0\}$ and $|\arg(\zeta)| < \pi/2 - \theta$, then*

$$\|e^{-A\zeta}\| \leq \inf_{\varphi\in(\theta,\pi/2-|\arg(\zeta)|)} \frac{Me\cos(\arg(\zeta))}{\pi\sin\varphi\cos(\varphi + |\arg(\zeta)|)}\, e^{-a\mathrm{Re}\,\zeta}.$$

This and Theorem 4.5.6 imply

Corollary 4.5.11 *If A is a sectorial operator and $a \in \mathbb{R}$ is such that $a < \mathrm{Re}\,\lambda$ for all $\lambda \in \sigma(A)$, then there exist $M < \infty$ and $\delta > 0$ such that*

$$\|e^{-A\zeta}\| \leq Me^{-a\mathrm{Re}\,\zeta} \quad \textit{for all} \quad \zeta \in \mathbb{C} \textit{ with } |\arg(\zeta)| < \delta.$$

Theorem 4.5.12 *If $A \in \mathfrak{A}(a, M, \theta, X)$, then $e^{-Aw}e^{-Az} = e^{-A(w+z)}$ for all complex numbers z and w which satisfy $|\arg(z)| < \pi/2 - \theta$, $|\arg(w)| < \pi/2 - \theta$.*

PROOF We may assume that $z \neq 0$, $w \neq 0$. Let $\alpha = \max\{|\arg(z)|, |\arg(w)|\}$. Choose $b \in (-\infty, a)$ and $\varphi \in (\theta, \pi/2 - \alpha)$ and define $\gamma(t) = b + |t|\cos\varphi - it\sin\varphi$, $\mu(t) = \gamma(t) + (a-b)/2$ for $t \in \mathbb{R}$. Choose $\ell \in \mathfrak{B}(X)^\star$ and note that

$$(2\pi i)^2\ell(e^{-Aw}e^{-Az}) = 2\pi i\int_{-\infty}^{\infty} \ell(e^{-Aw}(\gamma(t) - A)^{-1})e^{-\gamma(t)z}\gamma'(t)dt$$

$$= \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} ds\, \ell((\mu(s) - A)^{-1}(\gamma(t) - A)^{-1})e^{-\gamma(t)z - \mu(s)w}\gamma'(t)\mu'(s).$$

Using the resolvent identity,

$$(\mu - A)^{-1} - (\gamma - A)^{-1} = (\gamma - \mu)(\mu - A)^{-1}(\gamma - A)^{-1},$$

and the Fubini Theorem gives

$$\begin{aligned}(2\pi i)^2 \ell(\mathrm{e}^{-Aw}\mathrm{e}^{-Az}) = \\ \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\frac{e^{-\mu(s)w}\mu'(s)ds}{\mu(s)-\gamma(t)}\right)\ell((\gamma(t)-A)^{-1})e^{-\gamma(t)z}\gamma'(t)dt \\ + \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\frac{e^{-\gamma(t)z}\gamma'(t)dt}{\gamma(t)-\mu(s)}\right)\ell((\mu(s)-A)^{-1})e^{-\mu(s)w}\mu'(s)ds.\end{aligned}$$

The Cauchy formula (4.31) implies

$$\begin{aligned}(2\pi i)^2\ell(e^{-Aw}e^{-Az}) &= 2\pi i\int_{-\infty}^{\infty}\ell((\mu(s)-A)^{-1})e^{-\mu(s)(w+z)}\mu'(s)ds \\ &= (2\pi i)^2\ell(e^{-A(w+z)})\end{aligned}$$

and, since this is true for all $\ell \in \mathfrak{B}(X)^\star$, we have that $e^{-Aw}e^{-Az} = e^{-A(w+z)}$ by Corollary 1.5.10. □

This proves the semigroup property of $e^{-A\zeta}$. Analyticity of $e^{-A\zeta}$ follows from

Theorem 4.5.13 *If $A \in \mathfrak{A}(a, M, \theta, X)$, $\zeta \in \mathbb{C}\backslash\{0\}$, $|\arg(\zeta)| < \pi/2 - \theta$, then for $n \geq 0$,*

$$\left(\frac{d}{d\zeta}\right)^n e^{-A\zeta} = \frac{1}{2\pi i}\int_{-\infty}^{\infty}(-\gamma(t))^n(\gamma(t)-A)^{-1}e^{-\gamma(t)\zeta}\gamma'(t)dt$$

where γ is as in Definition 4.5.9 of $e^{-A\zeta}$.

PROOF The integral, denoted by $I_n(\zeta)$, clearly exists. If $h \in \mathbb{C}$ and $|h| < |\zeta|\cos(\varphi + |\arg(\zeta)|)$, then $\varphi + |\arg(\zeta + h)| < \pi/2$ and $\zeta + h \neq 0$; hence the path γ can be used also to evaluate $I_n(\zeta + h)$. Since

$$\begin{aligned}|e^{-\gamma(t)(\zeta+h)} - \mathrm{e}^{-\gamma(t)\zeta} + h\gamma(t)e^{-\gamma(t)\zeta}| \leq \\ |h\gamma(t)|^2\exp\{-b\mathrm{Re}\,\zeta + |bh| - |t\zeta|\cos(\varphi + |\arg(\zeta)|) + |th|\},\end{aligned}$$

we have that

$$\lim_{h\to 0}\left\|\frac{1}{h}(I_n(\zeta+h) - I_n(\zeta)) - I_{n+1}(\zeta)\right\| = 0.$$

□

The smoothing property of $e^{-A\zeta}$ is the consequence of the following:

Theorem 4.5.14 *If* $A \in \mathfrak{A}(a, M, \theta, X)$, $\zeta \in \mathbb{C}\backslash\{0\}$, $|\arg(\zeta)| < \pi/2 - \theta$, *then for* $n \geq 0$,

(1) the range of $e^{-A\zeta}$ *is contained in* $\mathcal{D}(A^n)$

(2) $A^n e^{-A\zeta} = (-\frac{d}{d\zeta})^n e^{-A\zeta}$

(3) $A^n e^{-A\zeta} x = e^{-A\zeta} A^n x$ *for* $x \in \mathcal{D}(A^n)$.

PROOF Statements (1) and (2) are obviously true for $n = 0$. If (1) and (2) are true for some n, then Theorem 4.5.13 implies that for all $x \in X$ we have

$$A^n e^{-A\zeta} x = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \gamma(t)^n (\gamma(t) - A)^{-1} x e^{-\gamma(t)\zeta} \gamma'(t) dt.$$

Since A is closed and $A(\gamma(t) - A)^{-1} x = -x + \gamma(t)(\gamma(t) - A)^{-1} x$, Theorem 4.2.10 implies that $A^n e^{-A\zeta} x \in \mathcal{D}(A)$ and

$$A^{n+1} e^{-A\zeta} x = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \gamma(t)^n (-x + \gamma(t)(\gamma(t) - A)^{-1} x) e^{-\gamma(t)\zeta} \gamma'(t) dt.$$

The Cauchy formula (4.31) implies

$$A^{n+1} e^{-A\zeta} x = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \gamma(t)^{n+1} (\gamma(t) - A)^{-1} x e^{-\gamma(t)\zeta} \gamma'(t) dt.$$

Therefore, (1) and (2) are true for $n + 1$ and hence for all $n \geq 0$.

Assuming (3) for some $n \geq 0$ implies that for $x \in \mathcal{D}(A^n)$,

$$A^n e^{-A\zeta} x = \frac{1}{2\pi i} \int_{-\infty}^{\infty} (\gamma(t) - A)^{-1} A^n x \, e^{-\gamma(t)\zeta} \gamma'(t) dt.$$

Hence, if $x \in \mathcal{D}(A^{n+1})$, then Theorem 4.2.10 implies (3) for $n + 1$. □

Theorem 4.5.15 *Suppose* $A \in \mathfrak{A}(a, M, \theta, X)$ *and* $\varepsilon \in (0, \pi/2 - \theta)$. *Define*

$$S = \{\zeta \in \mathbb{C}\backslash\{0\} \mid |\arg(\zeta)| < \pi/2 - \theta - \varepsilon\}.$$

Then,

(1) $\lim_{\zeta \to 0, \zeta \in S} \|e^{-A\zeta} x - x\| = 0$ *for each* $x \in X$

(2) $\lim_{\zeta \to 0, \zeta \in S} \|\frac{1}{\zeta}(x - e^{-A\zeta} x) - Ax\| = 0$ *for each* $x \in \mathcal{D}(A)$

(3) $\{e^{-At}\}_{t\geq 0}$ *is a strongly continuous semigroup whose generator is* $-A$.

PROOF Choose $y \in \mathcal{D}(A)$, $\zeta \in S$. If γ is as in Definition 4.5.9 , then

$$\begin{aligned} e^{-A\zeta}y - e^{-a\zeta}y &= \frac{1}{2\pi i}\int_{-\infty}^{\infty} \left((\gamma(t)-A)^{-1}y - (\gamma(t)-a)^{-1}y\right) e^{-\gamma(t)\zeta}\gamma'(t)dt \\ &= \frac{1}{2\pi i}\int_{-\infty}^{\infty} (\gamma(t)-a)^{-1}(\gamma(t)-A)^{-1}(Ay-ay)e^{-\gamma(t)\zeta}\gamma'(t)dt. \end{aligned}$$

Using (4.34) and (4.30), with $b = a - 1/|\zeta|$ and $\varphi \to \theta$, gives

$$\|e^{-A\zeta}y - e^{-a\zeta}y\| \leq |\zeta| \frac{Me\|Ay-ay\|}{\pi(\sin\theta)^2 \sin\varepsilon} e^{-a\mathrm{Re}\,\zeta}.$$

This, the bound in Theorem 4.5.10, the fact that

$$\|e^{-A\zeta}x - x\| \leq \|e^{-A\zeta}y - e^{-a\zeta}y\| + |e^{-a\zeta}-1|\|y\| + \|e^{-A\zeta}-1\|\|x-y\|$$

for every $x \in X$, and that $\mathcal{D}(A)$ is dense in X imply (1).

To show (2), observe that for $\zeta \in S$, $x \in \mathcal{D}(A)$, Theorem 4.5.14 implies $\frac{d}{dt}e^{-A\zeta t}x = -\zeta e^{-A\zeta t}Ax$. Hence, (1) and Theorem 4.2.11 imply

$$\begin{aligned} x - e^{-A\zeta}x &= \zeta\int_0^1 e^{-A\zeta t}Ax\,dt \\ \frac{1}{\zeta}(x - e^{-A\zeta}x) - Ax &= \int_0^1 (e^{-A\zeta t}Ax - Ax)dt. \end{aligned}$$

This and (1) imply assertion (2).

Let $-B$ be the generator of the semigroup $\{e^{-At}\}_{t\geq 0}$. (2) implies that B is an extension of A. Theorems 4.5.10 and 4.3.2 imply $a - 1 \in \rho(A) \cap \rho(B)$ and hence $A = B$ by Lemma 1.6.14. □

EXAMPLE 4.5.16 Consider the initial value problem for the **parabolic equation**:

$$\frac{\partial u}{\partial t} = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial}{\partial x_j}\left(a_{ij}(x)\frac{\partial u}{\partial x_i}\right) - \sum_{i=1}^{n} a_i(x)\frac{\partial u}{\partial x_i} - a_0(x)u \quad \text{for} \quad t > 0,\ x \in \Omega$$

$$u(x,0) = f(x) \quad \text{for} \quad x \in \Omega$$

$$u(x,t) = 0 \quad \text{for} \quad t \geq 0,\ x \in \partial\Omega.$$

Assume that Ω is a nonempty open subset of $\mathbb{R}^n$, that a_{ij}, a_i are bounded, complex valued, measurable functions on Ω such that $\mathrm{Im}\,a_{ij} = \mathrm{Im}\,a_{ji}$ and that the strong ellipticity condition (3.62) holds with some $\delta > 0$. Define the sectorial form by (3.61) on

$\mathcal{V} = W_0^1(\Omega)$, $\mathcal{H} = L^2(\Omega)$ and let A be the sectorial operator associated with it. The generalized solution of the parabolic equation is $u(x,t) = (e^{-At}f)(x)$ for $f \in L^2(\Omega)$.

Theorem 4.5.13 implies that $t \to u(\cdot,t)$ is an analytic, Hilbert-space-valued, function. Hence, $t \to (u(\cdot,t),v)$ is a complex valued analytic function for any $v \in L^2(\Omega)$. Note that when v is a characteristic function of a compact set $C \subset \Omega$, then $(u(\cdot,t),v)$ represents the average of $u(\cdot,t)$ in C multiplied by the volume of C.

When $a_{ij}, a_0 \in C^1(\Omega)$, Theorem 3.8.2 implies that $w \in \mathcal{D}(A)$ iff $w \in W_0^1(\Omega) \cap W_{loc}^2(\Omega)$ and

$$Aw \equiv -\sum_{i=1}^{n}\sum_{j=1}^{n} \mathsf{D}_j(a_{ij}\mathsf{D}_i w) + \sum_{i=1}^{n} a_i \mathsf{D}_i w + a_0 w \in L^2(\Omega).$$

It is easy to apply the induction argument to show that $\mathcal{D}(A^m) \subset W_{loc}^{2m}(\Omega)$ when $m \geq 1$ and $a_{ij}, a_i \in C^{2m-1}(\Omega)$. Corollary 3.6.4 implies that $W_{loc}^{2m}(\Omega) \subset C^j(\Omega)$ if $2m - n/2 > j \geq 0$ and Theorem 4.5.14 implies that $u(\cdot,t) \in \cap_{k=1}^{\infty}\mathcal{D}(A^k)$ for $t > 0$. Thus, if $a_{ij}, a_i \in C^\infty(\Omega)$, then, for $t > 0$, we have that $u(\cdot,t) \in C^\infty(\Omega)$ for any $f \in L^2(\Omega)$ and therefore e^{-At} is a smoothing operator.

Corollary 4.5.11 implies that if the spectrum of A lies in the half-plane $\mathrm{Re}\,\zeta > 0$, then u decays exponentially in t. On the other hand, Corollary 4.3.3 implies that if the spectrum of A contains any part of the half-plane $\mathrm{Re}\,\zeta < 0$, then u is unbounded in t for some $f \in L^2(\Omega)$. This stability-instability result is true for any sectorial operator.

Let B be a bounded operator on a complex Banach space X. Since B is a sectorial operator (Exercise 12), we have $e^{-B\zeta}$ defined for ζ in a sector. How does $e^{-B\zeta}$ relate to $Q(\zeta)$ which is defined in Example 4.1.4 for every $\zeta \in \mathbb{C}$? Since $\{Q(t)\}_{t\geq 0}$ is a strongly continuous semigroup whose generator is also $-B$, Theorems 4.3.1 and 4.5.15 imply that $Q(t) = e^{-Bt}$ for $t \geq 0$. $\ell(Q(\cdot)x)$ and $\ell(e^{-B\cdot}x)$ are complex-valued analytic functions in the sector and, since they are equal on $(0,\infty)$, they are equal in the whole sector. Moreover, since this is true for every $x \in X$ and $\ell \in X^\star$, we have that $Q(\zeta) = e^{-B\zeta}$ for all ζ in the sector. This also enables us to extend the definition of $e^{-B\zeta}$ to every $\zeta \in \mathbb{C}$, by

$$e^{-B\zeta} = \sum_{k=0}^{\infty} \frac{(-\zeta)^k}{k!} B^k \quad \text{for} \quad \zeta \in \mathbb{C}. \tag{4.36}$$

We shall often need the following bound on $Ae^{-A\zeta}$

Theorem 4.5.17 *If $A \in \mathfrak{A}(a, M, \theta, X)$, $\zeta \in \mathbb{C}\backslash\{0\}$, $|\arg(\zeta)| < \pi/2 - \theta$, then*

$$\|(A-a)e^{-A\zeta}\| \leq \frac{M}{\pi\cos(\theta + |\arg(\zeta)|)} \frac{e^{-a\mathrm{Re}\,\zeta}}{|\zeta|}.$$

PROOF Theorems 4.5.13 and 4.5.14, with γ as in Definition 4.5.9, imply that

$$\begin{aligned}(A-a)e^{-A\zeta} &= \frac{1}{2\pi i}\int_{-\infty}^{\infty}(\gamma(t)-a)(\gamma(t)-A)^{-1}e^{-\gamma(t)\zeta}\gamma'(t)dt\\ \|(A-a)e^{-A\zeta}\| &\le \frac{M}{2\pi}\int_{-\infty}^{\infty}|e^{-\gamma(t)\zeta}|dt\\ &\le \frac{M}{\pi|\zeta|\cos(\varphi+|\arg(\zeta)|)}\,e^{-b\mathrm{Re}\,\zeta}.\end{aligned}$$

Letting $b \to a$ and $\varphi \to \theta$ gives the bound. □

Corollary 4.5.18 *If A is a sectorial operator and $\tau \in (0,\infty)$, then for some $c < \infty$,*

$$\|Ae^{-At}\| \le ct^{-1} \quad \textit{for} \quad 0 < t \le \tau, \tag{4.37}$$

$$\|e^{-At} - e^{-As}\| \le cs^{-1}(t-s) \quad \textit{for} \quad 0 < s \le t \le \tau, \tag{4.38}$$

$$\|Ae^{-At} - Ae^{-As}\| \le ct^{-1}s^{-1}(t-s) \quad \textit{for} \quad 0 < s \le t \le \tau. \tag{4.39}$$

PROOF Theorems 4.5.10 and 4.5.17 imply (4.37). Theorem 4.5.14 implies

$$\|e^{-At} - e^{-As}\| \le \int_s^t \|Ae^{-Ar}\|dr$$

and hence (4.37) implies (4.38). Since

$$\|Ae^{-At} - Ae^{-As}\| \le \int_s^t \|A^2e^{-Ar}\|dr \le \int_s^t \|Ae^{-Ar/2}\|^2 dr,$$

(4.37) implies (4.39). □

The bound in (4.37) is a defining characteristic of analytic semigroups. This is due to the fact that if $-A$ is the generator of a strongly continuous semigroup $\{Q(t)\}_{t\ge 0}$, such that $t\|AQ(t)\|$ is bounded for $0 < t < 1$, then A is a sectorial operator. The proof of this fact consists of the justification of the following representation:

$$e^{-A\zeta} = \sum_{n=0}^{\infty}\frac{(t-\zeta)^n}{n!}(Ae^{-At/n})^n.$$

We shall not need this result, so no more details will be presented.

Theorem 4.5.19 *Let $\{Q(\zeta)\}_{\zeta\in S}$ be a family of bounded operators in a complex Banach space X where $S = \{\zeta \in \mathbb{C} \mid |\arg(\zeta)| < \beta, \zeta \ne 0\}$ and $\beta \in (0,\pi/2)$. Assume that*

(a) for some $M < \infty$ *and* $a \in \mathbb{R}$, *we have that* $\|Q(\zeta)\| \leq Me^{-a\mathrm{Re}\,\zeta}$ *for* $\zeta \in S$

(b) $\zeta \to \ell(Q(\zeta)x)$ *is an analytic function in* S *for every* $\ell \in X^\star$, $x \in X$

(c) $\{Q(t)\}_{t\geq 0}$ *is a strongly continuous semigroup whose generator is* $-A$.

Then A *is a sectorial operator in* X *and* $e^{-A\zeta} = Q(\zeta)$ *for* $\zeta \in S$.

PROOF Choose $\ell \in X^\star$, $x \in X$, $\lambda < a$ and let $f(\zeta) = \ell(Q(\zeta)x)e^{\lambda\zeta}$ for $\zeta \in S$. Since f is analytic, the Cauchy Theorem implies that

$$\int_\varepsilon^T f(t)dt + \int_{T\to Tz} f(\zeta)d\zeta = \int_\varepsilon^T f(zt)z\,dt + \int_{\varepsilon\to\varepsilon z} f(\zeta)d\zeta$$

for $0 < \varepsilon < T$ and $z \in S$. (a) implies that we can take the limits $T \to \infty$ and $\varepsilon \to 0+$ and obtain

$$\int_0^\infty f(t)dt = \int_0^\infty f(zt)z\,dt.$$

Theorem 4.3.2 then implies that for all $\ell \in X^\star$, $x \in X$, we have

$$\ell((A-\lambda)^{-1}x) = z\int_0^\infty \ell(Q(zt)x)e^{\lambda zt}dt \quad \text{for} \quad z \in S,\ \lambda < a. \tag{4.40}$$

(a) implies that for every $z \in S$ the right-hand side of (4.40), denoted by $RHS(\lambda)_z$, is an analytic function of λ in the half-plane $\mathrm{Re}\,(\lambda - a)z < 0$. (4.40) implies that $RHS(\lambda)_z = RHS(\lambda)_1$ for $\lambda \in (-\infty, a)$ and hence $RHS(\lambda)_z = RHS(\lambda)_1$ for all λ in their common domain. Thus, $RHS(\lambda)_1$ can be extended to an analytic function of λ in the region where $\mathrm{Re}\,(\lambda-a)z < 0$ for some $z \in S$, i.e. $|\arg(\lambda - a)| > \pi/2 - \beta$ and in this region we have the bound

$$\begin{aligned} |RHS(\lambda)_1| &\leq \inf_{z\in S} \frac{|z|M\|\ell\|\|x\|}{-\mathrm{Re}\,(\lambda-a)z} \\ &\leq \frac{M\|\ell\|\|x\|}{-|\lambda-a|\cos(|\arg(\lambda-a)| + \beta)}. \end{aligned} \tag{4.41}$$

Let α be the smallest nonnegative number such that if $|\arg(\lambda-a)| > \alpha$, then $\lambda \in \rho(A)$. Theorem 4.3.2 implies that $\alpha \leq \pi/2$. Since $(A-\lambda)^{-1}$ is analytic, we have that $\ell((A-\lambda)^{-1}x) = RHS(\lambda)_1$ in the sector $|\arg(\lambda-a)| > \max\{\alpha, \pi/2-\beta\}$ and hence (4.41) implies (see Exercise 11 of Chapter 1)

$$\|(A-\lambda)^{-1}\| \leq \frac{M}{-|\lambda-a|\cos(|\arg(\lambda-a)| + \beta)} \tag{4.42}$$

in the sector. If $\alpha > \pi/2 - \beta$, then Theorem 1.6.11 implies that, at each point λ in the sector $|\arg(\lambda-a)| > \alpha$, a disk of radius $-|\lambda-a|\cos(\alpha+\beta)/M$ also

lies in $\rho(A)$ and hence α could not be the smallest. Thus, $\alpha \leq \pi/2 - \beta$ and (4.42) implies that for every $\theta \in (\pi/2 - \beta, \pi/2)$ there exists $M_\theta < \infty$ such that $A \in \mathfrak{A}(a, M_\theta, \theta, X)$. $e^{-A\zeta}$ is then defined and analytic in S and, since Theorems 4.5.15 and 4.3.1 imply that $e^{-At} = Q(t)$ for $t \geq 0$, the analyticity implies

$$\ell(e^{-A\zeta}x) = \ell(Q(\zeta)x) \quad \text{for all} \quad \zeta \in S,\ \ell \in X^\star,\ x \in X.$$

Hence Corollary 1.5.10 implies that $e^{-A\zeta} = Q(\zeta)$ for $\zeta \in S$. □

4.6 Invariant Subspaces

A closed subspace Y of a Banach space X is said to be an **invariant subspace** of an operator A, defined in X, if $Ay \in Y$ for every $y \in \mathcal{D}(A) \cap Y$. In this case, the **restriction of A to Y** is defined to be an operator $A_{\downarrow Y}$ in Y given by

$$A_{\downarrow Y}y = Ay \quad \text{for} \quad y \in \mathcal{D}(A_{\downarrow Y}) \equiv \mathcal{D}(A) \cap Y.$$

Theorem 4.6.1 *Let $\{Q(t)\}_{t\geq 0}$ be a strongly continuous semigroup on a Banach space X, let $-A$ be its generator and let $a, M \in \mathbb{R}$ be such that $\|Q(t)\| \leq Me^{-at}$ for $t > 0$.*

If Y is an invariant subspace of $(A-\lambda)^{-1}$ for some $\lambda < a$, then Y is an invariant subspace of A and $Q(t)$ for all $t \geq 0$. Moreover, $\{Q(t)_{\downarrow Y}\}_{t\geq 0}$ is a strongly continuous semigroup on Y whose generator is $-A_{\downarrow Y}$.

PROOF If $|\mu - \lambda| < a - \lambda$, then Theorem 4.3.2 implies that

$$\|(\mu-\lambda)^n(A-\lambda)^{-n}\| \leq M|\mu-\lambda|^n(a-\lambda)^{-n} \quad \text{for} \quad n \geq 1$$

and, since Y is closed, (1.22) and Theorem 1.6.8 imply that Y is an invariant subspace of $(1-(\mu-\lambda)(A-\lambda)^{-1})^{-1}$. The resolvent identity,

$$(A-\mu)^{-1}(1-(\mu-\lambda)(A-\lambda)^{-1}) = (A-\lambda)^{-1},$$

implies that Y is an invariant subspace of $(A-\mu)^{-1}$ for $\mu \in (\lambda - a + \lambda, a)$ and, by induction, for all $\mu < a$.

If $t > 0$, then $-n/t < a$ for large n. Hence if $y \in Y$, then $(A + n/t)^{-n}y \in Y$ and, by Theorem 4.3.5, $Q(t)y \in Y$. Thus, Y is an invariant subspace of $Q(t)$ and $\{Q(t)_{\downarrow Y}\}_{t\geq 0}$ is obviously a strongly continuous semigroup on Y whose generator is $-A_{\downarrow Y}$. □

When working with real evolution equations, one would expect that real initial conditions would produce real solutions. To ensure this one can always try to set up the problem in a real function space. However, the preparatory analysis is often

more easily done by using a complex function space X. The real parts do not form a subspace of X; however, a complex vector space X can always be considered to be a real vector space and thus the real parts do form a *real subspace* of X and hence Theorem 4.6.1 applies. For example,

Corollary 4.6.2 *Let A be a sectorial operator in a complex Banach space X. Consider X to be a real Banach space. Let Y be a closed real subspace of X and suppose that Y is an invariant subspace of $(A+\lambda)^{-1}$ for some sufficiently large $\lambda \in \mathbb{R}$. Then Y is an invariant subspace of e^{-At} for $t \geq 0$.*

EXAMPLE 4.6.3 In Example 4.5.3 it was shown that $-\overline{\Delta}$ is a sectorial operator in $X = W_0^{0,p}(\mathbb{R}^n)$. Let X_r be the set of real valued members of X. From formula (4.20) for $(\lambda - \overline{\Delta})^{-1}$ we see that X_r is an invariant real subspace of $(\lambda - \overline{\Delta})^{-1}$, $\lambda > 0$. Corollary 4.6.2 implies that X_r is an invariant real subspace of $e^{\overline{\Delta}t}$ for $t \geq 0$ and hence $\{(e^{\overline{\Delta}t})_{\downarrow X_r}\}_{t\geq 0}$ is a strongly continuous semigroup on X_r whose generator is $\overline{\Delta}_{\downarrow X_r}$. Actually, $\{(e^{\overline{\Delta}t})_{\downarrow X_r}\}_{t\geq 0}$ is a contraction semigroup on X_r, see Example 4.3.10.

The same arguments apply to Δ_ℓ, see Example 4.5.4, when one wants to consider it in the space consisting of only real members of C_ℓ.

A bounded operator P on a Banach space is said to be a **projection** if $P^2 = P$. The following two Lemmas, Theorem 2.2.3 and Exercise 13 highlight the relationship between projections and closed subspaces.

Lemma 4.6.4 *When P is a projection on a Banach space X, then*

(a) $1 - P$ *is also a projection on* X

(b) $\mathcal{R}(P) = \mathcal{N}(1-P)$ *is a closed subspace of* X

(c) $\mathcal{R}(1-P) = \mathcal{N}(P)$ *is a closed subspace of* X

(d) $\mathcal{R}(P) \cap \mathcal{R}(1-P) = \{0\}$

(e) $X = \mathcal{R}(P) + \mathcal{R}(1-P)$

(f) $\mathcal{R}(P)$, $\mathcal{R}(1-P)$ *are invariant subspaces of B if $B \in \mathfrak{B}(X)$ and $BP = PB$.*

PROOF $(1-P)^2 = 1 - 2P + P^2 = 1 - P$ implies (a). If $x \in \mathcal{R}(P)$, then $x = Py$; hence $(1-P)x = (P-P^2)y = 0$. If $x \in \mathcal{N}(1-P)$, then $Px = x \in \mathcal{R}(P)$ and therefore $\mathcal{R}(P) = \mathcal{N}(1-P)$. Since the null space of a closed operator is closed we have (b) and, by (a), also (c). If $x \in \mathcal{R}(P) \cap \mathcal{R}(1-P)$, then $x = Px$ by (b) and $Px = 0$ by (c); hence, (d) follows. $x = Px + (1-P)x$ implies (e). If $x \in \mathcal{R}(P)$, then $Bx = BPx = PBx \in \mathcal{R}(P)$. If $x \in \mathcal{R}(1-P)$, then $Bx = B(x - Px) = Bx - PBx \in \mathcal{R}(1-P)$. □

Lemma 4.6.5 *If P is a projection on a Banach space X and A is a linear operator in X such that $(A-\lambda)^{-1}P = P(A-\lambda)^{-1}$ for some $\lambda \in \rho(A)$, then, $\mathcal{R}(P)$ and $\mathcal{R}(1-P)$ are invariant subspaces of A and $(A-\zeta)^{-1}$ for every $\zeta \in \rho(A)$.*

PROOF If $x \in \mathcal{D}(A) \cap \mathcal{R}(P)$ and $y = (A-\lambda)x$, then $x = Px = P(A-\lambda)^{-1}y = (A-\lambda)^{-1}Py$. Hence $(A-\lambda)x = Py$ and $Ax = Py + \lambda x \in \mathcal{R}(P)$.

If $x \in \mathcal{R}(P)$, $\zeta \in \rho(A)$ and $y = (A-\zeta)^{-1}x$, then

$$Py = P(A-\lambda)^{-1}(A-\lambda)y = (A-\lambda)^{-1}P(A-\lambda)y \in \mathcal{D}(A),$$

hence, $(A-\lambda)Py = P(A-\lambda)y$. Therefore, $(A-\zeta)Py = P(A-\zeta)y = Px = x$ and $Py = y \in \mathcal{R}(P)$.

This gives the invariance of $\mathcal{R}(P)$. $(A-\lambda)^{-1}(1-P) = (1-P)(A-\lambda)^{-1}$ implies the invariance of $\mathcal{R}(1-P)$. □

Recall that a **cycle** Γ is a finite collection of closed paths $\gamma_1, \ldots, \gamma_n$ in $\mathbb{C}$ and that the integration over Γ is defined by

$$\int_\Gamma f(\zeta)d\zeta = \sum_{k=1}^{n} \int_{\gamma_k} f(\zeta)d\zeta.$$

If each path γ_k lies in some open set Ω, we say that Γ is a cycle in Ω.

Let K be a nonempty compact subset of an open set Ω in $\mathbb{C}$. It is well known that a cycle Γ can be chosen in $\Omega \backslash K$ so that

$$\frac{1}{2\pi i}\int_\Gamma \frac{d\zeta}{\zeta - \lambda} = \begin{cases} 1 & \text{if } \lambda \in K \\ 0 & \text{if } \lambda \notin \Omega \end{cases}. \tag{4.43}$$

Any such cycle Γ is called a **contour that surrounds K in Ω**. If Γ' is another contour that surrounds K in Ω, then the Cauchy Theorem states that for any function f that is analytic in $\Omega \backslash K$, we have

$$\int_\Gamma f(\zeta)d\zeta = \int_{\Gamma'} f(\zeta)d\zeta. \tag{4.44}$$

A nonempty compact set σ in $\mathbb{C}$ is called a **spectral set** of a linear operator A in a complex Banach space X if $\sigma \subset \sigma(A)$ and $\sigma(A) \backslash \sigma$ is a closed set in $\mathbb{C}$. In this case define

$$P_\sigma = \frac{1}{2\pi i}\int_\Gamma (\zeta - A)^{-1}d\zeta \tag{4.45}$$

where Γ is a contour that surrounds σ in $(\sigma(A)\backslash\sigma)^c$. Since $(\sigma(A)\backslash\sigma)^c \backslash \sigma = \rho(A)$, the Bochner integral in (4.45) exists and $P_\sigma \in \mathfrak{B}(X)$. If $f(\zeta) = \ell((\zeta - A)^{-1})$, with $\ell \in \mathfrak{B}(X)^\star$, then (4.44) implies that P_σ does not depend on the particular choice of a contour Γ that surrounds σ in $(\sigma(A)\backslash\sigma)^c$. The following Theorem 4.6.6 allows us to call P_σ the **projection associated with the spectral set σ**.

Theorem 4.6.6 *Let σ be a spectral set of a linear operator A in a complex Banach space X and let P_σ be given by (4.45). Then*

(1) P_σ is a projection on X

(2) $(A-\lambda)^{-1}P_\sigma = P_\sigma(A-\lambda)^{-1}$ for every $\lambda \in \rho(A)$

(3) $\mathcal{R}(P_\sigma)$ and $\mathcal{R}(1-P_\sigma)$ are both invariant subspaces of A

(4) $\mathcal{R}(P_\sigma) \subset \mathcal{D}(A)$, $A_{\downarrow \mathcal{R}(P_\sigma)} \in \mathfrak{B}(\mathcal{R}(P_\sigma))$ and $\sigma(A_{\downarrow \mathcal{R}(P_\sigma)}) = \sigma$

(5) $\sigma(A_{\downarrow \mathcal{R}(1-P_\sigma)}) = \sigma(A)\backslash\sigma$.

PROOF Choose $\delta > 0$ such that $\sigma + \overline{B(0,\delta)} \subset (\sigma(A)\backslash\sigma)^c$. Let Γ be a contour that surrounds σ in $\sigma + B(0,\delta)$ and hence also in $(\sigma(A)\backslash\sigma)^c$. Let Γ' be a contour that surrounds $\sigma + \overline{B(0,\delta)}$ in $(\sigma(A)\backslash\sigma)^c$. Then

$$P_\sigma^2 = \frac{1}{2\pi i}\int_\Gamma P_\sigma(\zeta - A)^{-1}d\zeta = \frac{1}{(2\pi i)^2}\int_\Gamma\int_{\Gamma'}(\zeta' - A)^{-1}(\zeta - A)^{-1}d\zeta'\,d\zeta$$

and the resolvent identity implies

$$P_\sigma^2 = \frac{1}{(2\pi i)^2}\int_\Gamma\int_{\Gamma'}((\zeta' - A)^{-1} - (\zeta - A)^{-1})\frac{d\zeta'\,d\zeta}{\zeta - \zeta'}.$$

This, the Fubini Theorem and the fact that (4.43) implies

$$\frac{1}{2\pi i}\int_\Gamma \frac{d\zeta}{\zeta - \zeta'} = 0 \quad \text{for} \quad \zeta' \in \mathcal{R}(\Gamma'), \qquad \frac{1}{2\pi i}\int_{\Gamma'}\frac{d\zeta'}{\zeta' - \zeta} = 1 \quad \text{for} \quad \zeta \in \mathcal{R}(\Gamma)$$

give us that $P_\sigma^2 = P_\sigma$ - which proves (1).

(2) is obvious (Theorem 4.2.8). Lemma 4.6.5 implies (3).

Let us use abbreviations $A_1 = A_{\downarrow \mathcal{R}(P_\sigma)}$ and $A_2 = A_{\downarrow \mathcal{R}(1-P_\sigma)}$.

(4.45) and Theorem 4.2.10 imply that $\mathcal{R}(P_\sigma) \subset \mathcal{D}(A)$ and if $x \in \mathcal{R}(P_\sigma)$, then

$$\|A_1 x\| = \|AP_\sigma x\| = \left\|\frac{1}{2\pi i}\int_\Gamma A(\zeta - A)^{-1}x\,d\zeta\right\| \leq c\|x\|;$$

hence A_1 is a bounded operator on $\mathcal{R}(P_\sigma)$.

If $\lambda \in \rho(A)$, then (4) implies that $(A_2 - \lambda)(A-\lambda)^{-1}x = x$ for every $x \in \mathcal{R}(1-P_\sigma)$ and $(A-\lambda)^{-1}(A_2-\lambda)x = x$ for every $x \in \mathcal{D}(A_2)$. The same holds on $\mathcal{R}(P_\sigma)$; hence

$$\rho(A) \subset \rho(A_1) \cap \rho(A_2). \tag{4.46}$$

For $\lambda \in \sigma(A)$ define $Z_\lambda \in \mathfrak{B}(X)$ by

$$Z_\lambda = \frac{1}{2\pi i}\int_\Gamma (\zeta - A)^{-1}\frac{d\zeta}{\zeta - \lambda},$$

where Γ is a contour that surrounds σ in $(\sigma(A)\backslash\sigma)^c$. Clearly, $Z_\lambda P_\sigma = P_\sigma Z_\lambda$ which implies that $\mathcal{R}(P_\sigma)$ and $\mathcal{R}(1 - P_\sigma)$ are invariant subspaces of Z_λ. Theorem 4.2.10 and (4.43) imply

$$(A - \lambda)Z_\lambda = P_\sigma - \frac{1}{2\pi i}\int_\Gamma \frac{d\zeta}{\zeta - \lambda} = \begin{cases} P_\sigma - 1 & \text{if } \lambda \in \sigma \\ P_\sigma & \text{if } \lambda \in \sigma(A)\backslash\sigma \end{cases}.$$

If $\lambda \in \sigma(A)\backslash\sigma$ and $x \in \mathcal{R}(P_\sigma)$, then

$$x = P_\sigma x = (A_1 - \lambda)Z_\lambda x = Z_\lambda(A_1 - \lambda)x.$$

Hence $\lambda \in \rho(A_1)$, and by (4.46), $\sigma(A)\backslash\sigma \cup \rho(A) \subset \rho(A_1)$ and therefore

$$\sigma(A_1) \subset \sigma. \tag{4.47}$$

If $\lambda \in \sigma$ and $x \in \mathcal{R}(1 - P_\sigma)$, then $x = (1 - P_\sigma)x = -(A_2 - \lambda)Z_\lambda x$ and

$$Z_\lambda(A_2 - \lambda)x = (A_2 - \lambda)Z_\lambda x = -x \quad \text{for} \quad x \in \mathcal{D}(A_2).$$

Hence $\lambda \in \rho(A_2)$, and by (4.46), $\sigma \cup \rho(A) \subset \rho(A_2)$ which implies

$$\sigma(A_2) \subset \sigma(A)\backslash\sigma. \tag{4.48}$$

If $\lambda \notin \sigma(A_1) \cup \sigma(A_2)$, then $(A_1 - \lambda)^{-1}P_\sigma + (A_2 - \lambda)^{-1}(1 - P_\sigma)$ is the resolvent of A; hence $\lambda \in \rho(A)$. Thus, $\sigma(A) \subset \sigma(A_1) \cup \sigma(A_2)$, which together with (4.47) and (4.48) imply that $\sigma(A_1) = \sigma$ and $\sigma(A_2) = \sigma(A)\backslash\sigma$. □

Note that in (d) below we have estimates for both negative and positive times. See (4.36) for the definiton of $e^{-A_1 t}$ when $t < 0$.

Theorem 4.6.7 *Suppose that A is a sectorial operator in a Banach space X, that*

$$\{\zeta \in \mathbb{C} \mid \operatorname{Re}\zeta = \lambda\} \subset \rho(A) \quad \textit{for some } \lambda \in \mathbb{R}$$

and that $\sigma \equiv \{\zeta \in \sigma(A) \mid \operatorname{Re}\zeta < \lambda\}$ is not empty. Then, σ is a spectral set of A. Let P_σ be the projection associated with the spectral set σ and define $A_1 = A_{\downarrow\mathcal{R}(P_\sigma)}$, $A_2 = A_{\downarrow\mathcal{R}(1 - P_\sigma)}$. Then

(a) $P_\sigma e^{-At} = e^{-At}P_\sigma$ *for* $t \geq 0$

(b) $A_1 \in \mathfrak{B}(\mathcal{R}(P_\sigma))$ *and* $(e^{-At})_{\downarrow\mathcal{R}(P_\sigma)} = e^{-A_1 t}$ *for* $t \geq 0$

(c) A_2 is a sectorial operator in $\mathcal{R}(1-P_\sigma)$ and $(e^{-At})_{\downarrow \mathcal{R}(1-P_\sigma)} = e^{-A_2 t}$ for $t \geq 0$

(d) for some $\lambda_1 < \lambda < \lambda_2$ and $M < \infty$ we have that

$$\|e^{-A_1 t}\| \leq M e^{-\lambda_1 t} \quad \text{for} \quad t \leq 0$$

$$\|e^{-A_2 t}\| \leq M e^{-\lambda_2 t} \quad \text{for} \quad t \geq 0.$$

PROOF Since $\sigma(A)$ is closed and lies in a sector, Lemma 4.5.5 implies that σ is a spectral set. Lemma 4.5.5 and Theorem 4.6.6 imply that

$$\sigma(A_1) = \{\zeta \in \sigma(A) \mid \operatorname{Re}\zeta < \lambda_1\}, \quad \sigma(A_2) = \{\zeta \in \sigma(A) \mid \operatorname{Re}\zeta > \lambda_2\} \tag{4.49}$$

for some $\lambda_1 < \lambda < \lambda_2$.

(4.45), Theorem 4.2.8 and (5) of Theorem 4.3.1 imply (a).

Theorem 4.6.6 implies that $\mathcal{R}(P_\sigma)$ is an invariant subspace of $(A-\lambda)^{-1}$ for all $\lambda \in \rho(A)$. Hence, Theorem 4.6.1 implies that the generator of $\{(e^{-At})_{\downarrow \mathcal{R}(P_\sigma)}\}_{t\geq 0}$ is $-A_1$ and thus (b) follows. The same reasoning gives that the generator of $\{(e^{-At})_{\downarrow \mathcal{R}(1-P_\sigma)}\}_{t\geq 0}$ is $-A_2$. Hence, A_2 is densely defined and since the bound of its resolvent is no larger than the bound of the resolvent of A, we have that A_2 is sectorial, which completes the proof of (c).

(4.49) and Corollary 4.5.11 imply the bound on $e^{-A_2 t}$.

Since $\sigma(-A_1) = -\sigma(A_1)$, (4.49) and Corollary 4.5.11 imply $\|e^{-(-A_1)t}\| \leq M e^{\lambda_1 t}$ for $t \geq 0$. Since $e^{-A_1(-t)} = e^{-(-A_1)t}$ by (4.36), we have the rest of (d). □

4.7 The Inhomogeneous Problem - Part I

We shall consider the problem of finding a continuous $u : [0,\tau) \to X$ such that

$$u'(t) + Au(t) = f(t) \quad \text{for} \quad t \in (0,\tau), \quad u(0) = u_0. \tag{4.50}$$

Throughout this Section it is assumed that:

(1) X is a Banach space, $u_0 \in X$

(2) $-A$ is the generator of a strongly continuous semigroup $\{Q(t)\}_{t\geq 0}$ on X

(3) $0 < \tau \leq \infty$, $f \in L((0,t),X)$ for $t \in (0,\tau)$.

The special case when A is a sectorial operator is examined in more detail in Section 6.2.

Theorem 4.7.1 *If $u \in C([0,\tau),X)$ and it satisfies (4.50), then*

$$u(t) = Q(t)u_0 + \int_0^t Q(t-s)f(s)ds \quad \textit{for} \quad t \in [0,\tau). \tag{4.51}$$

PROOF Choose $t \in (0,\tau)$ and let $v(s) = Q(t-s)u(s)$ for $0 \le s \le t$. If $s, s+h \in (0,t]$, $h \neq 0$, then Theorem 4.3.1 implies

$$\begin{aligned}
\frac{v(s+h)-v(s)}{h} &= \frac{Q(t-s-h)-Q(t-s)}{h}u(s) + Q(t-s-h)u'(s) + \\
&\quad Q(t-s-h)\left(\frac{u(s+h)-u(s)}{h} - u'(s)\right) \\
&\xrightarrow{h\to 0} AQ(t-s)u(s) + Q(t-s)u'(s) \\
&= Q(t-s)f(s).
\end{aligned}$$

Since $\lim_{s\to 0} v(s) = Q(t)u_0$, Theorem 4.2.11 implies (4.51). □

So, it would seem that by defining u by (4.51) the problem is solved. However, things are more complicated. First of all, it has been shown that the Bochner integral in (4.51) exists only when (4.50) has a solution. Therefore, we need

Lemma 4.7.2 *The Bochner integral $F(t) \equiv \int_0^t Q(t-s)f(s)ds$ exists for all $t \in [0,\tau)$. Moreover, $F \in C([0,\tau),X)$.*

PROOF Choose $t \in (0,\tau)$ and let $f_1, f_2, \ldots$ be the simple functions approximating $f \in L((0,t),X)$ as in Corollary 4.2.4. Define

$$g_{kt}(s) = Q(t-s)f_k(s), \quad g_t(s) = Q(t-s)f(s) \quad \text{for} \quad s \in (0,t),\ k \ge 1.$$

It is easy to see that $g_{kt} \in L((0,t),X)$ and hence the Dominated Convergence Theorem 4.2.7 implies that $g_t \in L((0,t),X)$ and $\int_0^t g_{kt} \to \int_0^t g_t = F(t)$.

Let $g_t(s) = 0$ for $s > t$. Hence $F(t) = \int_0^\tau g_t(s)ds$. If $t_n \to t \in [0,\tau)$, then $g_{t_n}(s) \to g_t(s)$ for $s \neq t$; hence the DCT implies $F(t_n) \to F(t)$. □

Define the **mild solution** of (4.50) to be u given by (4.51). The mild solution may not be differentiable and it may not lie in the domain of A even when $f \equiv 0$, so it may not be an actual solution. However, if (4.50) has an actual solution, then the actual solution has to be equal to the mild solution by Theorem 4.7.1. The following Theorem 4.7.3 generalizes (3) of Theorem 4.3.1 and shows that the mild solution always satisfies (4.50) *on average*.

Theorem 4.7.3 *The mild solution u of (4.50) is the unique $u \in C([0,\tau),X)$ which satisfies $\int_0^t u(s)ds \in \mathcal{D}(A)$ and*

$$u(t) - u_0 + A\int_0^t u(s)ds = \int_0^t f(s)ds \quad \textit{for all} \quad t \in [0,\tau). \tag{4.52}$$

PROOF Let u be given by (4.51). Note that for $t \in [0,\tau)$,

$$\begin{aligned}\int_0^t u(r)dr &= \int_0^t Q(r)u_0\,dr + \int_0^t dr\int_0^r Q(r-s)f(s)ds \\ &= \int_0^t Q(r)u_0\,dr + \int_0^t ds\int_s^t Q(r-s)f(s)dr \\ &= \int_0^t Q(r)u_0\,dr + \int_0^t ds\int_0^{t-s} Q(r)f(s)dr.\end{aligned}$$

Hence (3) of Theorem 4.3.1 and Theorem 4.2.10 imply that $\int_0^t u \in \mathcal{D}(A)$ and

$$\begin{aligned}A\int_0^t u(r)dr &= u_0 - Q(t)u_0 + \int_0^t (f(s) - Q(t-s)f(s))ds \\ &= u_0 - u(t) + \int_0^t f(s)ds.\end{aligned}$$

Therefore the mild solution of (4.50) satisfies (4.52).

Let v be a continuous solution of (4.52) and let $w(t) = \int_0^t (u-v)$. Then $w' + Aw = 0$, $w(0) = 0$ and hence (8) of Theorem 4.3.1 implies that $w \equiv 0$. □

Corollary 4.7.4 *If $u_0 \in \mathcal{D}(A)$ and $f(t) = c + \int_0^t g$ for some $c \in X$, $g \in L((0,t),X)$ for all $t \in (0,\tau)$, then there exists u in $C^1([0,\tau),X)$ such that (4.50) holds.*

PROOF Let $v \in C([0,\tau),X)$ be such that

$$v(t) + Au_0 - c + A\int_0^t v(s)ds = \int_0^t g(s)ds \quad \text{for all} \quad t \in [0,\tau)$$

and let $u(t) = u_0 + \int_0^t v$. □

One cannot weaken the condition on f in the above Corollary to $f \in C([0,\tau),X)$, see Exercise 21.

Corollary 4.7.5 *If $u_0 \in \mathcal{D}(A)$ and $f, Af \in C([0,\tau),X)$, then there exists u in $C^1([0,\tau),X)$ such that (4.50) holds.*

PROOF Note that the mild solution u, given by (4.51), lies in $\mathcal{D}(A)$ by Theorem 4.2.10. Lemma 4.7.2 implies that $Au \in C([0,\tau),X)$; hence $A\int_0^t u = \int_0^t Au$ and we can differentiate (4.52). □

4.8 Exercises

1. Show that $Q : [0,\infty) \to \mathfrak{B}(C_u(\mathbb{R}))$, given in Example 4.1.3, is not continuous.

2. Let $\mathcal{H}$ be a Hilbert space, let $\{\varphi_1, \varphi_2, \ldots\}$ be a complete orthonormal set in $\mathcal{H}$ and let $\lambda_1, \lambda_2, \ldots$ be complex numbers. Define A by

$$Af = \sum_{k=1}^{\infty} \lambda_k (f, \varphi_k)\varphi_k \quad \text{for} \quad f \in \mathcal{D}(A) = \{f \in \mathcal{H} \mid \sum_{k=1}^{\infty} |\lambda_k (f, \varphi_k)|^2 < \infty\}.$$

   Show that $\sigma_p(A) = \{\lambda_1, \lambda_2, \ldots\}$ and that $\sigma(A) = \overline{\sigma_p(A)}$.

3. Suppose $f \in C^2[0, 2\pi]$, $f(0) = f(2\pi)$, $f_x(0) = f_x(2\pi)$ and let u satisfy

$$u_t = u_{xx} + (uf_x)_x \quad \text{for} \quad 0 \le x \le 2\pi, \ t \ge 0,$$

$$u(0,t) = u(2\pi, t), \ u_x(0,t) = u_x(2\pi, t) \quad \text{for} \quad t \ge 0,$$

$$u(\cdot, 0) \in L^2(0, 2\pi).$$

   Show that there exists c such that

$$\lim_{t\to\infty} u(x,t) = c\, e^{-f(x)} \quad \text{for} \quad 0 \le x \le 2\pi.$$

4. If A is as in Section 4.4, show that there exists $\delta \in (0,\infty)$ such that

$$(\operatorname{Im}\lambda)^2 \le 4\delta(\delta + \operatorname{Re}\lambda) \quad \text{for all} \quad \lambda \in \sigma(A).$$

5. Let $-A$ be the generator of a strongly continuous semigroup $\{Q(t)\}_{t\ge 0}$ on a Banach space X such that the spectrum of A contains a scalar λ with $\operatorname{Re}\lambda < 0$. Show that there exists $x \in \mathcal{D}(A)$ such that $Q(t)x$ is unbounded for $t \ge 0$.

6. Show that the generator of the semigroup in Example 4.1.5 is the operator $\overline{\Delta}$ defined in Example 4.3.10. (*Hint:* Compare the resolvents.)

7. Prove the 'Observe ...' statements in Example 4.5.4.

8. Consider

$$Au = -\Delta u + \sum_{i=1}^{n} a_i D_i u + a_0 u \quad \text{for} \quad u \in \mathcal{D}(A) \equiv \mathcal{S}$$

   as an operator in $W_0^{0,p}(\mathbb{R}^n)$, $1 \le p \le \infty$. Assume that $a_i \in L^\infty(\mathbb{R}^n)$ and when $p = \infty$, assume also that a_i are continuous. Show that A is closable and that $\overline{A}$ is equal to the operator T given in Example 4.5.8.

9. Let B be the linear operator in $L^2(\mathbb{R}^n)$ associated with the sectorial form

$$\mathfrak{F}(u,v) = \int_{\mathbb{R}^n} \sum_{i=1}^{n} (\mathrm{D}_i u \overline{\mathrm{D}_i v} + a_i(\mathrm{D}_i u)\overline{v}) + a_0 u\overline{v} \quad \text{for} \quad u, v \in W^1(\mathbb{R}^n),$$

where $a_i \in L^\infty(\mathbb{R}^n)$. Show that B is equal to the operator $\overline{A}$ given in Exercise 8 when $p = 2$.

10. Suppose that $u : (0,T) \to L^p(\Omega)$ has m continuous derivatives where $0 < T \leq \infty$, $1 \leq p \leq \infty$ and Ω is a nonempty open subset of $\mathbb{R}^n$. It would be nice to have that if

$$u^{(k)}(t) \in W^{m-k,p}_{loc}(\Omega) \quad \text{for} \quad 0 < t < T,\ 0 \leq k \leq m,$$

then $u \in W^{m,p}_{loc}(\Omega \times (0,T))$. Show that this is not true.

11. Suppose that $A \in \mathfrak{A}(a, M, \theta, X)$ and $\zeta \in \sigma(A)\backslash\{a\}$. Show that $M \sin\theta \geq 1$.

12. Show that every bounded operator on a complex Banach space is a sectorial operator.

13. Let M, N be closed subspaces of a Banach space X such that $X = M + N$ and $M \cap N = \{0\}$. Show that there exists a projection P on X such that $M = \mathcal{R}(P)$ and $N = \mathcal{R}(1 - P)$.

14. Show that a nonempty bounded subset σ of $\sigma(A)$ is a spectral set of a linear operator A in a complex Banach space iff there exists $\delta > 0$ such that

$$|\zeta - \lambda| \geq \delta \quad \text{for all} \quad \zeta \in \sigma,\ \lambda \in \sigma(A)\backslash\sigma.$$

15. In the proof of Theorem 4.6.6 it is assumed that A is closed. Why is this true?

16. Let $v \in L^2(0,\pi)$ be represented as $v(x) = \sum_{n=1}^{\infty} b_n \sin nx$. Define

$$u(x,t) = \sum_{n=1}^{\infty} b_n e^{-n^2 t} \sin nx \quad \text{for} \quad t > 0,\ x \in (0,\pi).$$

Show that $\|u(\cdot,t)\|_p \leq \|v\|_p$ for $p \in [1,\infty]$, $t > 0$. (*Hint:* estimate the resolvent in Theorem 4.3.5 by using Exercise 20 in Chapter 1.)

17. Let $-A$ be the generator of a contraction semigroup. Show that

$$\|Ax\|^2 \leq 4\|A^2 x\|\|x\| \quad \text{for} \quad x \in \mathcal{D}(A^2).$$

Hint: if $\{Q(t)\}_{t\geq 0}$ is the semigroup and $x \in \mathcal{D}(A^2)$, then

$$tAx = x - Q(t)x + \int_0^t (t-s)Q(s)A^2 x\, ds \quad \text{for} \quad t > 0.$$

18. Use the above problem to improve Lemma 3.4.11 to

$$\|f'\|_p \leq \varepsilon \|f''\|_p + \|f\|_p/\varepsilon \quad \text{for} \quad f \in C^2(\mathbb{R}),\ \varepsilon \in (0,\infty),\ p \in [1,\infty).$$

19. Show that if A is a sectorial operator with compact resolvent, then e^{-At} is compact for $t > 0$.

20. Find a semigroup $\{Q(t)\}_{t\geq 0}$ such that $Q(t)$ is not compact for any $t \geq 0$, yet its generator has a compact resolvent. (*Hint:* analyze examples given in the Introduction.)

21. Show that the condition on f in Corollary 4.7.4 cannot be weakened to requiring that $f \in C([0,\tau), X)$. (*Hint:* take $f(t) = Q(t)x$.)

Chapter 5

Weakly Nonlinear Evolution Equations

5.1 Introduction

The term **weakly nonlinear evolution equations** will be applied to abstract evolution equations of the form

$$u'(t) + Au(t) = F(t, u(t)) \quad \text{for} \quad t \in [0, \tau), \tag{5.1}$$

where A is a generator of a strongly continuous semigroup in a Banach space X and F is continuous and locally Lipschitz continuous in the last variable. There are no restrictions on A; hence both hyperbolic, including the wave equation, and parabolic problems can be treated. However, the restriction on F can be quite severe. Hence the term weakly nonlinear evolution equations is used. The restriction on F is not so bad when one can choose to work in various spaces of continuous functions. For example, equations like

$$u_t = \Delta u + e^u, \quad u_t = \Delta u + \frac{1}{1-u}$$

are studied in Section 5.3 in a closed subspace of $C_B(\mathbb{R}^n)$.

Basic existence, uniqueness, continuous dependence and stability results for weakly nonlinear evolution equations are presented in this chapter. The basic ideas used to prove these results are the same as in the ODE theory and are also the same as in the next chapter where semilinear parabolic problems are studied. However, technical preliminaries are much simpler here than in the next chapter.

Most of the chapter is devoted to studying approximations of solutions of such equations. By further extending these results, one can actually remove the condition that F is locally Lipschitz continuous in X (Example 5.6.2).

The following version of Gronwall's inequality will be sufficient here.

Theorem 5.1.1 (Gronwall) *If $\ell \in (0,\infty)$ and $f \in L^1(0,\ell)$ is real valued such that*

$$f(x) \le a + b\int_0^x f(s)ds \quad \text{for} \quad x \in (0,\ell) \text{ a.e.},$$

where $a \in \mathbb{R}$ and $b \in [0,\infty)$, then $f(x) \le ae^{bx}$ for almost all x in $(0,\ell)$.

PROOF If $F(x) = be^{-bx}\int_0^x f(s)ds$, then $F'(x) \le abe^{-bx}$ which implies that $a + e^{bx}F(x) \le ae^{bx}$. □

The key to all basic results presented in the next section is the following Theorem. Note that for (5.2) to be satisfied we must have that $\int_0^t u \in \mathcal{D}(A)$.

Theorem 5.1.2 *Suppose $-A$ is the generator of a strongly continuous semigroup $\{Q(t)\}_{t\ge 0}$ on a Banach space X, $a \in \mathbb{R}$, $M \in \mathbb{R}$, $\|Q(t)\| \le Me^{-at}$ for $t \ge 0$, $0 < \tau < \infty$, $H : [0,\tau] \times X \to X$ is continuous and for some $L < \infty$,*

$$\|H(t,x) - H(t,y)\| \le L\|x - y\| \quad \text{for} \quad 0 \le t \le \tau,\ x,y \in X.$$

Then for each $u_0 \in X$ there exists a unique $u \in C([0,\tau],X)$ which satisfies

$$u(t) - u_0 + A\int_0^t u(s)ds = \int_0^t H(s,u(s))ds \quad \text{for all} \quad t \in [0,\tau]. \tag{5.2}$$

Moreover, if $G : X \to C([0,\tau],X)$ is given by $G(u_0) = u$, then

$$\|G(x)(t) - G(y)(t)\| \le M\|x - y\|e^{(ML-a)t} \quad \text{for} \quad 0 \le t \le \tau,\ x,y \in X. \tag{5.3}$$

PROOF Abbreviate $Y = C([0,\tau],X)$ and note that Y is a Banach space with norm $\|\cdot\|_\infty$ (Example 1.3.2). Define $T : Y \to Y$ by

$$(Tu)(t) = Q(t)u_0 + \int_0^t Q(t-s)H(s,u(s))\,ds \quad \text{for} \quad u \in Y,\ 0 \le t \le \tau.$$

If $u, v \in Y$, then

$$\|(Tu)(t) - (Tv)(t)\| \le cL\int_0^t \|u(s) - v(s)\|ds \le cLt\|u - v\|_\infty$$

where $c = \sup_{0\le t\le\tau}\|Q(t)\|$ and hence

$$\|(T^n u)(t) - (T^n v)(t)\| \le \frac{(cLt)^n}{n!}\|u - v\|_\infty \tag{5.4}$$

for $n \ge 1$, $0 \le t \le \tau$. Thus, the Contraction Mapping Theorem 1.1.3 implies the existence of a unique $u \in Y$ such that $Tu = u$. Theorem 4.7.3 and closedness of A imply (5.2).

Choose $x, y \in X$ and let $u = G(x), v = G(y)$. For $0 \leq t \leq \tau$ we have

$$\begin{aligned} \|u(t) - v(t)\| &\leq Me^{-at}\|x - y\| + ML\int_0^t e^{-a(t-s)}\|u(s) - v(s)\|ds \\ e^{at}\|u(t) - v(t)\| &\leq M\|x - y\| + ML\int_0^t e^{as}\|u(s) - v(s)\|ds; \end{aligned}$$

hence, the Gronwall Theorem 5.1.1 implies (5.3). □

The following Corollary says that bounded perturbations of generators of strongly continuous semigroups are still generators of strongly continuous semigroups. This is an important result, but should not be surprising. However, it is interesting that it is a natural consequence of Theorem 5.1.2 which deals with nonlinear problems.

Corollary 5.1.3 *Suppose that $-A$ is the generator of a strongly continuous semigroup on a Banach space X and that $B \in \mathfrak{B}(X)$. Then $-A + B$ is the generator of a strongly continuous semigroup on X.*

PROOF Theorem 5.1.2 implies that for each $x \in X$ there exists a unique $u \in C([0, \infty), X)$ such that

$$u(t) = x + (-A + B)\int_0^t u(s)ds \quad \text{for all} \quad t \in [0, \infty). \tag{5.5}$$

Define $R(t)x = u(t)$ for $t \geq 0$. Uniqueness of u and formula (5.5) imply linearity of $R(t)$ for $t \geq 0$. (5.3) implies that $R(t) \in \mathfrak{B}(X)$ for $t \geq 0$. If $t, r \geq 0$, then obviously

$$u(t + r) = u(r) + (-A + B)\int_r^{t+r} u(s)ds;$$

hence $R(t)R(r)x = R(t)u(r) = u(t+r) = R(t+r)x$. Thus $\{R(t)\}_{t\geq 0}$ is a strongly continuous semigroup on X. Let $-S$ be its generator. Note that Theorem 4.3.2 implies that $(-\infty, a) \in \rho(A - B) \cap \rho(S)$ for some $a \in \mathbb{R}$. If $x \in \mathcal{D}(S)$ and u is as in (5.5), then

$$\frac{1}{h}\int_0^h u(s)ds \to x \quad \text{and} \quad (A - B)\frac{1}{h}\int_0^h u(s)ds = \frac{1}{h}(x - R(h)x) \to Sx$$

as $h \to 0+$ and hence the closedness of $A - B$ implies that $A - B$ is an extension of S. Lemma 1.6.14 implies that $A - B = S$. □

R_r given below is called a **retraction map**. It will enable us to apply Theorem 5.1.2 to a locally Lipschitz continuous F.

Theorem 5.1.4 *If X a Banach space, $r > 0$ and $R_r : X \to X$ is given by*

$$R_r(x) = x \text{ when } \|x\| \le r \quad \text{and} \quad R_r(x) = \frac{r}{\|x\|}x \text{ when } \|x\| > r,$$

then $\|R_r(x) - R_r(y)\| \le 2\|x - y\|$ for all $x, y \in X$.

PROOF Let $B = \{x \in X \mid \|x\| \le r\}$. If $x, y \in B$, there is nothing to prove.

If $x \in B$ and $y \in X \backslash B$, then

$$\begin{aligned} Rx - Ry &= x - y + (1 - \frac{r}{\|y\|})y \\ \|Rx - Ry\| &\le \|x - y\| + \|y\| - r \\ &\le 2\|x - y\| + \|x\| - r \le 2\|x - y\|. \end{aligned}$$

If $x, y \in X \backslash B$, then

$$\begin{aligned} Rx - Ry &= \frac{r}{\|x\|}(x - y) + \frac{r(\|y\| - \|x\|)}{\|x\|\|y\|}y \\ \|Rx - Ry\| &\le \|x - y\| + |\|y\| - \|x\|| \le 2\|x - y\|. \end{aligned}$$

□

5.2 Basic Theory

Throughout this section it is assumed that

(1) X is a Banach space.

(2) $-A$ is the generator of a strongly continuous semigroup $\{Q(t)\}_{t \ge 0}$ on X. Let $M \in [1, \infty)$, $a \in \mathbb{R}$ be such that $\|Q(t)\| \le Me^{-at}$ for $t \ge 0$.

(3) $T \in (0, \infty]$, $\mathcal{U}$ is an open set in X, $F : [0, T) \times \mathcal{U} \to X$ is continuous and for each $t \in (0, T)$ and each $z \in \mathcal{U}$ there exist $\delta > 0$ and $L < \infty$ such that

$$\|F(s, x) - F(s, y)\| \le L\|x - y\| \quad \text{for} \quad x, y \in B(z, \delta),\ s \in [0, t].$$

For $\tau \in (0, T]$, let $\mathcal{S}_m(\tau)$ be the collection of all $u \in C([0, \tau), \mathcal{U})$ which satisfy $\int_0^t u(s)ds \in \mathcal{D}(A)$ and

$$u(t) - u(0) + A\int_0^t u(s)ds = \int_0^t F(s, u(s))ds \quad \text{for all} \quad t \in [0, \tau). \tag{5.6}$$

Recall that Theorem 4.7.3 implies that $u \in \mathcal{S}_m(\tau)$ iff $u \in C([0, \tau), \mathcal{U})$ and

$$u(t) = Q(t)u(0) + \int_0^t Q(t - s)F(s, u(s))ds \quad \text{for} \quad t \in [0, \tau). \tag{5.7}$$

So $\mathcal{S}_m(\tau)$ is the set of mild solutions of (5.1).

Lemma 5.2.1 *Suppose $\tau \in (0,T)$ and $w \in C([0,\tau],\mathcal{U})$. Then there exist $\delta > 0$, $L < \infty$ and a continuous $H : [0,\tau] \times X \to X$ such that*

(i) if $t \in [0,\tau]$ and $z \in B(w(t),\delta)$, then $z \in \mathcal{U}$ and $H(t,z) = F(t,z)$

(ii) $\|H(t,x) - H(t,y)\| \le L\|x-y\|$ for $t \in [0,\tau]$, $x,y \in X$.

PROOF For $t \in [0,\tau]$ choose $\delta(t) > 0$ and $L(t) < \infty$ such that

$$\|F(s,x) - F(s,y)\| \le L(t)\|x-y\| \quad \text{for} \quad x,y \in B(w(t),\delta(t)),\ s \in [0,\tau]$$

and let $\mu(t) > 0$ be such that $\|w(s) - w(t)\| < \delta(t)/2$ if $|s-t| < \mu(t)$. Compactness of $[0,\tau]$ implies that $[0,\tau] \subset B(t_1,\mu(t_1)) \cup \cdots \cup B(t_n,\mu(t_n))$ for some $t_i \in [0,\tau]$. Let $L = 2\max_i L(t_i)$ and $\delta = \min_i \delta(t_i)/4$. Note that if $t \in [0,\tau]$ and $x,y \in B(w(t),2\delta)$, then $x,y \in \mathcal{U}$ and $\|F(t,x) - F(t,y)\| \le L\|x-y\|/2$.

Define $H(t,x) = F(t,w(t) + R_\delta(x - w(t)))$ for $t \in [0,\tau]$ and $x \in X$, where R_δ is the retraction map as given in Theorem 5.1.4. □

Existence for the initial value problem:

Theorem 5.2.2 *For each $u_0 \in \mathcal{U}$ there exist $\theta \in (0,T)$ and $u \in \mathcal{S}_m(\theta)$ such that $u(0) = u_0$. Furthermore, $\lim_{n\to\infty} \sup_{0\le t\le\theta} \|u_n(t) - u(t)\| = 0$ where $u_k \in C([0,\theta],\mathcal{U})$ are defined by $u_1(t) = u_0$ and*

$$u_{k+1}(t) = Q(t)u_0 + \int_0^t Q(t-s)F(s,u_k(s))ds \quad \text{for} \quad t \in [0,\theta],\ k = 1,2,\ldots$$

PROOF Pick $\tau \in (0,T)$ and let $w(t) = u_0$ for $t \in [0,\tau]$. Let δ and H be as given in Lemma 5.2.1 and let $u \in C([0,\tau],X)$ be as given by Theorem 5.1.2. Continuity of u implies that for some $\theta \in (0,\tau]$ we have for $t \in [0,\theta]$ that $\|u(t) - w(t)\| < \delta$. Hence $H(t,u(t)) = F(t,u(t))$ and therefore $u \in \mathcal{S}_m(\theta)$. If θ is small enough, then (5.4) implies the 'Furthermore' part. □

Uniqueness for the initial value problem:

Theorem 5.2.3 *If $0 < \mu \le \theta \le T$, $w \in \mathcal{S}_m(\mu)$, $u \in \mathcal{S}_m(\theta)$ and $w(0) = u(0)$, then $w(t) = u(t)$ for $t \in [0,\mu)$.*

PROOF Choose any $\tau \in (0,\mu)$ and let δ and H be as given in Lemma 5.2.1. For some $\tau' \in (0,\tau]$ we have $\|u(t) - w(t)\| < \delta$ for $t \in [0,\tau']$ and hence $H(t,u(t)) = F(t,u(t))$. Theorem 5.1.2 implies that $u = w$ on $[0,\tau']$, and hence we can choose $\tau' = \tau$. □

Continuous dependence on the initial value:

Theorem 5.2.4 *Suppose* $0 < \tau < \mu \leq T$ *and* $w \in \mathcal{S}_m(\mu)$. *Then there exist* $\varepsilon > 0$ *and* $C < \infty$ *such that for every* $x \in B(w(0), \varepsilon)$ *there exists* $u \in \mathcal{S}_m(\tau)$ *such that* $u(0) = x$ *and*

$$\|u(t) - w(t)\| \leq C\|u(0) - w(0)\| \quad \textit{for all} \quad t \in [0, \tau). \tag{5.8}$$

PROOF Let δ, L and H be as in Lemma 5.2.1. Let $C = \max\{Me^{(ML-a)\tau}, M\}$ and choose $\varepsilon \in (0, \delta/C)$. Suppose $x \in B(w(0), \varepsilon)$. Let $u = G(x)$, where G is as in Theorem 5.1.2 and note that (5.8) holds. Since $C\varepsilon < \delta$, Lemma 5.2.1 implies that $H(t, u(t)) = F(t, u(t))$ for $t \in [0, \tau]$ and hence $u \in \mathcal{S}_m(\tau)$. □

Solutions of real equations, with real initial conditions, are real:

Theorem 5.2.5 *Suppose that* Y *is a closed real subspace of* X *such that*

(1) $(A - \lambda)^{-1}Y \subset Y$ *for some* $\lambda < a$

(2) $F(t, x) \in Y$ *when* $t \in [0, T)$ *and* $x \in Y \cap \mathcal{U}$.

Suppose also that $u \in \mathcal{S}_m(\theta)$ *for some* $\theta \in (0, T]$ *and that* $u(0) \in Y$. *Then*

$$u(t) \in Y \quad \textit{for all} \quad t \in [0, \theta).$$

PROOF Choose any $\tau \in [0, \theta)$ such that $u(t) \in Y$ for $t \in [0, \tau]$. Theorem 5.2.2 implies that there exist $\delta \in (0, \theta - \tau)$ and $v \in C([0, \delta), \mathcal{U})$ such that

$$v(t) - u(\tau) + A\int_0^t v(s)ds = \int_0^t F(\tau + s, v(s))ds \quad \text{for all} \quad t \in [0, \delta)$$

and that $v = \lim_{n\to\infty} v_n$ for some $v_n \in C([0, \delta), \mathcal{U})$ given by $v_1(t) = u(\tau)$ and

$$v_{n+1}(t) = Q(t)u(\tau) + \int_0^t Q(t - s)F(\tau + s, v_n(s))ds$$

for $t \in [0, \delta)$ and $n = 1, 2, \ldots$ Since $u(\tau) \in Y$, Theorem 4.6.1 and Corollary 4.2.5 imply that the range of every v_n is in Y. Hence, the range of v is in Y. Let $w = u$ on $[0, \tau]$ and $w(t) = v(t - \tau)$ for $t \in [\tau, \tau + \delta)$. It is clear that $w \in \mathcal{S}_m(\tau + \delta)$. The uniqueness Theorem 5.2.3 implies that $w = u$ on $[0, \tau + \delta)$. Hence, $u(t) \in Y$ for $t \in [0, \tau + \delta)$. Therefore, the supremum of such τ has to be equal to θ. □

Maximal interval of existence:

Theorem 5.2.6 *Suppose $u_0 \in \mathcal{U}$. Then there exist $\tau \in (0,T]$, $u \in \mathcal{S}_m(\tau)$ such that $u(0) = u_0$, and which also have the property that $\tau \geq \mu$ for all $\mu \in (0,T]$ for which there exists $v \in \mathcal{S}_m(\mu)$ with $v(0) = u_0$. Moreover, if $\tau < T$, then either $\int_0^\tau \|F(t,u(t))\|dt = \infty$ or there exists $x \in \overline{\mathcal{U}}\backslash\mathcal{U}$ such that $\lim_{t\to\tau} u(t) = x$.*

PROOF Theorems 5.2.2 and 5.2.3 imply the first assertion. Suppose $\tau < T$ and $\int_0^\tau \|F(t,u(t))\|dt < \infty$. The Dominated Convergence Theorem 4.2.7 and (5.7) imply that there exists $x \in X$ such that $\lim_{t\to\tau} u(t) = x$. Clearly $x \in \overline{\mathcal{U}}$. The proof will be complete if we show that the assumption $x \in \mathcal{U}$ leads to a contradiction. So, assume $x \in \mathcal{U}$, choose $\tau' \in (\tau,T)$ and let $u(t) = x$ for $t \in [\tau,\tau']$. Let δ and H be as given in Lemma 5.2.1 applied to $u \in C([0,\tau'],\mathcal{U})$. Theorem 5.1.2 implies the existence of $v \in C([0,\tau'],X)$ such that

$$v(t) - u_0 + A\int_0^t v(s)ds = \int_0^t H(s,v(s))ds \quad \text{for} \quad t \in [0,\tau']$$

and its uniqueness implies that $v = u$ on $[0,\tau)$. For some $\mu \in (\tau,\tau']$ we have that $v(t) \in B(u(t),\delta)$ for $t \in [0,\mu]$ and hence $H(t,v(t)) = F(t,v(t))$ for $t \in [0,\mu]$. Therefore $v \in \mathcal{S}_m(\mu)$, which is a contradiction to the maximality of τ. □

Stability:

Theorem 5.2.7 *Suppose $a > 0$, $\varepsilon \in [0,a/M]$, $T = \infty$, $\delta > 0$, $B(0,\delta) \subset \mathcal{U}$ and*

$$\|F(t,x)\| \leq \varepsilon\|x\| \quad \textit{for} \quad t \geq 0,\ x \in B(0,\delta).$$

Then for each $x \in B(0,\delta/M)$ there exists $u \in \mathcal{S}_m(\infty)$ such that $u(0) = x$ and

$$\|u(t)\| \leq M\|x\|e^{(M\varepsilon-a)t} \quad \textit{for} \quad t \geq 0. \tag{5.9}$$

PROOF Pick the maximal τ for which there exists $u \in \mathcal{S}_m(\tau)$ satisfying $u(0) = x$ (Theorem 5.2.6). Note that for some $\mu \in (0,\tau]$ we have that $\|u(t)\| < \delta$ for $0 \leq t < \mu$. Let τ' be the largest of such μ. (5.7) implies

$$\begin{aligned} \|u(t)\| &\leq Me^{-at}\|x\| + M\varepsilon\int_0^t e^{-a(t-s)}\|u(s)\|ds \\ e^{at}\|u(t)\| &\leq M\|x\| + M\varepsilon\int_0^t e^{as}\|u(s)\|ds \quad \text{for} \quad 0 \leq t < \tau' \end{aligned}$$

and therefore the Gronwall Theorem 5.1.1 implies the bound (5.9) for $t \in [0,\tau')$. If $\tau' < \tau$, then (5.9) implies that $\|u(\tau')\| \leq M\|x\| < \delta$, which contradicts the maximality of τ'. Hence $\tau' = \tau$. If $\tau < \infty$, then obviously $\int_0^\tau \|F(t,u(t))\|dt < \infty$. Hence Theorem 5.2.6 implies that $\lim_{t\to\tau} u(t) = y \in \overline{\mathcal{U}}\backslash\mathcal{U}$, which is not possible because $\|y\| \leq M\|x\| < \delta$. Therefore $\tau = \infty$. □

5.3 Example: Nonlinear Heat Equation

We shall discuss the initial value problem for

$$u_t(x,t) = \Delta u(x,t) + f(u(x,t)) \quad \text{for} \quad x \in \mathbb{R}^n,\ t \geq 0$$

$$u(x,0) = u_0(x) \quad \text{for} \quad x \in \mathbb{R}^n,$$

where $f \in C^1((a,b),\mathbb{R})$ for some $-\infty \leq a < b \leq \infty$.

The problem will be set in the real Banach space C_ℓ. Let Δ_ℓ be the generator of the contraction semigroup on C_ℓ as given in Example 4.5.4 - but restricted to real functions (Example 4.6.3). Observe that Δ_ℓ is equal to the closure in C_ℓ of the operator Δ, with the domain consisting of functions in the form $c + v$ where $c \in \mathbb{R}$ and v is a real valued, rapidly decreasing function.

Let $\mathcal{U}$ consist of those $v \in C_\ell$ such that

$$v_- \equiv \inf_x v(x) > a \quad \text{and} \quad v_+ \equiv \sup_x v(x) < b.$$

If $v \in \mathcal{U}$, then $v_- - \varepsilon > a$ and $v_+ + \varepsilon < b$ for some $\varepsilon > 0$; hence

$$\|f(z) - f(w)\|_\infty \leq L\|z - w\|_\infty \quad \text{for} \quad z, w \in B(v,\varepsilon)$$

where $L = \max_{v_- - \varepsilon \leq x \leq v_+ + \varepsilon} |f'(x)|$.

Assume that $u_0 \in \mathcal{U}$.

Let $\tau \in (0,\infty]$ and $u \in \mathcal{S}_m(\tau)$ be as given by Theorem 5.2.6.

Theorem 5.2.3 implies that for every $\mu \in (0,\tau]$ this u is the unique element of $C([0,\mu), C_\ell)$ which satisfies $\int_0^t u(s)ds \in \mathcal{D}(\Delta_\ell)$ and

$$u(t) - u_0 = \Delta_\ell \int_0^t u(s)ds + \int_0^t f(u(s))ds \quad \text{for all} \quad t \in [0,\mu). \tag{5.10}$$

So u is the solution of the original problem on average. By showing its regularity, one can show, see Example 6.4.1 in the next chapter, that it is also the unique actual solution.

(5.10) obviously implies that if $r \in [0,\tau)$ and $\tilde{u}(t) = u(t+r)$ for $t \in [0,\tau - r)$, then $\tilde{u} \in \mathcal{S}_m(\tau - r)$. Theorem 5.2.3 implies that this $\tilde{u}$ is the unique solution starting at $u(r)$. Thus we have the semigroup property.

Theorem 5.2.6 also implies that if the solution does not exist for all time, i.e. $\tau < \infty$, then either $\int_0^\tau \sup_x |f(u(x,t))|dt = \infty$ or approaches to $\overline{\mathcal{U}}\backslash\mathcal{U}$, the boundary of $\mathcal{U}$. When, for example, $f(u) = \pm e^u$, then the solution either exists for all time or else $\sup_{0<t<\tau} \|u(t)\|_\infty = \infty$. When, for example, f is bounded and $(a,b) = \mathbb{R}$, then the solution exists for all time.

Let $h(t) \in \mathbb{R}$ be the value of $u(t)$ at infinity, i.e. $h(t) = \ell(u(t))$ where $\ell \in C_\ell{}^\star$ is as defined in Example 4.5.4. The fact that $\ell(\Delta_\ell v) = 0$ for all $v \in \mathcal{D}(\Delta_\ell)$, when applied to (5.10), implies that h satisfies an ODE,

$$h'(t) = f(h(t)) \quad \text{for} \quad t \in [0,\tau).$$

When u_0 is a constant function, then the uniqueness of u implies that $u(x,t) = h(t)$ for all $x \in \mathbb{R}$ and $t \in [0,\tau)$. When, for example, $f(u) = 1/(1-u)$, $b = -\infty$ and $a = 1$, then the solution of the ODE is

$$h(t) = 1 - \sqrt{(1-h(0))^2 - 2t},$$

which implies that $\tau \leq (1-h(0))^2/2$. The following Theorem 5.3.1 implies that $\tau \geq (1-\sup_x u_0(x))^2/2$ for this f.

Theorem 5.3.1 *If $v, w \in \mathcal{S}_m(\mu)$, $w(0) \geq v(0)$ and $f' \geq 0$, then $w(t) \geq v(t)$ for $0 \leq t < \mu$.*

PROOF Suppose that $r \in [0,\mu)$ and that $w \geq v$ on $[0,r]$. The proof will be complete if we show that there exists $\theta > 0$ such that $w \geq v$ on $[0, r+\theta]$.

By Theorem 5.2.2 we can choose $\tilde{v}_k \in C([0,\theta_v],\mathcal{U})$ such that $\tilde{v}_1(t) = v(r)$,

$$\tilde{v}_{k+1}(t) = e^{\Delta_\ell t} v(r) + \int_0^t e^{\Delta_\ell (t-s)} f(\tilde{v}_k(s))ds \quad \text{for} \quad t \in [0,\theta_v],\ k = 1,2,\ldots$$

and $\tilde{v}_k$ converge to some $\tilde{v} \in \mathcal{S}_m(\theta_v)$ with $\tilde{v}(0) = v(r)$. Here $\theta_v \in (0, \mu - r)$. The semigroup property implies that $\tilde{v}(t) = v(r+t)$ for $t \in [0,\theta_v)$.

Choose $\tilde{w}_k \in C([0,\theta_w],\mathcal{U})$ in the same way and let $\theta = \min\{\theta_v, \theta_w\} > 0$.

In view of the integral representation of $e^{\Delta_\ell t}$ (Exercise 4), we conclude that $e^{\Delta_\ell t}$ preserves positivity. Hence

$$e^{\Delta_\ell t}(w(r) - v(r)) \geq 0 \quad \text{for} \quad t \geq 0.$$

If $\tilde{w}_k \geq \tilde{v}_k$ on $[0,\theta]$, then using

$$\tilde{w}_{k+1}(t) - \tilde{v}_{k+1}(t) = e^{\Delta_\ell t}(w(r) - v(r)) + \int_0^t e^{\Delta_\ell (t-s)}(f(\tilde{w}_k(s)) - f(\tilde{v}_k(s)))ds$$

and the fact that pointwise evaluations are bounded linear functionals on C_ℓ gives that $\tilde{w}_{k+1} \geq \tilde{v}_{k+1}$ on $[0,\theta]$. Since $\tilde{w}_k, \tilde{v}_k$ converge to $\tilde{w}, \tilde{v}$ in C_ℓ, we have that $w \geq v$ on $[0, r+\theta]$. □

We have the same stability result as for ODEs:

Theorem 5.3.2 *If $f(c) = 0$ and $f'(c) < 0$ for some $c \in (a,b)$, then for each $\mu \in (0, -f'(c))$ there exists $\delta > 0$ such that if $\|u_0 - c\|_\infty < \delta$, then $u \in \mathcal{S}_m(\infty)$ and*

$$\|u(t) - c\|_\infty \leq \|u_0 - c\|_\infty e^{-\mu t} \quad \text{for} \quad t \geq 0.$$

PROOF Let $A = -\Delta_\ell - f'(c)$ and $F(v) = f(c+v) - f'(c)v$. Note that $\|e^{-At}\| \le e^{f'(c)t}$ for $t \ge 0$ and that for $\varepsilon = -f'(c) - \mu$ there exists $\delta > 0$ such that $\|F(v)\|_\infty \le \varepsilon \|v\|_\infty$ if $\|v\|_\infty < \delta$. Theorem 5.2.7 applies, giving $v \in \mathcal{S}_m(\infty)_{A,F}$ such that $v(0) = u_0 - c$ and $\|v(t)\|_\infty \le \|v(0)\|_\infty e^{-\mu t}$. Using (5.6), it is easy to see that $v + c \in \mathcal{S}_m(\infty)$ and, by uniqueness, $v + c = u$. □

As for ODEs, we also have instability when $f'(c) > 0$. The tools needed to prove this are developed in the next chapter. See Example 6.4.13.

5.4 Approximation for Evolution Equations

We shall approximate the solution u of

$$u'(t) + Au(t) = H(t, u(t)), \quad u(0) = x, \tag{5.11}$$

without using any a priori information about u. The original problem is set in a Banach space X. We shall seek approximations in spaces X_n, which are usually finite dimensional spaces in applications and hence this step is sometimes referred to as space discretization. The approximation u_n will satisfy

$$u_n'(t) + A_n u_n(t) = P_n H(t, E_n u_n(t)), \quad u_n(0) = P_n x, \tag{5.12}$$

where A_n is a bounded operator on X_n approximating A, P_n denotes a 'projection' from X to X_n and E_n is an 'interpolation' that associates with each element of X_n an element of X. Using some minimal assumptions, it will be shown that $E_n u_n \to u$. (5.12) is typically an ODE which has to be integrated numerically. In actual numerical applications it is usually preferred to discretize time t at the same time as space X. One popular way to do this is to use the explicit Euler method,

$$\frac{u_n(t+\tau_n) - u_n(t)}{\tau_n} + A_n u_n(t) \approx P_n H(t, E_n \tilde{u}_n(t)).$$

So, if v_n^k denotes the approximation of $u_n(k\tau_n)$, then

$$v_n^{k+1} = (1 - \tau_n A_n) v_n^k + \tau_n P_n H(k\tau_n, v_n^k), \quad v_n^0 = P_n x. \tag{5.13}$$

A variant of the implicit Euler method is

$$w_n^{k+1} = (1 + \varepsilon_n A_n)^{-1} w_n^k + \varepsilon_n P_n H(k\varepsilon_n, w_n^k), \quad w_n^0 = P_n x. \tag{5.14}$$

We shall also prove convergence of both $E_n v_n^{[t/\tau_n]}$ and $E_n w_n^{[t/\varepsilon_n]}$ to $u(t)$.

The following assumptions about space discretization will be used:

(A) (a) $X, X_1, X_2, \ldots$ are all real or all complex Banach spaces. All norms will be denoted by $\|\cdot\|$.

(b) $P_n \in \mathfrak{B}(X, X_n)$, $p \in [0, \infty)$ are such that $\|P_n x\| \leq p\|x\|$ for $n \geq 1$, $x \in X$.

(c) $E_n \in \mathfrak{B}(X_n, X)$, $q \in [0, \infty)$ are such that $\|E_n x\| \leq q\|x\|$ for $n \geq 1$, $x \in X_n$.

(d) $P_n E_n x = x$ for $n \geq 1$, $x \in X_n$.

About A_n it will be assumed that

(B) $A_n \in \mathfrak{B}(X_n)$ are such that for some $M < \infty$, $a \in \mathbb{R}$ we have

$$\|e^{-A_n t}\| \leq M e^{-at} \quad \text{for} \quad t \geq 0,\ n \geq 1.$$

This is called the stability condition of approximations A_n. While all norms are equivalent in finite dimensional spaces, the equivalence is in general not uniform in n. So the norms on spaces X_n have to be chosen carefully in order to insure that the bound in **B** does not depend on n. Observe that **B** and Theorem 4.3.2 imply that $(-\infty, a) \subset \rho(A_n)$ for $n \geq 1$. A bit more restrictive condition will be needed to prove convergence when the explicit Euler method is used.

The only connection between A and A_n is represented by

(C) A is a densely defined linear operator in X and there exists $\lambda_0 \in (-\infty, a) \cap \rho(A)$ such that for all x in a dense subset of X,

$$\lim_{n \to \infty} \|E_n (A_n - \lambda_0)^{-1} P_n x - (A - \lambda_0)^{-1} x\| = 0.$$

Note that **C** requires that the approximations A_n can be used to solve the time independent problem,

$$Au - \lambda_0 u = x,$$

for all x in a dense subset of X. It turns out that this formulation of assumptions is especially convenient to apply to the Galerkin method of approximating solutions of parabolic and wave type equations (Sections 5.6, 5.7, 5.8). For finite difference approximations (Section 5.5), it can be more convenient to verify the following condition instead.

(C') A is a densely defined linear operator in X, $\lambda_0 \in (-\infty, a) \cap \rho(A)$ and

$$\lim_{n \to \infty} \|A_n P_n u - P_n A u\| = \lim_{n \to \infty} \|E_n P_n u - u\| = 0 \quad \text{for all} \quad u \in \mathcal{D}(A).$$

Lemma 5.4.1 *Assume* **A** *and* **B**. *Then,* **C'** *implies* **C**.

PROOF Choose $x \in X$ and let $u = (A - \lambda_0)^{-1} x$. Note that

$$E_n (A_n - \lambda_0)^{-1} P_n x - u = E_n P_n u - u + E_n (A_n - \lambda_0)^{-1} (P_n A u - A_n P_n u),$$

hence **B**, the bound in Theorem 4.3.2 and **C'** imply **C**. □

The following is a version of the Trotter Theorem.

Theorem 5.4.2 *Assume* **A**, **B** *and* **C**. *Then* $-A$ *is the generator of a strongly continuous semigroup* $\{Q(t)\}_{t\geq 0}$ *on* X. *Moreover,* $\|Q(t)\| \leq pqMe^{-at}$ *for* $t \geq 0$ *and*

$$\lim_{n\to\infty} \sup_{t\geq 0} e^{bt}\|E_n e^{-A_n t}P_n x - Q(t)x\| = 0 \quad \textit{for all} \quad x \in X,\ b \in (-\infty, a).$$

PROOF Abbreviate $R_n(\lambda) = (A_n - \lambda)^{-1}$ and $R(\lambda) = (A-\lambda)^{-1}$. Theorem 4.3.2 implies

$$\|R_n(\lambda)^m\| \leq M(a-\lambda)^{-m} \quad \text{for} \quad \lambda < a,\ m \geq 1,\ n \geq 1. \tag{5.15}$$

This implies that $\lim_{n\to\infty} E_n R_n(\lambda_0)P_n x = R(\lambda_0)x$ for every $x \in X$.

If $\lambda \in (-\infty, a) \cap \rho(A)$, then

$$E_n R_n(\lambda)P_n - R(\lambda) = (1 + (\lambda - \lambda_0)E_n R_n(\lambda)P_n)(E_n R_n(\lambda_0)P_n - R(\lambda_0))(A - \lambda_0)R(\lambda).$$

Therefore,

$$\text{if } \lambda \in (-\infty, a) \cap \rho(A), \text{ then } \lim_{n\to\infty} E_n R_n(\lambda)P_n x = R(\lambda)x \text{ for } x \in X. \tag{5.16}$$

(5.16) and (5.15) imply

$$\text{if } \lambda \in (-\infty, a) \cap \rho(A), \text{ then } \|R(\lambda)\| \leq pqM(a-\lambda)^{-1}. \tag{5.17}$$

If $\mu \in (-\infty, a)$ and $\mu \notin \rho(A)$, then $\|R(\lambda)\|$ should approach $+\infty$ as λ goes from λ_0 to μ. However, by (5.17) this is not possible. Therefore, $(-\infty, a) \subset \rho(A)$.

Induction on m gives

$$\lim_{n\to\infty} \|E_n R_n(\lambda)^m P_n x - R(\lambda)^m x\| = 0 \quad \text{for} \quad x \in X,\ m \geq 1,\ \lambda < a. \tag{5.18}$$

This and (5.15) imply that

$$\|R(\lambda)^m\| \leq pqM(a-\lambda)^{-m} \quad \text{for} \quad m \geq 1,\ \lambda < a.$$

Therefore, Theorem 4.3.5 implies that $-A$ is the generator of a strongly continuous semigroup $\{Q(t)\}_{t\geq 0}$ with the stated bound.

Choose now $\lambda < a$, $b < a$, $x \in X$. Since

$$\frac{d}{ds} e^{-A_n(t-s)}R_n(\lambda)P_n Q(s)R(\lambda)x = e^{-A_n(t-s)}P_n(R(\lambda) - E_n R_n(\lambda)P_n)Q(s)x,$$

we have that

$$\begin{aligned} & e^{bt}\|E_nR_n(\lambda)(P_nQ(t)-e^{-A_nt}P_n)R(\lambda)x\| \\ &= \left\|\int_0^t e^{b(t-s)}E_ne^{-A_n(t-s)}P_n(R(\lambda)-E_nR_n(\lambda)P_n)e^{bs}Q(s)x\,ds\right\| \\ &\le pqM\int_0^\infty \|(R(\lambda)-E_nR_n(\lambda)P_n)e^{bs}Q(s)x\|ds \end{aligned}$$

and the DCT implies

$$\lim_{n\to\infty}\sup_{t\ge 0} e^{bt}\|E_nR_n(\lambda)(P_nQ(t)-e^{-A_nt}P_n)R(\lambda)x\| = 0. \tag{5.19}$$

Note that

$$\begin{aligned} & \mathrm{e}^{bt}(E_nR_n(\lambda)P_nQ(t)-Q(t)R(\lambda))R(\lambda)x \\ &= (E_nR_n(\lambda)P_n-R(\lambda))e^{bt}Q(t)R(\lambda)x \\ &= (E_nR_n(\lambda)P_n-R(\lambda))R(\lambda)x \\ &\quad +\int_0^t (E_nR_n(\lambda)P_n-R(\lambda))\mathrm{e}^{bs}Q(s)(b-A)R(\lambda)x\,ds. \end{aligned}$$

Hence the DCT and (5.19) imply that

$$\lim_{n\to\infty}\sup_{t\ge 0} e^{bt}\|(Q(t)R(\lambda)-E_nR_n(\lambda)e^{-A_nt}P_n)R(\lambda)x\| = 0. \tag{5.20}$$

Since

$$\begin{aligned} & E_nR_n(\lambda)e^{-A_nt}P_nR(\lambda)x-E_ne^{-A_nt}P_nR(\lambda)^2x \\ &\quad = E_ne^{-A_nt}P_n(E_nR_n(\lambda)P_n-R(\lambda))R(\lambda)x, \end{aligned}$$

we see that (5.20) implies

$$\lim_{n\to\infty}\sup_{t\ge 0} e^{bt}\|(Q(t)-E_ne^{-A_nt}P_n)R(\lambda)^2x\| = 0.$$

We are done because the range of $R(\lambda)^2$ is equal to $\mathcal{D}(A^2)$, which is dense in X by Theorem 4.3.1. □

For the explicit Euler method we will need a bit stronger assumption:

(B') $\tau_n \in (0,\infty)$, $A_n \in \mathfrak{B}(X_n)$ are such that for some $M<\infty$, $\tilde{a}\in\mathbb{R}$ we have

$$\|(1-\tau_nA_n)^k\| \le Me^{-\tilde{a}k\tau_n} \quad \text{for} \quad n\ge 1,\ k\ge 0;$$

moreover, $\lim_{i\to\infty}\tau_i = 0$.

Corollary 5.4.3 **B'** *implies* **B** *if* $a \leq (1 - e^{-\tau_n \tilde{a}})/\tau_n$ *for* $n \geq 1$.

This Corollary follows from the following Lemma

Lemma 5.4.4 *Suppose that* X *is a Banach space and that* $A \in \mathfrak{B}(X)$ *is such that*

$$\|(1 - \tau A)^k\| \leq M e^{-\tilde{a}k\tau} \quad \text{for} \quad k \geq 0$$

for some $\tau \in (0, \infty)$, $\tilde{a} \in \mathbb{R}$ *and* $M \in [1, \infty)$. *Let* $a \leq (1 - e^{-\tilde{a}\tau})/\tau$. *Then*

$$\|e^{-At}\| \leq M e^{-at} \quad \text{for} \quad t \geq 0$$

and

$$\|(1 - \tau A)^{[t/\tau]} x - e^{-At} x\| \leq \tau M^3 e^{-at + 2|a|\tau} (t \|A^2 x\| + \|Ax\|) \quad \text{for} \quad x \in X,\ t \geq 0.$$

PROOF Let $B = 1 - \tau A$. In view of (4.36) we have

$$\begin{aligned} e^{-At} &= e^{(B-1)t/\tau} = e^{-t/\tau} e^{Bt/\tau} \\ \|e^{-At}\| &\leq e^{-t/\tau} \sum_{k=0}^{\infty} \frac{t^k \|B^k\|}{\tau^k k!} \\ &\leq M e^{-t/\tau} \sum_{k=0}^{\infty} \frac{(te^{-\tilde{a}\tau}/\tau)^k}{k!} \leq M e^{-at}. \end{aligned}$$

Choose $x \in X$ and let $e_k = B^k x - e^{-Ak\tau} x$ for $k \geq 0$. Induction implies

$$e_k = \sum_{j=0}^{k-1} B^j e^{-A(k-1-j)\tau} e_1 \quad \text{for} \quad k \geq 1$$

and, since $a \leq \tilde{a}$, we have

$$\|e_k\| \leq M^2 \sum_{j=0}^{k-1} e^{-\tilde{a}j\tau} e^{-a(k-1-j)\tau} \|e_1\| \leq k M^2 e^{-a(k-1)\tau} \|e_1\|.$$

Note

$$e_1 = x - \tau A x - e^{-\tau A} x = \int_0^{\tau} (t - \tau) e^{-At} A^2 x dt,$$

$$\|e_1\| \leq \tau^2 M e^{-a\theta_1 \tau} \|A^2 x\|,$$

where $\theta_1 \in (0, 1)$. Hence,

$$\|e_k\| \leq \tau^2 k M^3 e^{-a(k-1+\theta_1)\tau} \|A^2 x\| \quad \text{for} \quad k \geq 0.$$

Thus, if $t \geq 0$ and $k = [t/\tau]$, hence, $0 \leq t - k\tau < \tau$, then

$$\|B^{[t/\tau]}x - e^{-At}x\| \leq \tau^2 k M^3 e^{-a(k-1+\theta_1)\tau}\|A^2 x\| + R \tag{5.21}$$

where

$$R = \|e^{-Ak\tau}x - e^{-At}x\| \leq M\int_{k\tau}^{t} e^{-as}\|Ax\|ds \leq \tau M e^{-at+|a|\tau}\|Ax\|.$$

Using this in (5.21) implies the conclusion of the Lemma. □

Theorem 5.4.5 *Assume* **A**, **B'** *and* **C**. *Then,* $-A$ *is the generator of a strongly continuous semigroup* $\{Q(t)\}_{t\geq 0}$ *on* X. *Moreover,* $\|Q(t)\| \leq pqMe^{-\tilde{a}t}$ *for* $t \geq 0$ *and*

$$\lim_{n\to\infty}\sup_{t\geq 0} e^{bt}\|E_n(1-\tau_n A_n)^{[t/\tau_n]}P_n x - Q(t)x\| = 0 \quad \textit{for all} \quad x \in X,\ b \in (-\infty, \tilde{a}).$$

PROOF We shall prove the statements of the Theorem first for the case when $\tilde{a}$ is replaced with a as given in Corollary 5.4.3. Taking the limit $n \to \infty$ then implies the original conclusion.

Abbreviate $R_n = (A_n - \lambda_0)^{-1}$ and $R = (A - \lambda_0)^{-1}$. Take $t \geq 0$, $n \geq 1$, $x \in X$ and note

$$(E_n(1-\tau_n A_n)^{[t/\tau_n]}P_n - Q(t))R^2 x = I_{1n}(t) - I_{2n}(t) + I_{3n}(t),$$

where

$$\begin{aligned} I_{1n}(t) &= E_n((1-\tau_n A_n)^{[t/\tau_n]} - e^{-A_n t})R_n^2 P_n x \\ I_{2n}(t) &= E_n((1-\tau_n A_n)^{[t/\tau_n]} - e^{-A_n t})P_n(E_n R_n^2 P_n - R^2)x \\ I_{3n}(t) &= (E_n e^{-A_n t}P_n - Q(t))R^2 x. \end{aligned}$$

Theorem 5.4.2 implies that

$$\lim_{n\to\infty}\sup_{t\geq 0} e^{bt}\|I_{3n}(t)\| = 0 \quad \text{for} \quad b < a.$$

Since

$$\|I_{2n}(t)\| \leq pqM(e^{-\tilde{a}\tau_n[t/\tau_n]} + e^{-at})\|E_n R_n^2 P_n x - R^2 x\|,$$

(5.18) implies

$$\lim_{n\to\infty}\sup_{t\geq 0} e^{bt}\|I_{2n}(t)\| = 0 \quad \text{for} \quad b < a.$$

Using Lemma 5.4.4 we obtain

$$\|I_{1n}(t)\| \leq \tau_n q M^3 e^{-at+2|a|\tau_n}(t\|A_n^2 R_n^2 P_n x\| + \|A_n R_n^2 P_n x\|)$$

and, since $\|R_n\| \le M/(a-\lambda_0)$, $\|A_nR_n\| \le 1 + |\lambda_0|M/(a-\lambda_0)$, we have

$$\lim_{n\to\infty} \sup_{t\ge 0} e^{bt}\|I_{1n}(t)\| = 0 \quad \text{for} \quad b < a.$$

This proves the limit in the Theorem for $x \in \mathcal{D}(A^2)$, which is dense in X by Theorem 4.3.1, and hence for all $x \in X$. □

Convergence for the implicit Euler method follows directly from convergence for the explicit Euler method:

Theorem 5.4.6 *Assume* **A**, **B**, **C** *and that* $\{\varepsilon_n\}_{n=1}^{\infty} \subset (0,\infty)$, $\lim_{n\to\infty}\varepsilon_n = 0$. *Then*

$$\lim_{n\to\infty} \sup_{t\ge 0} e^{bt}\|E_n(1+\varepsilon_nA_n)^{-[t/\varepsilon_n]}P_nx - Q(t)x\| = 0 \quad \textit{for all} \quad x \in X,\ b \in (-\infty, a).$$

PROOF Choose N so that $1 + a\varepsilon_n > 0$ for $n \ge N$. Define

$$A_n' = A_n(1+\varepsilon_nA_n)^{-1} \quad \text{for} \quad n \ge N.$$

Note $1 - \varepsilon_nA_n' = (1+\varepsilon_nA_n)^{-1}$; hence (5.15) implies

$$\|(1-\varepsilon_nA_n')^k\| \le M(1+\varepsilon_na)^{-k} \le Me^{-\tilde{a}\varepsilon_nk} \quad \text{for} \quad k \ge 0,\ n \ge N,$$

where $\tilde{a}\varepsilon_n \le \ln(1+a\varepsilon_n)$. Hence **B'** is satisfied with A_n', ε_n in place of A_n, τ_n for $n \ge N$. Since $\|A_n(A_n-\lambda_0)^{-1}\| \le 1 + |\lambda_0|M/(a-\lambda_0)$ and

$$(A_n'-\lambda_0)^{-1} = (A_n-\lambda_0)^{-1} + \varepsilon_nA_n^2(A_n-\lambda_0)^{-2}(1-\lambda_0\varepsilon_nA_n(A_n-\lambda_0)^{-1})^{-1}$$

when n is large enough, we see that **C** is also satisfied with A_n' in place of A_n. Thus, Theorem 5.4.5 implies the conclusion. □

For simplicity of presentation we shall assume that the nonlinearity satisfies

(D) $\tau \in (0,\infty)$, $H : [0,\tau] \times X \to X$ is continuous and such that for some $L < \infty$,

$$\|H(t,x) - H(t,y)\| \le L\|x-y\| \quad \text{for} \quad 0 \le t \le \tau,\ x,y \in X.$$

Using Lemma 5.2.1 one can handle more general nonlinearity, see the example in the next Section.

Note, when $-A$ is the generator of a strongly continuous semigroup on X, then **D** implies the existence of a unique mild solution of (5.11) on $[0,\tau]$ for every $x \in X$, see Theorem 5.1.2. Likewise, (5.12) has a unique mild solution on $[0,\tau]$. However, since A_n are bounded operators, Corollary 4.7.5 implies that the mild solution is differentiable and satisfies (5.12).

Theorem 5.4.7 *Assume* **A, B, C, D** *and pick* $x \in X$. *Let* u *be the mild solution of (5.11) and let* u_n *be the solutions of (5.12). Then*

$$\lim_{n\to\infty} \sup_{0\le t\le\tau} \|E_n u_n(t) - u(t)\| = 0.$$

PROOF Note that for $t \in [0,\tau]$, $n \ge 1$, we have

$$u_n(t) = e^{-A_n t}P_n x + \int_0^t e^{-A_n(t-s)}P_n H(s, E_n u(s))ds$$

$$u(t) = Q(t)x + \int_0^t Q(t-s)H(s,u(s))ds. \tag{5.22}$$

Hence

$$\begin{aligned} u(t) - E_n u_n(t) &= Q(t)x - E_n e^{-A_n t}P_n x \\ &\quad + \int_0^t (Q(t-s) - E_n e^{-A_n(t-s)}P_n)H(s,u(s))ds \\ &\quad + \int_0^t E_n e^{-A_n(t-s)}P_n(H(s,u(s)) - H(s,E_n u_n(s)))ds. \end{aligned}$$

Define

$$r_n(t) = \|u(t) - E_n u_n(t)\|,$$

$$\varepsilon_n = \sup_{0\le t\le\tau} \|Q(t)x - E_n e^{-A_n t}P_n x\|,$$

$$d_n(s) = \sup_{0\le t\le\tau} \|(Q(t) - E_n e^{-A_n t}P_n)H(s,u(s))\| \le 2C\|H(s,u(s))\|,$$

where $C = pqM(1 + e^{-a\tau})$. Note that $\lim_{n\to\infty}\varepsilon_n = 0$ by Theorem 5.4.2, which together with the DCT also implies that $\int_0^\tau d_n(s)ds = 0$. Since

$$r_n(t) \le \varepsilon_n + \int_0^\tau d_n(s)ds + LC\int_0^t r_n(s)ds \quad \text{for} \quad t \in [0,\tau],\ n \ge 1,$$

the Gronwall Theorem 5.1.1 implies

$$r_n(t) \le \left(\varepsilon_n + \int_0^\tau d_n(s)ds\right) e^{LC\tau},$$

which completes the proof. □

Theorem 5.4.8 *Assume* **A, B', C, D** *and pick* $x \in X$. *Let* u *be the mild solution of (5.11) and let* v_n^k *be given by (5.13). Then*

$$\lim_{n\to\infty} \sup_{0\le t\le\tau} \|E_n v_n^{[t/\tau_n]} - u(t)\| = 0.$$

PROOF Let $B_n = 1 - \tau_n A_n$. Observe that for $n \geq 1$, $k \geq 1$,

$$v_n^k = B_n^k P_n x + \tau_n \sum_{j=0}^{k-1} B_n^{k-1-j} P_n H(j\tau_n, E_n v_n^j).$$

Hence, for $t \geq 0$,

$$E_n v_n^{[t/\tau_n]} = E_n B_n^{[t/\tau_n]} P_n x + \int_0^{[t/\tau_n]\tau_n} D_n(t,s) H([s/\tau_n]\tau_n, E_n v_n^{[s/\tau_n]}) ds,$$

where $D_n(t,s) = E_n B_n^{[t/\tau_n]-1-[s/\tau_n]} P_n$. Using (5.22) we obtain

$$u(t) - E_n v_n^{[t/\tau_n]} = \sum_{i=1}^{6} I_{in}(t),$$

where

$$\begin{aligned}
I_{1n}(t) &= Q(t)x - E_n B_n^{[t/\tau_n]} P_n x \\
I_{2n}(t) &= \int_{[t/\tau_n]\tau_n}^{t} Q(t-s) H(s, u(s)) ds \\
I_{3n}(t) &= \int_0^{[t/\tau_n]\tau_n} (Q(t-s) - Q(t - [s/\tau_n]\tau_n - \tau_n)) H(s, u(s)) ds \\
I_{4n}(t) &= \int_0^{[t/\tau_n]\tau_n} (Q(t - [s/\tau_n]\tau_n - \tau_n) - D_n(t,s)) H(s, u(s)) ds \\
I_{5n}(t) &= \int_0^{[t/\tau_n]\tau_n} D_n(t,s)(H(s, u(s)) - H([s/\tau_n]\tau_n, u(s))) ds \\
I_{6n}(t) &= \int_0^{[t/\tau_n]\tau_n} D_n(t,s)(H([s/\tau_n]\tau_n, u(s)) - H([s/\tau_n]\tau_n, E_n v_n^{[s/\tau_n]})) ds.
\end{aligned}$$

Define $r_n(t) = \|u(t) - E_n v_n^{[t/\tau_n]}\|$, $\delta_{in} = \sup_{0 \leq t \leq \tau} \|I_{in}(t)\|$ and let C be the upper bound of

$$\|Q(t)\|, \ \|E_n B_n^{[t/\tau_n]} P_n\|, \ \|H(t, u(t))\| \quad \text{for} \quad t \in [0, \tau], \ n \geq 1.$$

Theorem 5.4.5 implies $\lim_{n\to\infty} \delta_{1n} = 0$.

Note that $\delta_{2n} \leq C^2 \tau_n \to 0$ as $n \to \infty$.

Note that

$$\delta_{3n} \leq C \int_0^{\tau} \|(Q([s/\tau_n]\tau_n + \tau_n - s) - 1) H(s, u(s))\| ds;$$

hence the strong continuity of Q and the DCT imply $\lim_{n\to\infty} \delta_{3n} = 0$.

Note that

$$\delta_{4n} \leq \int_0^\tau \sup_{0\leq t\leq \tau} \|(Q(t) - E_n B_n^{[t/\tau_n]} P_n) H(s, u(s))\| ds;$$

hence Theorem 5.4.5 and the DCT imply $\lim_{n\to\infty} \delta_{4n} = 0$.

Since $\delta_{5n} \leq C \int_0^\tau \|H(s, u(s)) - H([s/\tau_n]\tau_n, u(s))\| ds$, the continuity of H and the DCT imply $\lim_{n\to\infty} \delta_{5n} = 0$.

Since

$$r_n(t) \leq \sum_{i=1}^{5} \delta_{in} + LC \int_0^t r_n(s) ds,$$

the Gronwall Theorem 5.1.1 implies

$$r_n(t) \leq \sum_{i=1}^{5} \delta_{in} e^{LC\tau} \quad \text{for} \quad t \in [0, \tau],\ n \geq 1,$$

which completes the proof. □

Theorem 5.4.9 *Assume* **A**, **B**, **C**, **D**, *pick* $x \in X$ *and* $\{\varepsilon_n\}_{n=1}^\infty \subset (0, \infty)$ *such that* $\lim_{n\to\infty} \varepsilon_n = 0$. *Let* u *be the mild solution of (5.11) and let* w_n^k *be given by (5.14). Then*

$$\lim_{n\to\infty} \sup_{0\leq t\leq \tau} \|E_n w_n^{[t/\tau_n]} - u(t)\| = 0.$$

PROOF This proof is identical to that for Theorem 5.4.8, with the following obvious modifications: replace τ_n by ε_n, v by w, set new $B_n = (1 + \varepsilon_n A_n)^{-1}$ and use Theorem 5.4.6 in place of Theorem 5.4.5. □

5.5 Example: Finite Difference Method

We shall investigate the following parabolic equation:

$$\left.\begin{aligned} u_t(x,t) = a(x) u_{xx}(x,t) + F(u(x,t)) \quad &\text{for} \quad 0 \leq x \leq 1,\ t \geq 0 \\ u(0,t) = u(1,t) = 0 \quad &\text{for} \quad t \geq 0 \\ u(x,0) = u_0(x) \quad &\text{for} \quad 0 \leq x \leq 1, \end{aligned}\right\} \tag{5.23}$$

where $a \in C([0,1], (0,\infty))$, $F \in C^1(\mathbb{R}, \mathbb{R})$, $F(0) = 0$ and

$$u_0 \in X \equiv \{f \in C([0,1], \mathbb{R}) \mid f(0) = f(1) = 0\}.$$

If $u_{n,i}^k$ is to approximate $u(i\delta_n, k\tau_n)$, $\delta_n = \frac{1}{n+1}$, then an obvious finite difference discretization is

$$\frac{u_{n,i}^{k+1} - u_{n,i}^k}{\tau_n} = a(i\delta_n)\frac{u_{n,i+1}^k + u_{n,i-1}^k - 2u_{n,i}^k}{\delta_n^2} + F(u_{n,i}^k) \quad \text{for} \quad 1 \le i \le n,\ n \ge 1,\ k \ge 0$$

$$u_{n,0}^k = u_{n,n+1}^k = 0 \quad \text{for} \quad n \ge 1,\ k \ge 0$$

$$u_{n,i}^0 = u_0(i\delta_n) \quad \text{for} \quad 1 \le i \le n,\ n \ge 1.$$

Assuming only the stability condition,

$$2\tau_n a(i\delta_n) \le \delta_n^2 \quad \text{for} \quad 1 \le i \le n,\ n \ge 1, \tag{5.24}$$

we shall prove that linear interpolations of $u_{n,\cdot}^{[t/\tau_n]}$ converge to the mild solution u of (5.23). Moreover, it will be shown that the convergence is uniform on any finite time interval that is shorter than the time interval of existence of u.

Let $\|\cdot\|$ be the max norm on X and on $X_n \equiv \mathbb{R}^n$, $n \ge 1$. Define a 'projection' $P_n \in \mathfrak{B}(X, X_n)$ by

$$(P_n f)_i = f(i\delta_n) \quad \text{for} \quad 1 \le i \le n,\ n \ge 1.$$

For $n \ge 1$, define an 'interpolation' $E_n \in \mathfrak{B}(X_n, X)$ by

$$(E_n \mathrm{c})(x) = c_i + (\tfrac{x}{\delta_n} - i)(c_{i+1} - c_i) \quad \text{for} \quad i \le \tfrac{x}{\delta_n} \le i+1,\ 0 \le i \le n,$$

where $\mathrm{c} = (c_1, \ldots, c_n)^\top \in X_n$ and $c_0 = c_{n+1} = 0$. Note that assumption **A** of Section 5.4 is satisfied, with $p = q = 1$.

For $n \ge 1$ define $A_n \in \mathfrak{B}(X_n)$ by

$$(A_n \mathrm{c})_i = a(i\delta_n)\frac{2c_i - c_{i+1} - c_{i-1}}{\delta_n^2} \quad \text{for} \quad 1 \le i \le n,$$

where $\mathrm{c} = (c_1, \ldots, c_n)^\top \in X_n$ and $c_0 = c_{n+1} = 0$. Note that

$$((1 - \tau_n A_n)\mathrm{c})_i = (1 - 2\tau_n \tfrac{a(i\delta_n)}{\delta_n^2})c_i + \tau_n \tfrac{a(i\delta_n)}{\delta_n^2} c_{i+1} + \tau_n \tfrac{a(i\delta_n)}{\delta_n^2} c_{i-1};$$

hence

$$\|1 - \tau_n A_n\| \le \max_i \{|1 - 2\tau_n \tfrac{a(i\delta_n)}{\delta_n^2}| + 2\tau_n \tfrac{a(i\delta_n)}{\delta_n^2}\}.$$

Therefore choosing $\tau_n > 0$ so that they satisfy the stability condition (5.24) implies

$$\|1 - \tau_n A_n\| \le 1 \quad \text{for} \quad n \ge 1.$$

Thus, assumption **B'** of Section 5.4 is satisfied, with $M = 1$, $\tilde{a} = 0$. Corollary 5.4.3 implies that **B** is also satisfied.

Define A by

$$Af = -af'' \quad \text{for} \quad f \in \mathcal{D}(A) \equiv \{f \in C^2[0,1] \mid f, f'' \in X\}.$$

It is easy to show that $0 \in \rho(A)$ by showing that

$$(A^{-1}f)(x) = \int_0^x \frac{(1-x)y}{a(y)} f(y)dy + \int_x^1 \frac{(1-y)x}{a(y)} f(y)dy \quad \text{for} \quad f \in X,\ x \in [0,1].$$

Since the resolvent set is open, we have that $\lambda \in \rho(A)$ when $|\lambda|$ is small. If $f \in \mathcal{D}(A)$ and $n \geq 1$, then for $1 \leq i \leq n$ we have

$$\begin{aligned}(A_n P_n f - P_n A f)_i &= \tfrac{a(i\delta_n)}{\delta_n^2}(\delta_n^2 f''(i\delta_n) + 2f(i\delta_n) - f(i\delta_n + \delta_n) - f(i\delta_n - \delta_n)) \\ &= \tfrac{a(i\delta_n)}{\delta_n^2} \int_{-\delta_n}^{\delta_n} (\delta_n - |t|)(f''(i\delta_n) - f''(i\delta_n + t))dt\end{aligned}$$

and hence the uniform continuity of f'' implies that

$$\lim_{n\to\infty} \|A_n P_n f - P_n A f\| = 0 \quad \text{for} \quad f \in \mathcal{D}(A).$$

If $f \in \mathcal{D}(A)$ and $n \geq 1$, then for $i \leq \frac{x}{\delta_n} \leq i+1$, $0 \leq i \leq n$ we have

$$\begin{aligned}(E_n P_n f - f)(x) &= f(i\delta_n) + (\tfrac{x}{\delta_n} - i)(f(i\delta_n + \delta_n) - f(i\delta_n)) - f(x) \\ &= \tfrac{1}{\delta_n} \int_0^{x - i\delta_n} \int_0^{\delta_n} (f'(i\delta_n + t) - f'(i\delta_n + s))dt\, ds \\ \|E_n P_n f - f\| &\leq \delta_n^2 \|f''\|.\end{aligned}$$

Therefore assumption **C'** of Section 5.4 is satisfied and so is **C** by Lemma 5.4.1.

Theorem 5.4.5 implies that $-A$ is the generator of a contraction semigroup. This was shown independently in Example 4.3.9. Thus, Theorem 5.2.6 applies, giving the mild solution u of (5.23) on the maximal interval of existence $[0, \tau_{max})$. Note, if $\tau_{max} < \infty$, then u and $F(u)$ have to blow-up as $t \to \tau_{max}$. Lemma 5.2.1 implies that for any $\tau \in (0, \tau_{max})$ we can find a Lipschitz continuous nonlinearity H which agrees with F in a neighborhood of u on $[0,\tau]$. Theorem 5.4.8 implies that

$$\lim_{n\to\infty} \sup_{0\leq t\leq\tau} \|E_n v_n^{[t/\tau_n]} - u(t)\| = 0$$

where v_n^k is given by (5.13). Thus, for large enough n, $E_n v_n^{[\cdot/\tau_n]}$ is in the neighborhood of u where F and H agree and therefore $v_n^{[t/\tau_n]} = u_{n,\cdot}^{[t/\tau_n]}$ for $t \in [0,\tau]$. This proves that linear interpolations of $u_{n,\cdot}^{[\cdot/\tau_n]}$ converge uniformly to u on $[0,\tau]$.

It is natural to expect that in order to approximate u to a given tolerance it would be optimal to choose about the same size space (δ_n) and time (τ_n) steps. However, the stability condition (5.24) requires that the time steps be much smaller than the space steps. This is due to the use of the explicit Euler method for space discretization (5.13). If one would choose the implicit version (5.14) instead, then there would be no such restriction (any $\varepsilon_n \in (0,\infty)$, such that $\lim_{n\to\infty} \varepsilon_n = 0$, would do).

5.6 Example: Galerkin Method for Parabolic Equations

Here the results of Section 5.4 will be applied to

$$u'(t) + Au(t) = H(t, u(t)), \quad u(0) = x, \tag{5.25}$$

where A is the operator associated with a sectorial form $\mathfrak{F}$. This covers most parabolic equations set in a Hilbert space. See Section 3.7.

Assume that $\mathcal{H}$ is a Hilbert space with an inner product $(\cdot, \cdot)$ and norm $\|\cdot\|$, $[\cdot, \cdot]$ is an inner product on $\mathcal{V} \subset \mathcal{H}$ with the corresponding norm denoted by $|\cdot|$ and $\mathfrak{F}$ is a sectorial sesquilinear form on $\mathcal{V}$ such that hypotheses **H1**, **H2**, **H3** of the section on Sectorial Forms are satisfied. Here, assume in addition that

(**H4**) $\mathcal{V}_1, \mathcal{V}_2, \cdots$ are finite dimensional subspaces of $\mathcal{V}$ such that

$$\lim_{n \to \infty} \inf_{z \in \mathcal{V}_n} |y - z| = 0$$

for all y in a dense (in $|\cdot|$ norm) subset of $\mathcal{V}$.

Observe that hypothesis **H4** is exactly the same one as the one in Theorem 2.8.11, where convergence of approximations of solutions of elliptic problems is proved. **H1** - **H4** are assumed to be in effect throughout this Section. We shall first demonstrate how assumptions **A**, **B** and **C** of Section 5.4 can be satisfied without imposing any other conditions. Then a version of Theorem 5.4.7 that is suitable for obtaining Galerkin approximations will be presented. Finally, in Example 5.6.2, it will be shown that the Galerkin approximations can be used to deduce existence of solutions even when the nonlinearity is not locally Lipschitz continuous.

Let $X_n \equiv \mathcal{V}_n$ be equipped with norm $\|\cdot\|$. Let P_n be the orthogonal projection of $\mathcal{H}$ onto X_n. Let E_n be the identity map from X_n to $X \equiv \mathcal{H}$. Thus, assumption **A** of Section 5.4 is clearly satisfied, with $p = q = 1$.

Observe that assumptions **H1**, **H2** and **H3** remain valid, with the same constants, if we replace both $\mathcal{H}$ and $\mathcal{V}$ with X_n. Let the operator $A_n \in \mathfrak{B}(X_n)$ be the operator associated with $\mathfrak{F}$-restricted to X_n. Hence, A_n is defined by

$$(A_n y, z) = \mathfrak{F}(y, z) \quad \text{for all} \quad y, z \in X_n, \ n \geq 1. \tag{5.26}$$

(1) of Theorem 2.8.2 and Figure 2.1 imply that

$$\|(A_n - \lambda)^{-1}\| \leq \frac{1}{a + M_3 M_1^{-2} - \lambda} \quad \text{for} \quad \lambda < a + M_3 M_1^{-2}.$$

Hence, the Hille-Yosida Theorem 4.3.5 implies that

$$\|e^{-A_n t}\| \leq e^{-(a + M_3 M_1^{-2})t} \quad \text{for} \quad t \geq 0, \ n \geq 1.$$

This shows that assumption **B** of Section 5.4 is satisfied.

Theorem 2.8.11 implies that

$$\lim_{n\to\infty} |(A_n - \lambda)^{-1} P_n x - (A - \lambda)^{-1} x| = 0 \quad \text{for all} \quad x \in X,\ \lambda < a + M_3 M_1^{-2};$$

hence **C** of Section 5.4 is also satisfied.

Therefore, all results of Section 5.4 are applicable. For example, by using (5.26) to characterize the solution of (5.12), we can restate Theorem 5.4.7 as follows:

Theorem 5.6.1 *Assume* **D** *of Section 5.4, choose any* $x \in \mathcal{H}$ *and let* u *denote the mild solution of (5.25). Then for every* $n \geq 1$ *there exists a unique* $u_n \in C^1([0,\tau], \mathcal{V}_n)$ *such that for all* $z \in \mathcal{V}_n$ *we have* $(u_n(0), z) = (x, z)$ *and*

$$(u_n'(t), z) + \mathfrak{F}(u_n(t), z) = (H(t, u_n(t)), z) \quad \textit{for} \quad t \in [0,\tau].$$

Moreover,

$$\lim_{n\to\infty} \sup_{0\leq t\leq\tau} \|u_n(t) - u(t)\| = 0.$$

EXAMPLE 5.6.2 We shall first prove convergence of finite element approximations to the following parabolic equation:

$$u_t(x,t) = u_{xx}(x,t) + G(u(x,t)) \quad \text{for} \quad 0 \leq x \leq 1,\ t \geq 0$$
$$u(0,t) = u(1,t) = 0 \quad \text{for} \quad t \geq 0$$
$$u(x,0) = u_0(x) \quad \text{for} \quad 0 \leq x \leq 1,$$

where $u_0 \in L^2(0,1)$ is real valued, $G \in C^1(\mathbb{R},\mathbb{R})$ and G' is bounded. Then it will be shown how the condition that G' is bounded can be removed without losing the existence of the solution and the convergence of approximations.

Let $\mathcal{H}$ be the real $L^2(0,1)$ and let $\mathcal{V}$, $\mathfrak{F}$, $[\cdot,\cdot]$ be as in Example 2.8.1; hence, **H1**, **H2** and **H3** are satisfied.

For $n \geq 1$, let $\delta_n = 1/(n+1)$ and define $\mathcal{V}_n \subset \mathcal{V}$ as follows: $v \in \mathcal{V}_n$ if and only if

$$v \in C[0,1],\ v(0) = v(1) = 0,\ v \text{ is linear on } [i\delta_n, (i+1)\delta_n] \text{ for each } i = 0, 1, \ldots, n.$$

Lemma 2.11.1 implies that **H4** is satisfied for all $y \in \mathcal{D}(A)$, which is dense in $\mathcal{V}$ by Theorem 2.8.2, and therefore we can apply Theorem 5.6.1.

Let ψ be the hat function as in Section 2.11 and let $\psi_{n,i}(x) = \psi(x/\delta_n - i)$. Note that u_n can be represented as

$$u_n(x,t) = \sum_{i=1}^{n} c_{n,i}(t)\psi_{n,i}(x). \tag{5.27}$$

Taking z in Theorem 5.6.1 to be $\psi_{n,j}$, $1 \leq j \leq n$, $n \geq 1$ gives the ODEs,

$$\sum_{i=1}^{n} c_{n,i}'(t) \int_0^1 \psi_{n,i}(x)\psi_{n,j}(x)dx + \sum_{i=1}^{n} c_{n,i}(t) \int_0^1 \psi_{n,i}'(x)\psi_{n,j}'(x)dx = \delta_n g_{n,j}(c_{n,\cdot}(t)),$$

where

$$\delta_n g_{n,j}(c_{n,\cdot}) = \int_0^1 G\left(\sum_{i=1}^n c_{n,i}\psi_{n,i}(x)\right)\psi_{n,j}(x)dx$$

with the initial conditions

$$\sum_{i=1}^n c_{n,i}(0)\int_0^1 \psi_{n,i}(x)\psi_{n,j}(x)dx = \delta_n b_{n,j} \equiv \int_0^1 u_0(x)\psi_{n,j}(x)dx$$

which completely determine $c_{n,j}(t)$ for $t \geq 0$. Evaluation of those integrals gives

$$\frac{4c'_{n,j} + c'_{n,j+1} + c'_{n,j-1}}{6} + \frac{2c_{n,j} - c_{n,j+1} - c_{n,j-1}}{\delta_n^2} = g_{n,j}(c_{n,\cdot}) \tag{5.28}$$

$$\frac{4c_{n,j}(0) + c_{n,j+1}(0) + c_{n,j-1}(0)}{6} = b_{n,j}$$

for $1 \leq j \leq n$, $n \geq 1$, where $c_{n,0} = c_{n,n+1} = 0$ and

$$g_{n,j}(c_{n,\cdot}) = \int_{-1}^1 G(c_{n,j-1}\psi(x+1) + c_{n,j}\psi(x) + c_{n,j+1}\psi(x-1))\psi(x)dx$$

$$b_{n,j} = \int_{-1}^1 u_0(j\delta_n + x\delta_n)\psi(x)dx.$$

After solving (5.28) we obtain u_n via (5.27), and Theorem 5.6.1 implies convergence of u_n to the mild solution u on $[0,\infty)$. Note that it is numerically more difficult to solve (5.28) than the corresponding finite difference scheme in Section 5.5. However, the approach used here is directly applicable to a much larger class of parabolic problems. Note also that, here, the initial value u_0 can be any L^2 function.

Let us remove the condition that G' be bounded from now on. Note that G may not be even locally Lipschitz continuous in L^2 and hence nothing in this chapter may seem applicable. However, observe that one can still solve the ODEs (5.28) and hence obtain Galerkin approximations u_n on at least some finite time interval. Now suppose that one can show that

$$\gamma \equiv \sup_{0\leq x\leq 1, 0\leq t\leq\tau, n\geq 1} |u_n(x,t)| < \infty \tag{5.29}$$

for some $\tau > 0$. Choose $\phi \in C^1(\mathbb{R})$ such that $\phi(x) = x$ for $|x| \leq \gamma$ and $\phi(x) = 0$ for $|x| > \gamma + 1$. Define

$$\tilde{G}(x) = G(\phi(x)) \quad \text{for} \quad x \in \mathbb{R}$$

and note that $\tilde{G} \in C^1(\mathbb{R},\mathbb{R})$ and $\tilde{G}'$ is bounded. So we can apply the above theory with $\tilde{G}$ in place of G and obtain $\tilde{u}_n$ converging to the mild solution $\tilde{u}$. However, (5.29) implies that $c_{n,\cdot}$ also satisfy the modified (5.28) and hence, by the uniqueness for ODEs, $\tilde{c}_{n,\cdot} = c_{n,\cdot}$ on $[0,\tau]$ and therefore $\tilde{u}_n = u_n$ on $[0,\tau]$. Thus, u_n converge to $\tilde{u}$ on $[0,\tau]$ and $\tilde{u}$ can be interpreted as the mild solution on $[0,\tau]$ of the original problem.

Thus, when G' is not bounded, everything depends on being able to establish the bound (5.29). In some cases numerical evidence may be acceptable. It will be shown

now that bound (5.29) can always be satisfied for some small $\tau > 0$ when $u_0 \in L^\infty(0,1)$. To do this we will be using the max norm $\|\cdot\|_\infty$ on $\mathbb{R}^n$. Define an $n \times n$ matrix B_n by

$$B_n(i,i) = 2,\ B_n(i,i-1) = -1,\ B_n(i,i+1) = -1.$$

Note that (5.28) can be written as

$$(1 - \tfrac{1}{6}B_n)c'_{n,\cdot} + \delta_n^{-2}B_n c_{n,\cdot} = g_{n,\cdot}(c_{n,\cdot}).$$

A long but elementary calculation (Exercise 5) gives that

$$\|(1 - \tfrac{1}{6}B_n)^{-1}\| \leq 3 \quad \text{for} \quad n \geq 1. \tag{5.30}$$

Hence

$$c'_{n,\cdot} + D_n c_{n,\cdot} = (1 - \tfrac{1}{6}B_n)^{-1} g_{n,\cdot}(c_{n,\cdot}), \tag{5.31}$$

where $D_n = \delta_n^{-2}(1 - \frac{1}{6}B_n)^{-1}B_n$. It can be shown (Exercise 6) that

$$\|e^{-D_n t}\| \leq 2 \quad \text{for} \quad t \geq 0,\ n \geq 1.$$

This, (5.30) and (5.31) imply that

$$c_{n,\cdot}(t) = e^{-D_n t}c_{n,\cdot}(0) + \int_0^t e^{-D_n(t-s)}(1 - \tfrac{1}{6}B_n)^{-1} g_{n,\cdot}(c_{n,\cdot}(s))ds$$

$$\|c_{n,\cdot}(t)\|_\infty \leq 2\|c_{n,\cdot}(0)\|_\infty + 6\int_0^t \|g_{n,\cdot}(c_{n,\cdot}(s))\|_\infty ds.$$

Since $c_{n,\cdot}(0) = (1 - \frac{1}{6}B_n)^{-1}b_{n,\cdot}$ and $\|b_{n,\cdot}\|_\infty \leq \|u_0\|_\infty$ we have

$$\|c_{n,\cdot}(t)\|_\infty \leq 6\|u_0\|_\infty + 6\int_0^t \|g_{n,\cdot}(c_{n,\cdot}(s))\|_\infty ds. \tag{5.32}$$

Pick any $\gamma > 6\|u_0\|_\infty$ and let $G_\gamma = \max_{|x|\leq\gamma}|G(x)|$. (5.32) implies that

$$\|c_{n,\cdot}(t)\|_\infty \leq 6\|u_0\|_\infty + 6G_\gamma t$$

as long as $\|c_{n,\cdot}(s)\|_\infty \leq \gamma$ for $0 \leq s \leq t$. Hence, if τ is such that

$$6\|u_0\|_\infty + 6G_\gamma \tau = \gamma,$$

then bound (5.29) holds and the Galerkin approximations u_n converge on $[0,\tau]$ to the mild solution of the original problem.

5.7 Example: Galerkin Method for the Wave Equation

Here the results of Section 5.4 will be applied to

$$u''(t) + Au(t) = f(t,u(t),u'(t)), \quad u(0) = u_0, u'(0) = v_0, \tag{5.33}$$

where A is the operator associated with a sectorial form $\mathfrak{F}$.

Assume that $\mathcal{H}$ is a Hilbert space, with an inner product $(\cdot,\cdot)$ and norm $\|\cdot\|$, $[\cdot,\cdot]$ is an inner product on $\mathcal{V} \subset \mathcal{H}$ with its corresponding norm denoted by $|\cdot|$ and $\mathfrak{F}$ is a sectorial sesquilinear form on $\mathcal{V}$ such that hypotheses **H**1, **H**2 on **H**3 of the section on Sectorial Forms are satisfied. Here it is also assumed that

(**H4**) $\mathcal{V}_1, \mathcal{V}_2, \cdots$ are finite dimensional subspaces of $\mathcal{V}$ such that

$$\lim_{n\to\infty} \inf_{z\in\mathcal{V}_n} |y - z| = 0$$

for all y in a dense (in $|\cdot|$ norm) subset of $\mathcal{V}$.

These assumptions about $\mathcal{H}$, $\mathcal{V}$, $\mathcal{V}_n$, $\mathfrak{F}$ are the same as in Section 5.6, where the Galerkin method for parabolic equations was studied. In order for the wave equation to make sense we also assume, as in Section 4.4, that there exists $b < \infty$ such that

$$|\mathfrak{F}(x,y) - \overline{\mathfrak{F}(y,x)}| \le 2b|x|\|y\| \quad \text{for all} \quad x, y \in \mathcal{V}.$$

Assume, in addition, that $\tau \in (0,\infty)$, $f : [0,\tau] \times \mathcal{V} \times \mathcal{H} \to \mathcal{H}$ is continuous and such that for some $L < \infty$,

$$\|f(t,x_1,y_1) - f(t,x_2,y_2)\| \le L(|x_1 - x_2| + \|y_1 - y_2\|)$$

for all $0 \le t \le \tau$, $x_i \in \mathcal{V}$, $y_i \in \mathcal{H}$. In applications one can, as in Example 5.6.2, considerably weaken this condition on f.

Theorem 5.7.1 *Choose any $u_0 \in \mathcal{V}$, $v_0 \in \mathcal{H}$. Then for each $n \ge 1$ there exists a unique $u_n \in C^2([0,\tau],\mathcal{V}_n)$ such that for all $z \in \mathcal{V}_n$ we have*

$$(u_n''(t),z) + \mathfrak{F}(u_n(t),z) = (f(t,u_n(t),u_n'(t)),z) \quad \textit{for} \quad t \in [0,\tau]$$

$$[u_n(0),z] = [u_0,z], \quad (u_n'(0),z) = (v_0,z).$$

Moreover, there exists $u \in C([0,\tau],\mathcal{V}) \cap C^1([0,\tau],\mathcal{H})$ such that

$$\lim_{n\to\infty} \sup_{0\le t\le\tau} (|u_n(t) - u(t)| + \|u_n'(t) - u'(t)\|) = 0,$$

$$u(0) = u_0, \quad u'(0) = v_0,$$

$$(u(\cdot),z) \in C^2([0,\tau],\mathbb{C}) \quad \textit{for} \quad z \in \mathcal{V},$$

$$\frac{d^2}{dt^2}(u(t),z) + \mathfrak{F}(u(t),z) = (f(t,u(t),u'(t)),z) \quad \textit{for} \quad t \in [0,\tau],\ z \in \mathcal{V}.$$

PROOF Let $X_n = \mathcal{V}_n \times \mathcal{V}_n$. X_n is a Hilbert space with an inner product

$$(\{x,y\},\{z,w\}) = [x,z] + (y,w).$$

Let this also be the inner product on $X = \mathcal{V} \times \mathcal{H}$.

Let P_n' be the orthogonal (in $\mathcal{V}$) projection of $\mathcal{V}$ onto $\mathcal{V}_n$.

Let P_n'' be the orthogonal (in $\mathcal{H}$) projection of $\mathcal{H}$ onto $\mathcal{V}_n$.

Define $P_n \in \mathfrak{B}(X, X_n)$ by $P_n\{x, y\} = \{P_n' x, P_n'' y\}$ and note that P_n is the orthogonal (in X) projection of X onto X_n.

Let E_n be the identity map from X_n to X and note that assumption **A** of Section 5.4 is satisfied.

Define $S : \mathcal{D}(S) \to X$ by $\mathcal{D}(S) = \{\{x, y\} \mid x \in \mathcal{D}(A), y \in \mathcal{V}\}$ and

$$S\{x, y\} = \{-y, Ax\} \quad \text{for} \quad \{x, y\} \in \mathcal{D}(S).$$

Observe that assumptions **H1**, **H2**, **H3** remain valid, with the same constants, if we replace both $\mathcal{H}$ and $\mathcal{V}$ with X_n. Let the operator $A_n \in \mathfrak{B}(X_n)$ be the operator associated with $\mathfrak{F}$-restricted to X_n. Hence, A_n is defined by (5.26). Define

$$S_n\{x, y\} = \{-y, A_n x\} \quad \text{for} \quad \{x, y\} \in X_n.$$

Let us show that **B** and **C** of Section 5.4, with S, S_n in place of A, A_n, hold.

Theorem 4.4.1 implies that there exist $c, M \in (0, \infty)$, which depend only on b, a, M_1, M_2, M_3, such that

$$\|(S_n - \lambda)^{-k}\| \le M(|\lambda| - c)^{-k} \quad \text{for} \quad |\lambda| > c,\ \lambda \in \mathbb{R},\ n, k \ge 1.$$

Hence, the Hille-Yosida Theorem 4.3.5 implies that

$$\|e^{-S_n t}\| \le M e^{ct} \quad \text{for} \quad t \ge 0,\ n \ge 1.$$

This shows that assumption **B** of Section 5.4 is satisfied.

Choose $\lambda < -c$ and $\{z, w\} \in X$. Define

$$\{x_n, y_n\} = (S_n - \lambda)^{-1} P_n\{z, w\}, \quad \{x, y\} = (S - \lambda)^{-1}\{z, w\}.$$

In view of Theorem 4.4.1's proof note that $-\lambda^2 < a + M_3 M_1^{-2}$ and that $x = (A + \lambda^2)^{-1}(w - \lambda z)$, $y = -\lambda x - z$, and similarly for x_n, y_n. Hence

$$\begin{aligned} |x_n - x| &= |(A_n + \lambda^2)^{-1}(P_n'' w - \lambda P_n' z) - (A + \lambda^2)^{-1}(w - \lambda z)| \\ &\le |(A_n + \lambda^2)^{-1} P_n'' w - (A + \lambda^2)^{-1} w| \qquad (5.34) \\ &\quad + |\lambda||(A_n + \lambda^2)^{-1} P_n'' z - (A + \lambda^2)^{-1} z| \qquad (5.35) \\ &\quad + |\lambda||(A_n + \lambda^2)^{-1}(P_n' z - P_n'' z)|. \qquad (5.36) \end{aligned}$$

Theorem 2.8.11 implies that expressions (5.34) and (5.35) converge to 0 as $n \to \infty$. Lemma 2.8.10 implies that for some $c_1 < \infty$, depending only on $\lambda, a, M_1, M_2, M_3$, we have

$$\begin{aligned} |(A_n + \lambda^2)^{-1}(P_n' z - P_n'' z)| &\le c_1 \|P_n' z - P_n'' z\| \\ &\le c_1(\|P_n' z - z\| + \|z - P_n'' z\|) \\ &\le 2c_1 \|P_n' z - z\| \\ &\le 2c_1 M_1 |P_n' z - z| \end{aligned}$$

which, by **H**4, converges to 0 as $n \to \infty$. Thus, expression (5.36) and hence $|x_n - x|$ converge to 0 as $n \to \infty$. Since

$$\|y_n - y\| = \|\lambda(x - x_n) + z - P_n' z\| \leq M_1|\lambda||x - x_n| + M_1|z - P_n' z|,$$

we have that $\|y_n - y\|$ also converges to 0 as $n \to \infty$ and therefore assumption **C** of Section 5.4 is satisfied.

Define $H(t,x) = \{0, f(t,x)\}$ for $0 \leq t \leq \tau$, $x \in X$ and note that **D** of Section 5.4 holds. Let $\{u,v\}$ be the mild solution of (5.33) as given by Theorem 5.1.2, i.e. $\{u,v\} \in C([0,\tau], X)$ is such that

$$\{u(t), v(t)\} - \{u_0, v_0\} + S\int_0^t \{u,v\} = \int_0^t \{0,f\}.$$

Hence, $u \in C([0,\tau], \mathcal{V})$ and

$$u(t) - u_0 - \int_0^t v(s)ds = 0$$

$$v(t) - v_0 + A\int_0^t u(s)ds = \int_0^t f(s, u(s), v(s))ds.$$

Therefore $u(0) = u_0$, $u \in C^1([0,\tau], \mathcal{H})$, $u' = v$, $u'(0) = v_0$ and

$$u'(t) - v_0 + A\int_0^t u(s)ds = \int_0^t f(s, u(s), u'(s))ds \quad \text{for} \quad t \in [0,\tau].$$

If $z \in \mathcal{V}$, then, for $t \in [0,\tau]$,

$$(u'(t), z) - (v_0, z) + \mathfrak{F}\left(\int_0^t u(s)ds, z\right) = \int_0^t (f(s, u(s), u'(s)), z)ds.$$

Theorem 2.8.12 gives the representation $\mathfrak{F}(x,y) = (BGx, Gy)$ for $x, y \in \mathcal{V}$; hence

$$\mathfrak{F}\left(\int_0^t u(s)ds, z\right) = \left(BG\int_0^t u(s)ds, Gz\right).$$

Since $u \in C([0,\tau], \mathcal{V})$, we have that $Gu \in C([0,\tau], \mathcal{H})$ by Theorem 2.8.12, and Theorem 4.2.10 implies $G\int u = \int Gu$. Hence

$$\mathfrak{F}\left(\int_0^t u(s)ds, z\right) = \left(\int_0^t BGu(s)ds, Gz\right) = \int_0^t \mathfrak{F}(u(s), z)ds$$

and therefore

$$(u'(t), z) - (v_0, z) + \int_0^t \mathfrak{F}(u(s), z)ds = \int_0^t (f(s, u(s), u'(s)), z)ds,$$

implying that u has all of the claimed properties.

Let $\{u_n, v_n\} \in C^1([0,\tau], X_n)$ satisfy

$$\{u_n, v_n\}' + S_n\{u_n, v_n\} = P_n H(t, \{u_n, v_n\}), \quad \{u_n(0), v_n(0)\} = P_n\{u_0, v_0\},$$

i.e. $v_n = u_n'$ and

$$u_n'' + A_n u_n = P_n'' f(t, u_n, u_n'), \quad u_n(0) = P_n' u_0, \quad u_n'(0) = P_n'' v_0,$$

which is, in view of (5.26), equivalent to the formulation given in the Theorem. Theorem 5.4.7 implies that $\{u_n, v_n\}$ converge to $\{u, v\}$ and this completes the proof. □

5.8 Friedrichs Extension and Galerkin Approximations

Theorem 2.12.6 showed that an elementary algebraic condition implies convergence of Galerkin approximations to the solution of an elliptic problem, where the elliptic operator is the Friedrichs extension determined by the basis functions. The same holds for parabolic and wave type equations. In the case of parabolic equations we have:

Theorem 5.8.1 *Suppose that*

(a) $\mathcal{H}$ *is a complex Hilbert space*

(b) $\phi_n \in \mathcal{H}$ *for* $n \geq 1$, $\mathcal{V}_n = \text{span}\{\phi_1, \ldots, \phi_n\}$ *and* $\cup_{n=1}^{\infty} \mathcal{V}_n$ *is dense in* $\mathcal{H}$

(c) S *is a linear operator in* $\mathcal{H}$ *with domain* $\mathcal{D}(S) = \cup_{n=1}^{\infty} \mathcal{V}_n$ *and such that there exist* $r \in \mathbb{R}$ *and* $M \in (0, \infty)$ *such that*

$$|\text{Im}\,(Sx, x)| \leq M \text{Re}\,(Sx - rx, x) \quad \textit{for all} \quad x \in \mathcal{D}(S)$$

(d) $\tau \in (0, \infty)$, $H : [0,\tau] \times \mathcal{H} \to \mathcal{H}$ *is continuous and such that for some* $L < \infty$

$$\|H(t,x) - H(t,y)\| \leq L\|x - y\| \quad \textit{for} \quad 0 \leq t \leq \tau,\ x, y \in \mathcal{H}$$

(e) $x \in \mathcal{H}$.

Then for every $n \geq 1$ *there exists a unique* $u_n \in C^1([0,\tau], \mathcal{V}_n)$ *such that for all* $z \in \mathcal{V}_n$ *we have* $(u_n(0), z) = (x, z)$ *and*

$$(u_n'(t), z) + (Su_n(t), z) = (H(t, u_n(t)), z) \quad \textit{for} \quad t \in [0,\tau].$$

Moreover,

$$\lim_{n\to\infty} \sup_{0\leq t\leq\tau} \|u_n(t) - u(t)\| = 0,$$

where u *is the mild solution of* $u'(t) + Au(t) = H(t, u(t))$, $u(0) = x$ *and* A *is the Friedrichs extension of* S.

PROOF Theorem 2.12.1 implies that all assumptions of Section 5.6 are satisfied. Hence, the assertions of the Theorem follow from Theorem 5.6.1. □

EXAMPLE 5.8.2 To see how easy it is to apply the above Theorem to even abstruse problems, consider the following PDE:

$$u_t = (p_0 u_x)_x + p_1 u_x + p_2 u + p_3 \sin(\mathrm{Re}\, u) \quad \text{for} \quad x \in \mathbb{R},\ t \geq 0$$

$$u(x,0) = g(x) \quad \text{for} \quad x \in \mathbb{R},$$

where

(1) p_0, p_1, p_2, p_3 are complex valued functions in $L^\infty(\mathbb{R})$

(2) $p_0 \in C^1(\mathbb{R})$ and $p_0' \in L^\infty(\mathbb{R})$

(3) there exists $c \in (0,\infty)$ such that

$$|p_1(x)|^2 + |\mathrm{Im}\, p_0(x)| \leq c^2 \mathrm{Re}\, p_0(x) \quad \text{for} \quad x \in \mathbb{R}$$

(4) $g \in L^2(\mathbb{R})$ is complex valued.

Note that p_0 can vanish on an interval.

Let $\mathcal{H}$ be the complex $L^2(\mathbb{R})$ and let

$$\phi_n(x) = x^{n-1} e^{-x^2/2} \quad \text{for} \quad n \geq 1,\ x \in \mathbb{R}.$$

Define $\mathcal{V}_n = \mathrm{span}\{\phi_1, \ldots, \phi_n\}$ and $\mathcal{D}(S) = \cup_{n=1}^\infty \mathcal{V}_n$. Note that $v \in \mathcal{D}(S)$ iff $v(x) = P(x)e^{-x^2/2}$ where P is a polynomial. Theorem 2.9.8 implies that $\mathcal{D}(S)$ is dense in $\mathcal{H}$. Define

$$Sv = -(p_0 v')' - p_1 v' - p_2 v \quad \text{for} \quad v \in \mathcal{D}(S).$$

If $v \in \mathcal{D}(S)$, then

$$(Sv, v) = \int_{-\infty}^{\infty} p_0 |v'|^2 - p_1 v' \overline{v} - p_2 |v|^2.$$

Hence

$$\mathrm{Re}\,(Sv, v) \geq w^2 - cw\|v\|_2 - \|p_2\|_\infty \|v\|_2^2$$

$$|\mathrm{Im}\,(Sv, v)| \leq c^2 w^2 + cw\|v\|_2 + \|p_2\|_\infty \|v\|_2^2,$$

where $w^2 = \int |v'|^2 \mathrm{Re}\, p_0$, $w > 0$. Thus, for any $M > c^2$ one can find r so that *(c)* is true and hence Theorem 5.8.1 applies.

5.9 Exercises

1. Show that if X is a Hilbert space and R_r is the retraction map given in Theorem 5.1.4, then $\|R_r(x) - R_r(y)\| \leq \|x - y\|$ for all $x, y \in X$, $r > 0$.

2. Give the sharp upper bound for the blow up time of the solution of

$$u_t(x,t) = \Delta_\ell u(x,t) + e^{u(x,t)} \quad \text{for} \quad x \in \mathbb{R}^n,\ t \geq 0$$

$$u(x,0) = u_0(x) \quad \text{for} \quad x \in \mathbb{R}^n,$$

where $u_0 \in \mathbb{C}(\mathbb{R}^n, \mathbb{R})$ is such that there exists $\lim_{|x|\to\infty} u_0(x) \in \mathbb{R}$.

3. Give the sharp upper bound for the blow up time of the solution of

$$u_t(x,t) = \Delta_\ell u(x,t) + u(x,t)^2 \quad \text{for} \quad x \in \mathbb{R}^n,\ t \geq 0$$

$$u(x,0) = u_0(x) \quad \text{for} \quad x \in \mathbb{R}^n,$$

where $u_0 \in \mathbb{C}(\mathbb{R}^n, \mathbb{R})$ is such that there exists $\lim_{|x|\to\infty} u_0(x) > 0$.

4. Prove that $e^{\Delta_\ell t}$ has the same integral representation as Q in Example 4.1.5.

5. Suppose that

$$4c_i + c_{i+1} + c_{i-1} = 6d_i \quad \text{for} \quad 1 \leq i \leq n,\ c_0 = c_{n+1} = 0.$$

Show that $c_i = \sum_{j=1}^n G_n(i,j)d_j$ for $1 \leq i \leq n$, where

$$G_n(i,j) = G_n(j,i) = 6\frac{(-1)^{i+j} f(j) f(n+1-i)}{f(1) f(n+1)} \quad \text{for} \quad 1 \leq j \leq i \leq n,$$

$f(i) = \lambda^i - \lambda^{-i}$ and $\lambda = 2 + \sqrt{3}$. Show that this implies

$$\sum_{j=1}^n |G_n(i,j)| = 3 - 3\frac{\lambda^{i-\frac{n+1}{2}} + \lambda^{-i+\frac{n+1}{2}}}{\lambda^{\frac{n+1}{2}} + \lambda^{-\frac{n+1}{2}}} \quad \text{for} \quad 1 \leq i \leq n.$$

Hence

$$\max_{1\leq i\leq n} |c_i| \leq 3 \max_{1\leq i\leq n} |d_i|.$$

6. Show that the solution of the following system of ODEs,

$$4c_i' + c_{i+1}' + c_{i-1}' + 2c_i - c_{i+1} - c_{i-1} = 0 \quad \text{for} \quad 1 \leq i \leq n,\ c_0 = c_{n+1} = 0,$$

is

$$c_i(t) = \sum_{j=1}^n \frac{2}{n+1} \left(\sum_{k=1}^n e^{-\omega_k t} \sin(ik\alpha) \sin(jk\alpha) \right) c_j(0) \quad \text{for} \quad 1 \leq i \leq n,\ t \geq 0,$$

where

$$\omega_k = \frac{1-\cos k\alpha}{2+\cos k\alpha}, \quad \alpha = \frac{\pi}{n+1}.$$

Show that

$$\sup_{1\le i\le n, t\ge 0} |c_i(t)| \le 1.09677269 \max_{1\le i\le n} |c_i(0)|.$$

7. Derive and prove convergence of the finite element approximation for

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{1}{x}\frac{\partial u}{\partial x} + f(x,t)$$

$$u(1,t) = \frac{\partial u}{\partial x}(0,t) = 0$$

$$u(x,0) = u_0(x),$$

where $0 \le x \le 1$, $t \ge 0$ and f, u_0 are continuous. Let $\mathcal{H}$ consist of all measurable functions v for which

$$\int_0^1 x|v(x)|^2 dx < \infty$$

and let the approximations be linear on intervals $[\frac{i-1}{n}, \frac{i}{n}]$, $1 \le i \le n$, $n \ge 2$.

8. Using basis functions $\cos(n-1)\pi x$, $n \ge 1$, prove convergence of Galerkin approximations for

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{1}{x}\frac{\partial u}{\partial x} + f(x,t)$$

$$\frac{\partial u}{\partial x}(0,t) = \frac{\partial u}{\partial x}(1,t) = 0$$

$$u(x,0) = u_0(x),$$

where $0 \le x \le 1$, $t \ge 0$ and f, u_0 are continuous.

9. Using basis functions $x^n(1-x)$, $n \ge 1$, prove convergence of Galerkin approximations for

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(x,t)$$

$$u(0,t) = u(1,t) = 0$$

$$u(x,0) = u_0(x),$$

where $0 \le x \le 1$, $t \ge 0$ and f, u_0 are continuous.

10. State and prove the analog of Theorem 5.8.1 for the wave equation.

Chapter 6

Semilinear Parabolic Equations

In this chapter we will extend the results of the preceding chapter to allow for less restrictive nonlinearities. While the exact definition of semilinear parabolic equations is left for Section 6.4, its typical representative is

$$u_t = \Delta u + f(u, \nabla u).$$

The dominant term in the equation is the Laplacian, so, we need to control the nonlinearity in terms of the Laplacian which leads one to study its fractional powers. An equivalent approach is to define the nonlinearity in an interpolation space - a space between the domain of the Laplacian and the basic space.

6.1 Fractional Powers of Operators

Definition 6.1.1 Suppose that $-A$ is the generator of a strongly continuous semigroup $\{Q(t)\}_{t\geq 0}$ on a Banach space X and that

$$\|Q(t)\| \leq Me^{-at} \quad \text{for} \quad t \geq 0$$

where $a > 0$ and $M < \infty$. For such A define $A^\alpha \in \mathfrak{B}(X)$, for $\alpha \in (-\infty, 0)$, by

$$A^\alpha x = \frac{1}{\Gamma(-\alpha)} \int_0^\infty t^{-\alpha-1} Q(t)x\,dt \quad \text{for} \quad x \in X.$$

Let $\mathfrak{G}(X)$ denote the set of all such operators A in X.

Observe that when α is a negative integer Theorem 4.3.2 implies that the above definition of A^α agrees with its usual definition.

EXAMPLE 6.1.2 Let $\{\varphi_1, \varphi_2, \ldots\}$ be a complete orthonormal set in a Hilbert space $\mathcal{H}$, let $\lambda_1, \lambda_2, \ldots$ be complex numbers such that $\inf_k \operatorname{Re} \lambda_k > 0$ and let a linear operator A in $\mathcal{H}$ be defined by

$$Ax = \sum_{k=1}^{\infty} \lambda_k (x, \varphi_k)\varphi_k \quad \text{for} \quad x \in \mathcal{D}(A) = \{x \in \mathcal{H} \mid \sum_{k=1}^{\infty} |\lambda_k (x, \varphi_k)|^2 < \infty\}.$$

In view of Example 4.1.6 we have that $A \in \mathfrak{G}(X)$ and

$$(A^\alpha x, \varphi_k) = \frac{1}{\Gamma(-\alpha)} \int_0^\infty t^{-\alpha-1} e^{-\lambda_k t}(x, \varphi_k) dt = \lambda_k^\alpha (x, \varphi_k) \quad \text{for} \quad k \geq 1.$$

Hence,

$$A^\alpha x = \sum_{k=1}^\infty \lambda_k^\alpha (x, \varphi_k)\varphi_k \quad \text{for} \quad \alpha < 0,\ x \in \mathcal{H}.$$

The formula in the following Theorem can actually be used to define fractional powers of operators for operators in a bit larger class than $\mathfrak{G}(X)$.

Theorem 6.1.3 *If $A \in \mathfrak{G}(X)$, then*

$$A^{-\alpha} = \frac{\sin \pi\alpha}{\pi} \int_0^\infty \lambda^{-\alpha}(A + \lambda)^{-1} d\lambda \quad \textit{for} \quad \alpha \in (0, 1).$$

PROOF Theorem 4.3.2 implies that for all $x \in X$

$$\begin{aligned}
(A + \lambda)^{-1} x &= \int_0^\infty e^{-\lambda t} Q(t) x \, dt \quad (\lambda \geq 0) \\
\int_0^\infty \lambda^{-\alpha}(A + \lambda)^{-1} x \, d\lambda &= \int_0^\infty \int_0^\infty \lambda^{-\alpha} e^{-\lambda t} Q(t) x \, dt \, d\lambda \\
&= \int_0^\infty t^{\alpha-1} \Gamma(1 - \alpha) Q(t) x \, dt \\
&= \Gamma(\alpha)\Gamma(1 - \alpha) A^{-\alpha} x.
\end{aligned}$$

□

Lemma 6.1.4 *If $A \in \mathfrak{G}(X)$, then $A^\alpha A^\beta = A^{\alpha+\beta}$ for $\alpha, \beta < 0$.*

PROOF For all $x \in X$ we have

$$\begin{aligned}
A^\alpha A^\beta x &= \frac{1}{\Gamma(-\alpha)} \int_0^\infty t^{-\alpha-1} Q(t) A^\beta x \\
&= \frac{1}{\Gamma(-\alpha)\Gamma(-\beta)} \int_0^\infty \int_0^\infty t^{-\alpha-1} Q(t) s^{-\beta-1} Q(s) x \, ds \, dt \\
&= \frac{1}{\Gamma(-\alpha)\Gamma(-\beta)} \int_0^\infty \int_t^\infty t^{-\alpha-1} (r - t)^{-\beta-1} Q(r) x \, dr \, dt \\
&= \frac{1}{\Gamma(-\alpha)\Gamma(-\beta)} \int_0^\infty \int_0^r t^{-\alpha-1} (r - t)^{-\beta-1} Q(r) x \, dt \, dr \\
&= \frac{1}{\Gamma(-\alpha - \beta)} \int_0^\infty r^{-\alpha-\beta-1} Q(r) x \, dr = A^{\alpha+\beta} x.
\end{aligned}$$

Lemma 6.1.5 *If $A \in \mathfrak{G}(X)$, then, for all $\alpha < 0$, A^α is one-to-one and the range of A^α is dense in X.*

PROOF Pick $-n < \alpha$. Since $A^{-n} = A^\alpha A^{-n-\alpha}$ we have that $\mathcal{R}(A^{-n}) \subset \mathcal{R}(A^\alpha)$. Theorem 4.3.1 implies that $\mathcal{R}(A^{-n}) = \mathcal{D}(A^n)$ is dense in X and so is $\mathcal{R}(A^\alpha)$.

If $A^\alpha x = 0$, then $A^{-n}x = A^{-n-\alpha}A^\alpha x = 0$, hence $A^{-n+1}x = AA^{-n}x = 0, \ldots,$ and therefore $x = 0$. □

This makes possible the following:

Definition 6.1.6 For $A \in \mathfrak{G}(X)$ and $\alpha \in (0, \infty)$ let A^α be the inverse map of $A^{-\alpha}$, i.e.

$$\mathcal{D}(A^\alpha) = \mathcal{R}(A^{-\alpha}), \quad A^\alpha A^{-\alpha}x = x \quad \text{for} \quad x \in X.$$

Define A^0 to be the identity map on X.

EXAMPLE 6.1.7 When A is defined in a Hilbert space $\mathcal{H}$, as in Example 6.1.2, then it is easy to see that for every $\alpha > 0$

$$A^\alpha x = \sum_{k=1}^{\infty} \lambda_k^\alpha (x, \varphi_k)\varphi_k \quad \text{for} \quad x \in \mathcal{D}(A) \equiv \{x \in \mathcal{H} \mid \sum_{k=1}^{\infty} |\lambda_k^\alpha (x, \varphi_k)|^2 < \infty\}.$$

Theorem 6.1.8 *If $A \in \mathfrak{G}(X)$, then*

(1) A^α is closed, densely defined, $0 \in \rho(A^\alpha)$, $(A^\alpha)^{-1} = A^{-\alpha}$ for all $\alpha > 0$

(2) $A^\alpha Q(t)x = Q(t)A^\alpha x$ for $\alpha \in \mathbb{R}$, $t \geq 0$, $x \in \mathcal{D}(A^\alpha)$

(3) $\mathcal{D}(A^\alpha) \subset \mathcal{D}(A^\beta)$ if $\alpha > \beta$

(4) $A^{\alpha+\beta}x = A^\alpha A^\beta x$ for $\alpha, \beta \in \mathbb{R}$ and $x \in \mathcal{D}(A^\beta) \cap \mathcal{D}(A^{\alpha+\beta})$

(5) $\|A^\alpha x\| \leq 2(M+1)\|Ax\|^\alpha \|x\|^{1-\alpha}$ for all $\alpha \in [0,1]$ and $x \in \mathcal{D}(A)$

(6) for each $\lambda \in \mathbb{R}$ there exists $c < \infty$ such that

$$\|x - e^{\lambda t}Q(t)x\| \leq c(1 + e^{\lambda t})\, t^\alpha \|A^\alpha x\| \quad \text{for} \quad \alpha \in [0,1],\ x \in \mathcal{D}(A^\alpha),\ t > 0.$$

PROOF Definition 6.1.6 and Lemma 6.1.5 imply (1).

Definition 6.1.1 and Theorem 4.2.8 imply (2) when $\alpha < 0$ and for $\alpha > 0$ it follows from the fact that $(A^\alpha)^{-1} = A^{-\alpha}$.

If $\alpha > \beta > 0$, then, by Lemma 6.1.4, $A^{-\alpha} = A^{-\beta}A^{\beta-\alpha}$ which implies (3).

When $\alpha \leq 0$, $\beta \leq 0$, then Lemma 6.1.4 implies (4).

When $\alpha \le 0$, $\beta > 0$, $\alpha + \beta > 0$, then $A^{-\alpha-\beta}A^{\alpha} = A^{-\beta}$ applied to $A^{\beta}x$ implies $A^{-\alpha-\beta}A^{\alpha}A^{\beta}x = x$ which implies (4).

When $\alpha \le 0$, $\beta > 0$, $\alpha + \beta \le 0$, then $A^{\alpha+\beta}A^{-\beta}(A^{\beta}x) = A^{\alpha}(A^{\beta}x)$ implies (4).

When $\alpha > 0$, $\beta \le 0$, $\alpha + \beta > 0$, then $A^{\beta}A^{-\alpha-\beta} = A^{-\alpha}$ applied to $A^{\alpha+\beta}x$ implies $A^{\beta}x = A^{-\alpha}A^{\alpha+\beta}x$ which implies (4).

When $\alpha > 0$, $\beta \le 0$, $\alpha + \beta \le 0$, then $A^{-\alpha}A^{\alpha+\beta}x = A^{\beta}x$ for $x \in X$ which implies (4).

When $\alpha > 0$, $\beta > 0$, then $A^{-\beta}A^{-\alpha} = A^{-\alpha-\beta}$ applied to $A^{\alpha+\beta}x$ implies $A^{-\beta}A^{-\alpha}A^{\alpha+\beta}x = x$ which implies (4).

To prove (5) assume that $x \in \mathcal{D}(A)$, $\alpha \in (0,1)$. Theorem 6.1.3 implies

$$\|A^{\alpha}x\| = \|A^{\alpha-1}Ax\| \le \frac{\sin\pi(1-\alpha)}{\pi}\int_0^{\infty}\lambda^{\alpha-1}\|(A+\lambda)^{-1}Ax\|d\lambda.$$

Since $\|(A+\lambda)^{-1}\| \le M/\lambda$, $\|A(A+\lambda)^{-1}\| \le M+1$ for $\lambda > 0$ we have

$$\begin{aligned}\|A^{\alpha}x\|\frac{\pi}{\sin\pi(1-\alpha)} &\le \int_0^{\varepsilon}\lambda^{\alpha-1}(M+1)\|x\|d\lambda + \int_{\varepsilon}^{\infty}\lambda^{\alpha-2}M\|Ax\|d\lambda\\ &\le \varepsilon^{\alpha}\alpha^{-1}(M+1)\|x\| + \varepsilon^{\alpha-1}(1-\alpha)^{-1}M\|Ax\|\\ \|A^{\alpha}x\| &\le \varepsilon^{\alpha}(M+1)\|x\| + \varepsilon^{\alpha-1}M\|Ax\|\end{aligned}$$

for all $\varepsilon > 0$. Minimizing this expression, see (3.15), gives the bound in (5).

To prove (6) let $z = x - e^{\lambda t}Q(t)x$. (2), (4) and (5) imply

$$\|z\| = \|A^{1-\alpha}A^{\alpha-1}z\| \le 2(M+1)\|A^{\alpha}z\|^{1-\alpha}\|A^{\alpha-1}z\|^{\alpha}.$$

Hence, using $\|A^{\alpha}z\| = \|A^{\alpha}x - e^{\lambda t}Q(t)A^{\alpha}x\| \le M(1+e^{\lambda t})\|A^{\alpha}x\|$ and

$$\begin{aligned}A^{\alpha-1}z &= (A-\lambda)A^{-1}\int_0^t e^{\lambda s}Q(s)A^{\alpha}x\,ds\\ \|A^{\alpha-1}z\| &\le c_1\int_0^t e^{\lambda s}\|A^{\alpha}x\|ds \le c_1(1+e^{\lambda t})t\|A^{\alpha}x\|\end{aligned}$$

implies (6). □

The following Theorem is the key tool for obtaining bounds in terms of fractional powers of operators - which are used to control the nonlinear term in semilinear parabolic equations.

Theorem 6.1.9 *If $A \in \mathfrak{G}(X)$ and B is a closed linear operator from X into another Banach space Y such that $\mathcal{D}(B) \supset \mathcal{D}(A)$ and that for some $\beta \in [0,1)$, $c < \infty$ we have*

$$\|Bx\|_Y \le c\|Ax\|^{\beta}\|x\|^{1-\beta} \quad \text{for} \quad x \in \mathcal{D}(A), \tag{6.1}$$

then $\mathcal{D}(B) \supset \mathcal{D}(A^{\alpha})$ and $BA^{-\alpha} \in \mathfrak{B}(X,Y)$ for $\alpha > \beta$.

PROOF Choose $\gamma \in (\beta, 1)$, $\gamma \le \alpha$ and $x \in \mathcal{D}(A^{\gamma})$. Theorem 6.1.3 implies

$$x = A^{-\gamma}A^{\gamma}x = \int_0^{\infty} f(\lambda)d\lambda \quad \text{where} \quad f(\lambda) = \frac{\sin \pi\gamma}{\pi}\lambda^{-\gamma}(A+\lambda)^{-1}A^{\gamma}x.$$

(6.1) implies that $Bf \in C((0,\infty), Y)$ and

$$\begin{aligned} \|Bf(\lambda)\|_Y &\le c\|\lambda^{-\gamma}A(A+\lambda)^{-1}A^{\gamma}x\|^{\beta}\|\lambda^{-\gamma}(A+\lambda)^{-1}A^{\gamma}x\|^{1-\beta} \\ &\le c\lambda^{-\gamma}\|A(A+\lambda)^{-1}\|^{\beta}\|(A+\lambda)^{-1}\|^{1-\beta}\|A^{\gamma}x\|. \end{aligned}$$

Since $\|(A+\lambda)^{-1}\| \le M/(\lambda+a)$, $\|A(A+\lambda)^{-1}\| \le M+1$ for $\lambda > 0$, we have

$$\begin{aligned} \|Bf(\lambda)\|_Y &\le c(M+1)\lambda^{-\gamma}(\lambda+a)^{\beta-1}\|A^{\gamma}x\| \\ \int_0^{\infty} \|Bf(\lambda)\|_Y d\lambda &\le c_1\|A^{\gamma}x\|. \end{aligned}$$

Hence Bf is Bochner integrable on $(0,\infty)$ and therefore Theorem 4.2.10 implies that $x = \int f \in \mathcal{D}(B)$ and $\|Bx\|_Y \le c_1\|A^{\gamma}x\|$. Clearly, $\mathcal{D}(B) \supset \mathcal{D}(A^{\alpha})$ and

$$\|BA^{-\alpha}\| = \|BA^{-\gamma}A^{\gamma-\alpha}\| \le c_1\|A^{\gamma-\alpha}\|.$$

□

EXAMPLE 6.1.10 Suppose $p \in (n/2, \infty)$, $p \ge 1$ and let I be the identity map from $L^p(\mathbb{R}^n)$ into $C_B(\mathbb{R}^n)$. Note that I is closed. Using (3.14) and the definition of $\overline{\Delta}$ in Example 4.3.10 implies that

$$\|Iu\|_{\infty} \le c\|\overline{\Delta}u\|_p^{\frac{n}{2p}}\|u\|_p^{1-\frac{n}{2p}} \quad \text{for} \quad u \in \mathcal{D}(\overline{\Delta}).$$

Choose $a > 0$. Note that $\|e^{\overline{\Delta}t}\| \le 1$ for $t \ge 0$, $\|\overline{\Delta}(a-\overline{\Delta})^{-1}\| \le 2$ hence

$$\|Iu\|_{\infty} \le 2c\|(a-\overline{\Delta})u\|_p^{\frac{n}{2p}}\|u\|_p^{1-\frac{n}{2p}} \quad \text{for} \quad u \in \mathcal{D}(\overline{\Delta}).$$

Theorem 6.1.9 implies that, for every $\alpha > n/(2p)$, $\mathcal{D}((a-\overline{\Delta})^{\alpha}) \subset C_B(\mathbb{R}^n)$ and there exists $c_1 < \infty$ such that

$$\|Iu\|_{\infty} \le c_1\|(a-\overline{\Delta})^{\alpha}u\|_p \quad \text{for} \quad u \in \mathcal{D}((a-\overline{\Delta})^{\alpha}).$$

EXAMPLE 6.1.11 Suppose $1 \le q \le p \le \infty$, $\beta = \frac{1}{2} + \frac{n}{2}(\frac{1}{q} - \frac{1}{p}) < 1$ and $1 \le k \le n$. Let $\mathcal{D}(B)$ consists of all $u \in L^q(\mathbb{R}^n)$ such that $\mathsf{D}_k u$ exists and $\mathsf{D}_k u \in L^p(\mathbb{R}^n)$. Define $B : \mathcal{D}(B) \subset L^q(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ by $Bu = \mathsf{D}_k u$ for $u \in \mathcal{D}(B)$. Theorem 3.4.3 implies that B is closed. Using (3.16) and the definition of $\overline{\Delta}$ in Example 4.3.10 implies that

$$\|Bu\|_p \le c\|\overline{\Delta}u\|_q^{\beta}\|u\|_q^{1-\beta} \quad \text{for} \quad u \in \mathcal{D}(\overline{\Delta}).$$

Theorem 6.1.9 implies that, for every $a > 0, \alpha > \beta$, $\mathcal{D}((a - \overline{\Delta})^{\alpha}) \subset \mathcal{D}(B)$ and there exists $c_1 < \infty$ such that

$$\|Bu\|_p \le c_1\|(a - \overline{\Delta})^{\alpha}u\|_q \quad \text{for} \quad u \in \mathcal{D}((a - \overline{\Delta})^{\alpha}).$$

Theorem 6.1.12 *Suppose $A, B \in \mathfrak{G}(X)$ are such that $\mathcal{D}(A) = \mathcal{D}(B)$ and*

$$\|(A - B)x\| \le c\|Ax\|^{\alpha}\|x\|^{1-\alpha} \quad \textit{for} \quad x \in \mathcal{D}(A) \tag{6.2}$$

for some $\alpha \in [0, 1)$, $c < \infty$. Then $\mathcal{D}(A^{\beta}) = \mathcal{D}(B^{\beta})$ for all $\beta \in [0, 1]$.

PROOF Assume $\beta \in (0, 1)$, $x \in X$. Theorem 6.1.3 implies that

$$B^{-\beta}x - A^{-\beta}x = \frac{\sin \pi\beta}{\pi} \int_0^{\infty} f(\lambda)d\lambda \tag{6.3}$$

where $f(\lambda) = \lambda^{-\beta}(B+\lambda)^{-1}(A-B)(A+\lambda)^{-1}x$. Note that for some $a > 0, M < \infty$

$$\|(A + \lambda)^{-1}\| \le M/(\lambda + a), \quad \|(B + \lambda)^{-1}\| \le M/(\lambda + a) \quad \text{for} \quad \lambda > 0.$$

Hence (5) of Theorem 6.1.8 implies

$$\|B^{\beta}(B + \lambda)^{-1}\| \le c_1(\lambda + a)^{\beta-1} \quad \text{for} \quad \lambda > 0$$

and (6.2) implies

$$\|(A - B)(A + \lambda)^{-1}\| \le c_1(\lambda + a)^{\alpha-1} \quad \text{for} \quad \lambda > 0$$

for some $c_1 < \infty$. Thus $\|B^{\beta}f(\lambda)\| \le c_1^2\lambda^{-\beta}(\lambda + a)^{\alpha+\beta-2}\|x\|$. Hence $B^{\beta}f$ is Bochner integrable on $(0, \infty)$ and Theorem 4.2.10 implies that $\int f \in \mathcal{D}(B^{\beta})$. (6.3) implies $A^{-\beta}x \in \mathcal{D}(B^{\beta})$ and therefore $\mathcal{D}(A^{\beta}) \subset \mathcal{D}(B^{\beta})$.

Since we also have (Chapter 1, Exercise 15)

$$\|(A - B)x\| \le c\|AB^{-1}\|^{\alpha}\|Bx\|^{\alpha}\|x\|^{1-\alpha} \quad \text{for} \quad x \in \mathcal{D}(B)$$

the above conclusion implies $\mathcal{D}(B^{\beta}) \subset \mathcal{D}(A^{\beta})$. □

Corollary 6.1.13 *If $A \in \mathfrak{G}(X)$, $b > 0$ and $\alpha \in [0, 1]$, then $\mathcal{D}(A^{\alpha}) = \mathcal{D}((A + b)^{\alpha})$.*

One can actually show that $(A+b)^\alpha x$ has a limit as $b \to -a$ (see Definition 6.1.1) for every $x \in \mathcal{D}(A^\alpha)$, $\alpha \in (0,1)$. This enables one to extend the definition of A^α, $\alpha > 0$ for the case when $-A$ is the generator of a bounded strongly continuous semigroup. A Theorem similar to Theorem 6.1.9 can be proved in this case.

Corollary 6.1.14 *Under the assumptions of Theorem 2.8.6 we have*

$$\mathcal{D}((A-a)^\beta) = \mathcal{D}((A^*-a)^\beta) = \mathcal{D}((A_r-a)^\beta) \quad \textit{for} \quad \beta \in [0,1].$$

PROOF Using the notation and assumptions as in the Sectorial Forms section, we have for $x \in \mathcal{D}(A)$, $y \in \mathcal{V}$

$$\begin{aligned}
((A-A_r)x,y) &= (\mathfrak{F}(x,y) - \overline{\mathfrak{F}(y,x)})/2 \\
|((A-A_r)x,y)| &\leq b|x\|\|y\| \\
\|(A-A_r)x\| &\leq b|x| \\
\|(A-A_r)x\|^2 &\leq b^2(\mathfrak{F}_r(x,x) - a\|x\|^2)/M_3 \\
&= b^2((A_r-a)x,x)/M_3 \\
\|(A^*-A_r)x\| = \|(A-A_r)x\| &\leq bM_3^{-1/2}\|(A_r-a)x\|^{1/2}\|x\|^{1/2}.
\end{aligned}$$

Hence Theorem 6.1.12 implies the conclusion. □

Theorem 6.1.15 *Suppose A is a sectorial operator and that $a \leq b < \operatorname{Re}\lambda$ for all $\lambda \in \sigma(A)$. Then, for every $n \geq 1$ there exists $c < \infty$ such that*

$$\|(A-a)^\alpha e^{-At}\| \leq c\,t^{-\alpha}e^{-bt} \quad \textit{for all} \quad t > 0,\ \alpha \in [0,n].$$

PROOF Lemma 4.5.5 and Theorem 4.5.6 imply that $A \in \mathfrak{A}(b+\delta, M, \theta, X)$ for some $\delta > 0$, M and θ. Theorem 4.5.10 implies $\|e^{-At}\| \leq c_1 e^{-(b+\delta)t}$ and Theorem 4.5.17 implies

$$\|(A-a)e^{-At}\| \leq c_2(1+t^{-1})e^{-(b+\delta)t} \leq c_3 t^{-1}e^{-bt} \quad \text{for} \quad t > 0.$$

Note that $A - a \in \mathfrak{G}(X)$. Theorem 4.5.14 and (5) of Theorem 6.1.8 imply

$$\|(A-a)^\alpha e^{-At}\| \leq c_4 t^{-\alpha}e^{-bt} \quad \text{for all} \quad t > 0,\ \alpha \in [0,1],$$

where $c_4 = 2(c_1+1)c_3^\alpha c_1^{1-\alpha}$. If $0 \leq \alpha \leq n$, then for all $t > 0$

$$\begin{aligned}
(A-a)^\alpha e^{-At} &= \left((A-a)^{\frac{\alpha}{n}} e^{-A\frac{t}{n}}\right)^n \\
\|(A-a)^\alpha e^{-At}\| &\leq \left(c_4(t/n)^{-\frac{\alpha}{n}} e^{-b\frac{t}{n}}\right)^n \\
&= c_4^n n^\alpha t^{-\alpha} e^{-bt}.
\end{aligned}$$

□

To simplify notation the following definition will be used.

Definition 6.1.16 Let A be a sectorial operator in a Banach space X and $a < \operatorname{Re}\lambda$ for all $\lambda \in \sigma(A)$. For $\alpha \geq 0$, define $X^\alpha = \mathcal{D}((A-a)^\alpha)$ and

$$\|x\|_\alpha = \|(A-a)^\alpha x\| \quad \text{for} \quad x \in X^\alpha.$$

(1) of Theorem 6.1.8 implies that $0 \in \sigma((A-a)^\alpha)$, hence, X^α is a Banach space with the norm $\|\cdot\|_\alpha$. Corollary 6.1.13 implies that X^α does not depend on the particular choice of a. Different choices of a give equivalent norms $\|\cdot\|_\alpha$ on X^α (Chapter 1, Exercise 15).

Lemma 6.1.17 *If A is a sectorial operator and $n \geq 1$, then, there exists $c < \infty$ such that*

$$\|e^{-A(t+h)}x - e^{-At}x\|_\alpha \leq c(e^{-a(t+h)} + e^{-at})\, h^\delta t^{\beta-\alpha-\delta}\|x\|_\beta$$

whenever $h > 0$, $t > 0$, $\delta \in [0,1]$, $\alpha \in [0,n]$, $\beta \in [0, \alpha+\delta]$, $x \in X^\beta$.

PROOF Choose $a \in \mathbb{R}$ such that $a < \operatorname{Re}\lambda$ for all $\lambda \in \sigma(A)$. Define $A_1 = A - a$ and $z = A_1^\alpha(e^{-A(t+h)} - e^{-At})x$. Note that

$$z = \left(A_1^{\alpha+\delta-\beta}e^{-At}\right)(e^{-A_1 h}e^{-ah} - 1)A_1^{-\delta}A_1^\beta x.$$

Theorem 6.1.15 implies

$$\|z\| \leq c_1 t^{\beta-\alpha-\delta}e^{-at}\|(e^{-A_1 h}e^{-ah} - 1)A_1^{-\delta}A_1^\beta x\|$$

and hence the bound follows from (6) of Theorem 6.1.8. □

6.2 The Inhomogeneous Problem - Part II

Let us assume that A is a sectorial operator in a Banach space X, $\tau \in (0,\infty)$ and $f : [0,\tau] \to X$ is Bochner integrable and bounded. The purpose of this section is to derive some basic properties of

$$v(t) \equiv \int_0^t e^{-A(t-s)}f(s)ds$$

which will then be needed to obtain regularity of mild solutions of semilinear parabolic equations.

Theorem 6.2.1 *$v \in C^{1-\alpha}([0,\tau], X^\alpha) \cap C^\alpha([0,\tau], X)$ for all $\alpha \in (0,1)$.*

PROOF Suppose $a < \operatorname{Re}\lambda$ for all $\lambda \in \sigma(A)$ and $A_1 = A - a$. Note that

$$v(t+h) - v(t) = \int_0^t g_1(s)ds + \int_t^{t+h} g_2(s)ds$$

for $0 \le t \le t+h \le \tau$, where

$$g_1(s) = (e^{-A(t+h-s)} - e^{-A(t-s)})f(s), \quad g_2(s) = e^{-A(t+h-s)}f(s).$$

Note that

$$A_1^\alpha g_1(s) = -AA_1^{-1}\int_{t-s}^{t+h-s} A_1^{1+\alpha}e^{-Au}f(s)du.$$

Hence the boundedness of f and Theorem 6.1.15 imply

$$\begin{aligned}
\|A_1^\alpha g_1(s)\| &\le c_1\int_{t-s}^{t+h-s} u^{-1-\alpha}du \\
&= c_1((t-s)^{-\alpha} - (t+h-s)^{-\alpha})/\alpha \\
\int_0^t \|A_1^\alpha g_1(s)\|ds &\le h^{1-\alpha}c_1/(\alpha-\alpha^2) \\
\int_t^{t+h} \|A_1^\alpha g_2(s)\|ds &\le c_2\int_t^{t+h}(t+h-s)^{-\alpha}ds \\
&= h^{1-\alpha}c_2/(1-\alpha)
\end{aligned}$$

and Theorem 4.2.10 implies that $v \in C^{1-\alpha}([0,\tau], X^\alpha)$. Since

$$\|v(t+h) - v(t)\| \le \|A_1^{\alpha-1}\|\|v(t+h) - v(t)\|_{1-\alpha}$$

the above result implies that $v \in C^\alpha([0,\tau], X)$. □

Define

$$G(t) = \int_0^t e^{-A(t-s)}(f(s) - f(t))ds$$

and note that

$$v(t) = G(t) + \int_0^t e^{-As}f(t)ds \quad \text{for} \quad t \in [0,\tau]. \tag{6.4}$$

Lemma 6.2.2 *If $\nu \in (0,1)$, $\alpha \in [0,\nu)$ and $f \in C^\nu([0,\tau], X)$, then, $G(t) \in X^{1+\alpha}$ for $t \in [0,\tau]$ and $AG \in C^{\nu-\alpha}([0,\tau], X^\alpha)$.*

PROOF Choose $0 \le t < t+h \le \tau$ and note

$$G(t+h) - G(t) = \int_0^t g_1(s)ds + \int_0^t g_2(s)ds + \int_t^{t+h} g_3(s)ds \tag{6.5}$$

where

$$g_1(s) = (e^{-A(t+h-s)} - e^{-A(t-s)})(f(s) - f(t))$$
$$g_2(s) = e^{-A(t+h-s)}(f(t) - f(t+h))$$
$$g_3(s) = e^{-A(t+h-s)}(f(s) - f(t+h)).$$

Since

$$A(e^{-A(t+h-s)} - e^{-A(t-s)}) = e^{-A(t-s)/2}A(e^{-A(h+(t-s)/2)} - e^{-A(t-s)/2})$$

Theorem 6.1.15 and (4.39) imply that

$$\|Ag_1(s)\|_\alpha \le c_2 h(t-s+2h)^{-1}(t-s)^{\nu-\alpha-1}.$$

Hence

$$\int_0^t \|Ag_1(s)\|_\alpha ds \le c_2 h^{\nu-\alpha} \int_0^{t/h} (2+r)^{-1} r^{\nu-\alpha-1} dr \le c_3 h^{\nu-\alpha}. \tag{6.6}$$

(3) of Theorem 4.3.1 implies

$$A\int_0^t g_2(s)ds = e^{-Ah}(1 - e^{-At})(f(t) - f(t+h)).$$

Hence

$$\left\| A\int_0^t g_2(s)ds \right\|_\alpha \le c_4 h^{\nu-\alpha}. \tag{6.7}$$

Theorem 6.1.15 and (4.37) imply

$$\int_t^{t+h} \|Ag_3(s)\|_\alpha ds \le c_5 \int_t^{t+h} (t+h-s)^{\nu-\alpha-1} ds = c_5(\nu-\alpha)^{-1}h^{\nu-\alpha}.$$

This, (6.7), (6.6), (6.5) and Theorem 4.2.10 imply first that $G(t+h) - G(t) \in \mathcal{D}(A)$ and then that $AG \in C^{\nu-\alpha}([0,\tau], X^\alpha)$. □

Theorem 6.2.3 *If $\nu \in (0,1)$, $\alpha \in [0,\nu)$, $\varepsilon \in (0,\tau)$ and $f \in C^\nu([0,\tau], X)$, then*

(1) $v : (0,\tau) \to X^\alpha$ is differentiable, $v' \in C^{\nu-\alpha}([\varepsilon,\tau], X^\alpha)$, $v' \in C([0,\tau], X)$

(2) $v \in C([0,\tau], X^1)$ and $Av \in C^\nu([\varepsilon,\tau], X)$

(3) $v' + Av = f$ on $[0,\tau]$.

PROOF (6.4) and Lemma 6.2.2 imply that $v(t) \in \mathcal{D}(A)$ and

$$Av(t) = AG(t) + f(t) - e^{-At}f(t) \quad \text{for} \quad t \in [0,\tau]. \tag{6.8}$$

Thus, $Av \in C([0,\tau],X)$. Theorems 4.7.3 and 4.2.10 imply that

$$v(t) + \int_0^t Av(s)ds = \int_0^t f(s)ds.$$

Hence $v' + Av = f$ on $[0,\tau]$ and $v' \in C([0,\tau],X)$. (4.38) implies that $e^{-A\cdot}f \in C^\nu([\varepsilon,\tau],X)$ which implies that $Av \in C^\nu([\varepsilon,\tau],X)$. (3) and (6.8) imply

$$v'(t) + AG(t) = e^{-At}f(t).$$

Hence Lemmas 6.2.2, 6.1.15 and 6.1.17 imply that $v' \in C^{\nu-\alpha}([\varepsilon,\tau],X^\alpha)$. This and Theorem 4.2.10 imply that $v : (0,\tau) \to X^\alpha$ is differentiable. □

EXAMPLE 6.2.4 Consider an abstract delay equation of the form

$$u'(t) + Au(t) = F(u(t-1)) \quad \text{for} \quad t > 0, \tag{6.9}$$

$$u(t) = \varphi(t) \quad \text{for} \quad t \in [-1,0],$$

where A is a sectorial operator in a Banach space X, $F : X^\alpha \to X$ is locally Lipschitz for some $\alpha \in [0,1)$, $\varphi \in C_H([-1,0],X^\alpha)$ and $\varphi(0) \in X^\beta$ for some $\beta > \alpha$.

Define $v_0(t) = \varphi(t-1)$ for $t \in [0,1]$. Having $v_n \in C([0,1],X^\alpha)$ we can define $v_{n+1} \in C([0,1],X^\alpha)$ by

$$v_{n+1}(t) = e^{-At}v_n(1) + \int_0^t e^{-A(t-s)}F(v_n(s))ds \quad \text{for} \quad t \in [0,1],\ n \geq 0.$$

Since $v_n(1) = v_{n+1}(0)$ we can define $u \in C([-1,\infty),X^\alpha)$ by

$$u(t) = v_n(t-n+1) \quad \text{for} \quad t \in [n-1,n].$$

It can be easily seen that

$$u(t) = e^{-At}\varphi(0) + \int_0^t e^{-A(t-s)}F(u(s-1))ds \quad \text{for} \quad t \geq 0$$

and hence Lemma 6.1.17 and Theorem 6.2.1 imply that $u \in C_H([-1,\infty),X^\alpha)$. Theorems 6.2.3 and 4.5.14 then imply (6.9).

6.3 Global Version

When the nonlinearity is globally Lipschitz one can easily obtain global existence, uniqueness and continuous dependence. This is done in Theorem 6.3.3. Using the retraction map one can extend these results to locally Lipschitz nonlinearities - which is done in the next section.

Lemma 6.3.1 *Suppose $\ell \in (0,\infty)$, $p > 1$, $h \in L^p(0,\ell)$ and $1 \le q \le \infty$. Define*

$$(Kg)(x) = \int_0^x h(x-s)g(s)ds \quad \text{for} \quad g \in L^q(0,\ell),\ x \in (0,\ell). \tag{6.10}$$

Then, $K \in \mathfrak{B}(L^q(0,\ell))$, $\lim_{n\to\infty} \|K^n\|^{1/n} = 0$ and also $K^m \in \mathfrak{B}(L^1(0,\ell), L^\infty(0,\ell))$ where $1/m \le 1 - 1/p$.

PROOF Let $g = h = 0$ in $\mathbb{R}\backslash(0,\ell)$. Since $h \in L^1(\mathbb{R})$ and $Kg = h * g$ on $(0,\ell)$, Young's Lemma 3.1.5 implies that $K \in \mathfrak{B}(L^q(0,\ell))$ and $\|K\| \le \|h\|_1$.

Note $K^n g = h_n * g$ on $(0,\ell)$ where $h_1 = h$ and $h_{n+1} = h_n * h$ for $n \ge 1$. Define $1/p_n = 1 - n/m$ for $1 \le n \le m$. We claim that $h_i \in L^{p_i}(\mathbb{R})$ for $1 \le i \le m$. This is obvious if $i = 1$. If the claim is true for some $i < m$, then $1 + 1/p_{i+1} = 1/p_1 + 1/p_i$ and hence Lemma 3.1.5 implies that the claim is true for $i+1$. Therefore $h_m \in L^\infty(\mathbb{R})$, $K^m \in \mathfrak{B}(L^1(0,\ell), L^\infty(0,\ell))$ and

$$|(K^m g)(x)| \le \|h_m\|_\infty \int_0^x |g(s)|ds.$$

Hence, for $i \ge 1$,

$$|(K^{im} g)(x)| \le \frac{\|h_m\|_\infty^i}{(i-1)!} \int_0^x (x-s)^{i-1}|g(s)|ds \quad \text{for} \quad x \in (0,\ell)$$

and $\|K^{im}\| \le \|\ell h_m\|_\infty^i / i!$. If $\varepsilon > 0$ and $n = im + j$, with $0 \le j \le m-1$, then

$$\varepsilon^{-n}\|K^n\| \le (\|K\|/\varepsilon)^j \|\varepsilon^{-m}\ell h_m\|_\infty^i / i! \overset{i\to\infty}{\longrightarrow} 0$$

and hence $\|K^n\|^{1/n} < \varepsilon$ for large enough n. □

Theorem 6.3.2 (Gronwall) *Suppose $\ell \in (0,\infty)$ and $f \in L^1(0,\ell)$ is such that*

$$f(x) \le g(x) + \int_0^x h(x-s)f(s)ds \quad \text{for} \quad x \in (0,\ell) \text{ a.e.},$$

where $h \in L^p(0,\ell)$, $p > 1$, $h \ge 0$, $g \in L^q(0,\ell)$, $q \in [1,\infty]$. Let $K \in \mathfrak{B}(L^q(0,\ell))$ be given by (6.10). Then, $f \le (1-K)^{-1}g = g + Kg + K^2 g + \cdots$.

PROOF Note that $f \le g + Kg + \cdots + K^n g + K^{n+1} f$ for $n \ge 0$. Hence Lemma 6.3.1 and Theorem 1.6.8 imply the assertion. □

Theorem 6.3.3 *Suppose A is a sectorial operator in a Banach space X, $\alpha \in [0,1)$, $\tau \in (0,\infty)$, $H : [0,\tau] \times X^\alpha \to X$ is continuous and such that for some $L < \infty$*

$$\|H(t,x) - H(t,y)\| \le L\|x - y\|_\alpha \quad \text{for} \quad 0 \le t \le \tau,\ x, y \in X^\alpha.$$

Then, for each $u_0 \in X^\alpha$ there exists a unique $u \in C([0,\tau], X^\alpha)$ which satisfies

$$u(t) = e^{-At}u_0 + \int_0^t e^{-A(t-s)}H(s,u(s))\,ds \quad \textit{for} \quad 0 \le t \le \tau.$$

Moreover, if $G : X^\alpha \to C([0,\tau], X^\alpha)$ is given by $G(u_0) = u$, then there exists $c < \infty$ such that

$$\|G(x)(t) - G(y)(t)\|_\alpha \le c\|x - y\|_\alpha \quad \textit{for} \quad 0 \le t \le \tau,\ x, y \in X^\alpha. \tag{6.11}$$

PROOF Abbreviate $Y = C([0,\tau], X^\alpha)$ and note that Y is a Banach space with the sup norm (Example 1.3.2). For $u \in Y$ define Tu by

$$(Tu)(t) = e^{-At}u_0 + \int_0^t e^{-A(t-s)}H(s,u(s))\,ds \quad \text{for} \quad 0 \le t \le \tau.$$

Continuity in X^α of $t \to e^{-At}u_0$ and Theorem 6.2.1 imply that $Tu \in Y$ for $u \in Y$. If $u, v \in Y$, then Theorems 4.2.10 and 6.1.15 imply

$$\|(Tu)(t) - (Tv)(t)\|_\alpha \le c_1 \int_0^t (t-s)^{-\alpha}\|u(s) - v(s)\|_\alpha ds. \tag{6.12}$$

Hence $\|T^j u - T^j v\|_\alpha \le K^j \|u - v\|_\alpha$ where K is given by (6.10), with $h(s) = c_1 s^{-\alpha}$, $q = \infty$. Lemma 6.3.1 implies that $\|K^n\| < 0.5$ for some n and hence

$$\|(T^n u)(t) - (T^n v)(t)\|_\alpha \le 0.5 \sup_{0 \le s \le \tau} \|u(s) - v(s)\|_\alpha \quad \text{for} \quad 0 \le t \le \tau.$$

Thus, the Contraction Mapping Theorem 1.1.3 implies existence of a unique $u \in Y$ such that $Tu = u$.

Choose $x, y \in X^\alpha$ and let $u = G(x), v = G(y)$. For $0 \le t \le \tau$ we have

$$\|u(t) - v(t)\|_\alpha \le c_2\|x - y\|_\alpha + c_1 \int_0^t (t-s)^{-\alpha}\|u(s) - v(s)\|_\alpha ds.$$

Hence the Gronwall Theorem 6.3.2 implies (6.11). □

6.4 Main Results

In this section assume that

(1) A is a sectorial operator in a Banach Space X and $a < \operatorname{Re}\lambda$ for all $\lambda \in \sigma(A)$

(2) $\alpha \in [0,1)$ and $\mathcal{U}$ is a nonempty open subset of X^α

(3) $T \in (0,\infty]$, $F : [0,T) \times \mathcal{U} \to X$ is such that for every $t \in [0,T)$ and $x \in \mathcal{U}$ there exist $\vartheta \in (0,1]$ and $c, \delta \in (0,\infty)$ such that

$$\|F(t_2,x_2) - F(t_1,x_1)\| \le c(|t_2 - t_1|^{\vartheta} + \|x_2 - x_1\|_\alpha) \tag{6.13}$$

whenever $t_i \in [0,T)$, $x_i \in \mathcal{U}$ and $|t_i - t| + \|x_i - x\|_\alpha < \delta$ for $i = 1,2$.

For $\tau \in (0,T]$ let $\mathcal{S}(\tau)$ denote the collection of $u \in C([0,\tau),\mathcal{U})$ such that

$$u'(t) \text{ exists, } u(t) \in \mathcal{D}(A) \text{ and } u'(t) + Au(t) = F(t,u(t)) \quad \text{for all} \quad t \in (0,\tau). \tag{6.14}$$

We shall refer to such equations as **semilinear parabolic equations**.

Observe that for u to belong to $\mathcal{S}(\tau)$, the set of solutions of (6.14), it is required that u be continuous in $\mathcal{U}$ and hence in X^α. In applications one can usually choose for α any number in a certain interval. It will be shown that the solution of an initial value problem does not depend on a particular choice of α. Note also that it is required that $u(0) \in \mathcal{U}$. Hence, solutions with 'rough' initial values are excluded though they may exist (Exercise 8). One can define $\mathcal{S}(\tau)$ in various other ways. However, in order to get uniqueness for a given initial value, some care is needed as Example 6.4.1 shows [17]. Several otherwise excellent texts make false uniqueness claims (Henry [11, page 54], Pazy [20, page 196]).

EXAMPLE 6.4.1 Assume that $a = 0$, $0 < \alpha < 1$, $\mathcal{U} = X^\alpha$ and that $F(x) = \|A^\alpha x\|^p x$. For $p > 1/\alpha$ it is easy to find concrete examples such that $\int_0^1 g(t)dt = \infty$ where $g(t) = \|A^\alpha e^{-At}z\|^p$ for some $z \in X$ (Exercise 9). In such a case define $u(0) = 0$ and

$$u(t) = \left(1 + p\int_t^1 g(s)ds\right)^{-1/p} e^{-At}z.$$

Then, $u \not\equiv 0$ and

(a) $u \in C([0,1),X) \cap C^1((0,1),X)$

(b) $u(t) \in \mathcal{D}(A)$ and $u'(t) + Au(t) = F(u(t))$ for $0 < t < 1$

(c) $F(u(\cdot))$ is Lipschitz continuous on $(\delta,1)$ for every $\delta \in (0,1)$

(d) $\int_0^1 \|F(u(t))\|dt < \infty$.

Thus, this u and 0 are two different solutions of 6.14 with the same initial value. u does not belong to the solution set $\mathcal{S}(1)$ because u is not continuous in $\mathcal{U}$ at $t = 0$. For a PDE version of this example see Exercise 10.

EXAMPLE 6.4.2 Assuming only that $f \in C^1(\mathbb{R}^{n+1},\mathbb{R})$ and $f(0) = 0$ let us show that

$$u_t = \Delta u + f(u,\nabla u) \quad \text{in } \mathbb{R}^n,\ t > 0$$

can be formulated as a semilinear parabolic equation in the complex $X = C_{\ell 0}$.

Fix $\alpha \in (1/2, 1)$ and let $X^\alpha = \mathcal{D}((1-\overline{\Delta})^\alpha)$. If $u \in X^\alpha$, then, in view of Example 6.1.11, $D_i u \in X$ and we can define

$$F(u) = f(\mathrm{Re}\, u, \mathrm{Re}\, \nabla u) \in X \quad \text{for all} \quad u \in X^\alpha.$$

Choose any $r \in (0, \infty)$. In view of Example 6.1.11, there exists $c < \infty$ such that if $w \in X^\alpha$, then $\|w\| \le c\|w\|_\alpha$ and $\|D_i w\| \le c\|w\|_\alpha$ for all $i = 1, \dots, n$. Let L be the bound of all first order partial derivatives of f while each of its $n+1$ variables lies in the interval $[-cr, cr]$. Note that

$$\|F(u) - F(v)\| \le (n+1)Lc\|u - v\|_\alpha \quad \text{if } \|u\|_\alpha < r \text{ and } \|v\|_\alpha < r.$$

Theorem 6.4.3 *Suppose $\tau \in (0, T]$. Then $u \in \mathcal{S}(\tau)$ if and only if $u \in C([0,\tau), \mathcal{U})$ and*

$$u(t) = \mathrm{e}^{-At}u(0) + \int_0^t \mathrm{e}^{-A(t-s)}F(s, u(s))ds \quad \textit{for all} \quad t \in [0, \tau). \tag{6.15}$$

Moreover, if $u \in \mathcal{S}(\tau)$, then

(a) $u', Au, F(\cdot, u) \in C_H((0,\tau), X)$. Moreover, if one can always choose $\vartheta \ge 1-\alpha$ in (6.13), then, $u : (0,\tau) \to X^\eta$ is differentiable and $u' \in C_H((0,\tau), X^\eta)$ for every $\eta \in [0, 1-\alpha)$

(b) if $\gamma \in [\alpha, 1]$ and $u(0) \in X^\gamma$, then $u \in C([0,\tau), X^\gamma)$.

PROOF If $u \in C([0,\tau),\mathcal{U})$ and $f(t) = F(t, u(t))$, then $f \in C([0,\tau), X)$. Thus, if $u \in \mathcal{S}(\tau)$, then Theorem 4.7.1 implies (6.15).

Assume now that $u \in C([0,\tau),\mathcal{U})$ and that (6.15) holds. Note that

$$u(t) = e^{-At}u(0) + g(t) \quad \text{where} \quad g(t) = \int_0^t e^{-A(t-s)}f(s)ds. \tag{6.16}$$

Choose any $0 < a < b < \tau$. Since $f \in C([0,\tau), X)$, Theorem 6.2.1 implies

$$g \in C^{1-\beta}([0,b], X^\beta) \quad \text{for all} \quad \beta \in (0,1). \tag{6.17}$$

Lemma 6.1.17 implies that for $a \le s \le t \le b$ and $\beta \in (0,1)$ we have

$$\|e^{-At}u(0) - e^{-As}u(0)\|_\beta \le c(t-s)^{1-\beta}$$

which together with (6.16) and (6.17) implies that $u \in C^{1-\beta}([a,b], X^\beta)$. This and the assumed regularity of F imply that f is locally Hölder continuous on $[a,b]$. Hence, $f \in C^\nu([a,b], X)$ for some $\nu \in (0,1)$. Actually, if one can always choose $\vartheta \ge 1-\alpha$ in (6.13), then any $\nu \in (0,1)$ with $\nu \le 1-\alpha$ will do. Let

$$v(t) = e^{-At}u(a) + \int_0^t e^{-A(t-s)}f(a+s)ds \quad \text{for} \quad t \in [0, b-a].$$

Theorem 6.2.3 implies that $v'(t)$ exists, $v(t) \in \mathcal{D}(A)$, $v'(t) + Av(t) = f(a+t)$ for $t \in (0, b-a)$ and that

$$Av \in C^\nu([\varepsilon, b-a], X), \; v' \in C^{\nu-\eta}([\varepsilon, b-a], X^\eta) \quad \text{for} \quad \varepsilon \in (0, b-a), \; \eta \in [0, \nu).$$

On the other hand it is easy to see, by rearranging (6.15), that $v(t) = u(t+a)$ and since a, b are arbitrary we see that $u \in \mathcal{S}(\tau)$ and (a) holds.

If $\gamma \in (\alpha, 1)$, then (b) follows from (6.16) and (6.17) and the continuity of $t \to e^{-At}u(0)$ in X^γ.

If $\gamma = 1$, then $t \to e^{-At}u(0)$ is in $C^{1-\alpha}([0,b], X^\alpha)$ by Lemma 6.1.17. Hence, (6.16) and (6.17) imply that $u \in C^\beta([0,b], X^\alpha)$ for some $\beta \in (0,1)$. Thus, $f \in C^\nu([0,b], X)$ for some $\nu \in (0,1)$ and therefore $Au \in C([0,b], X)$ by Theorem 6.2.3. □

Now the results of Section 5.2 are going to be adapted to semilinear parabolic equations.

Lemma 6.4.4 *Suppose $\tau \in (0,T)$ and $w \in C([0,\tau], \mathcal{U})$. Then there exist $\delta > 0$, $L < \infty$ and a continuous $H : [0,\tau] \times X^\alpha \to X$ such that*

(i) if $t \in [0,\tau]$ and $\|z - w(t)\|_\alpha < \delta$, then $z \in \mathcal{U}$ and $H(t,z) = F(t,z)$

(ii) $\|H(t,x) - H(t,y)\| \le L\|x-y\|_\alpha$ for $t \in [0,\tau]$, $x, y \in X^\alpha$.

PROOF For $t \in [0,\tau]$ choose $\delta(t) > 0$ and $L(t) < \infty$ such that

$$x, y \in \mathcal{U} \quad \text{and} \quad \|F(s,x) - F(s,y)\| \le L(t)\|x-y\|_\alpha$$

when $\|x - w(t)\|_\alpha < \delta(t)$, $\|y - w(t)\|_\alpha < \delta(t)$ and $|s-t| < \delta(t)$. Let $\mu(t) \in (0, \delta(t))$ be such that $\|w(s) - w(t)\|_\alpha < \delta(t)/2$ if $|s-t| < \mu(t)$. Compactness of $[0,\tau]$ implies that $[0,\tau] \subset B(t_1, \mu(t_1)) \cup \cdots \cup B(t_n, \mu(t_n))$ for some $t_i \in [0,\tau]$. Let $L = 2\max_i L(t_i)$ and $\delta = \min_i \delta(t_i)/4$. A short verification gives that if $t \in [0,\tau]$ and $x, y \in B(w(t), 2\delta)$, then $x, y \in \mathcal{U}$ and $\|F(t,x) - F(t,y)\| \le L\|x-y\|_\alpha/2$.

Define $H(t,x) = F(t, w(t) + R_\delta(x - w(t)))$ for $t \in [0,\tau]$ and $x \in X^\alpha$, where R_δ is the retraction map in X^α (Theorem 5.1.4). □

Existence for the initial value problem:

Theorem 6.4.5 *For each $u_0 \in \mathcal{U}$ there exist $\theta \in (0,T)$ and $u \in \mathcal{S}(\theta)$ such that $u(0) = u_0$. Furthermore, $\lim_{n\to\infty} \sup_{0\le t<\theta} \|u_n(t) - u(t)\|_\alpha = 0$ where $u_k \in C([0,\theta), \mathcal{U})$ are defined by $u_1(t) = u_0$ and*

$$u_{k+1}(t) = e^{-At}u_0 + \int_0^t e^{-A(t-s)}F(s, u_k(s))ds \quad \text{for} \quad t \in [0,\theta), \; k = 1, 2, \ldots$$

PROOF Pick $\tau \in (0,T)$ and let $w(t) = u_0$ for $t \in [0,\tau]$. Let δ and H be as given in Lemma 6.4.4 and let $u \in C([0,\tau], X^\alpha)$ be as given by Theorem 6.3.3. Continuity of u implies that for some $\theta \in (0,\tau]$ we have for $t \in [0,\theta]$ that $\|u(t) - w(t)\|_\alpha < \delta/2$. Hence $H(t,u(t)) = F(t,u(t))$ and therefore $u \in \mathcal{S}(\theta)$ by Theorem 6.4.3. (6.12) implies that T is a contraction when τ is small enough, which implies that $\|u_k - w\|_\alpha < \delta$ on $[0,\theta]$ for $k \geq 1$ and that $u_k \to u$. □

Uniqueness for the initial value problem:

Theorem 6.4.6 *If $0 < \mu \leq \theta \leq T$, $w \in \mathcal{S}(\mu)$, $u \in \mathcal{S}(\theta)$ and $w(0) = u(0)$, then, $w(t) = u(t)$ for $t \in [0,\mu)$.*

PROOF Choose any $\tau \in (0,\mu)$ and let δ and H be as given in Lemma 6.4.4. Let τ' be the largest number in $[0,\tau]$ such that $\|u(t) - w(t)\|_\alpha < \delta$ for $t \in [0,\tau')$. Note that $\tau' > 0$. Note that $H(t,u(t)) = F(t,u(t))$ and $H(t,w(t)) = F(t,w(t))$ for $t \in [0,\tau')$. Hence, Theorems 6.4.3 and 6.3.3 imply that $u = w$ on $[0,\tau')$. Continuity of $u - w$ and maximality of τ' imply that $\tau' = \tau$. □

Solutions of real equations, with real initial conditions, are real:

Theorem 6.4.7 *Suppose that Y is a closed real subspace of X such that*

(1) $(A - \lambda)^{-1} Y \subset Y$ for some $\lambda < a$

(2) $F(t,x) \in Y$ when $t \in [0,T)$ and $x \in Y \cap \mathcal{U}$.

Suppose also that $u \in \mathcal{S}(\theta)$ for some $\theta \in (0,T]$ and that $u(0) \in Y$. Then

$$u(t) \in Y \quad \text{for all} \quad t \in [0,\theta).$$

PROOF Choose any $\tau \in [0,\theta)$ such that $u(t) \in Y$ for $t \in [0,\tau]$. Theorems 6.4.5 and 6.4.3 imply that there exist $\delta \in (0,\theta - \tau)$ and $v \in C([0,\delta),\mathcal{U})$ such that

$$v(t) = e^{-At}u(\tau) + \int_0^t e^{-A(t-s)}F(\tau + s, v(s))ds \quad \text{for all} \quad t \in [0,\delta)$$

and that $v = \lim_{n\to\infty} v_n$ for some $v_n \in C([0,\delta),\mathcal{U})$ given by $v_1(t) = u(\tau)$ and

$$v_{n+1}(t) = e^{-At}u(\tau) + \int_0^t e^{-A(t-s)}F(\tau + s, v_n(s))ds$$

for $t \in [0,\delta)$ and $n = 1,2,\ldots$. Since $u(\tau) \in Y$, Theorem 4.6.1 and Corollary 4.2.5 imply that the range of every v_n is in Y. Hence, the range of v is in Y. Let $w = u$ on $[0,\tau]$ and $w(t) = v(t - \tau)$ for $t \in [\tau, \tau + \delta)$. It is easy to see that $w \in \mathcal{S}(\tau + \delta)$. The uniqueness Theorem 6.4.6 implies that $w = u$ on $[0,\tau + \delta)$. Hence, $u(t) \in Y$ for $t \in [0,\tau + \delta)$. Therefore, the supremum of such τ has to be equal to θ. □

EXAMPLE 6.4.8 The solution of the generalized heat equation, given in Example 6.4.2, is real when the initial value is real (see also Example 4.6.3).

Continuous dependence on the initial value:

Theorem 6.4.9 *Suppose* $0 < \tau < \mu \leq T$ *and* $w \in \mathcal{S}(\mu)$. *Then there exist* $\varepsilon > 0$ *and* $c < \infty$ *such that for every* $x \in X^\alpha$, *with* $\|x - w(0)\|_\alpha < \varepsilon$, *there exists* $u \in \mathcal{S}(\tau)$ *such that* $u(0) = x$ *and*

$$\|u(t) - w(t)\|_\alpha \leq c\|u(0) - w(0)\|_\alpha \quad \textit{for all} \quad t \in [0, \tau). \tag{6.18}$$

PROOF Let δ and H be as in Lemma 6.4.4. Let c be as in (6.11) and let $\varepsilon \in (0, \delta/c)$. Suppose $\|x - w(0)\|_\alpha < \varepsilon$. Let $u = G(x)$, where G is as in Theorem 6.3.3. (6.18) follows from (6.11) and the fact that $w = G(w(0))$ on $[0, \tau]$. Since $c\varepsilon < \delta$ Lemma 6.4.4 implies that $H(t, u(t)) = F(t, u(t))$ for $t \in [0, \tau]$ and hence Theorem 6.4.3 implies that $u \in \mathcal{S}(\tau)$. □

Maximal interval of existence:

Theorem 6.4.10 *Suppose* $u_0 \in \mathcal{U}$. *Then, there exist* $\tau \in (0, T]$, $u \in \mathcal{S}(\tau)$ *such that* $u(0) = u_0$ *and which also have the property that* $\tau \geq \mu$ *for all* $\mu \in (0, T]$ *for which there exists* $v \in \mathcal{S}(\mu)$ *with* $v(0) = u_0$. *Moreover, if* $\tau < T$, *then either* $\sup_{0<t<\tau} \|F(t, u(t))\| = \infty$ *or there exists* $x \in \overline{\mathcal{U}} \backslash \mathcal{U}$ *such that* $\lim_{t\to\tau} \|u(t) - x\|_\alpha = 0$.

PROOF Theorems 6.4.5 and 6.4.6 imply the first assertion. Suppose $\tau < T$ and $\sup_{0<t<\tau} \|F(t, u(t))\| < \infty$. Theorems 6.4.3 and 6.2.1 imply that there exists $x \in X^\alpha$ such that $\lim_{t\to\tau} \|u(t) - x\|_\alpha = 0$. Clearly $x \in \overline{\mathcal{U}}$. The proof will be complete if we show that the assumption $x \in \mathcal{U}$ leads to a contradiction. So, assume $x \in \mathcal{U}$, choose $\tau' \in (\tau, T)$ and let $u(t) = x$ for $t \in [\tau, \tau']$. Let δ and H be as given in Lemma 6.4.4 but applied to $\tau', u \in C([0, \tau'], \mathcal{U})$ in place of τ, w. Theorem 6.3.3 implies existence of $v \in C([0, \tau'], X)$ such that

$$v(t) = e^{-At}u(0) + \int_0^t e^{-A(t-s)}H(s, v(s))ds \quad \text{for} \quad t \in [0, \tau']$$

and its uniqueness implies that $v = u$ on $[0, \tau)$. Continuity of $v - u$ implies that for some $\mu \in (\tau, \tau']$ we have that $\|v(t) - u(t)\|_\alpha < \delta$ for $t \in [0, \mu]$ and hence $H(t, v(t)) = F(t, v(t))$ for $t \in [0, \mu]$. Therefore Theorem 6.4.3 implies that $v \in \mathcal{S}(\mu)$ - which is a contradiction to the maximality of τ. □

Stability:

Theorem 6.4.11 *If $a > 0$, $T = \infty$ and if for each $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$x \in \mathcal{U} \text{ and } \|F(t,x)\| \leq \varepsilon \|x\|_\alpha \text{ whenever } x \in X^\alpha,\ \|x\|_\alpha < \delta \text{ and } t \geq 0,$$

then there exist $\mu > 0$ and $c < \infty$ such that for every $x \in X^\alpha$, with $\|x\|_\alpha < \mu$, there exists $u \in \mathcal{S}(\infty)$ such that $u(0) = x$ and

$$\|u(t)\|_\alpha \leq c\|x\|_\alpha e^{-at} \quad \text{for} \quad t \geq 0.$$

PROOF In view of Theorem 6.1.15 we can choose $M, \theta \in (0, \infty)$ such that

$$\|e^{-At}\| \leq Me^{-(a+\theta)t}, \quad \|(A-a)^\alpha e^{-At}\| \leq Mt^{-\alpha}e^{-(a+\theta)t} \quad \text{for} \quad t > 0.$$

Choose $\varepsilon > 0$ so that $\varepsilon M\theta^{\alpha-1}\Gamma(1-\alpha) < 1$ and let the corresponding $\delta > 0$ be given. Choose $\mu \in (0, \delta)$ so that

$$2M\mu < \delta(1 - \varepsilon M\theta^{\alpha-1}\Gamma(1-\alpha)). \tag{6.19}$$

Choose $x \in X^\alpha$ with $\|x\|_\alpha < \mu$ and pick the maximal $\tau > 0$ such that there exists $u \in \mathcal{S}(\tau)$ with $u(0) = x$ (Theorem 6.4.10). Note that for some $\eta \in (0, \tau]$ we have that $\|u(t)\|_\alpha < \delta$ for $0 \leq t < \eta$. Let τ' be the largest of such η. Note that for all $t \in [0, \tau')$

$$\begin{aligned}
\|u(t)\|_\alpha &\leq Me^{-(a+\theta)t}\|x\|_\alpha + \varepsilon M \int_0^t (t-s)^{-\alpha} e^{-(a+\theta)(t-s)} \|u(s)\|_\alpha ds \\
e^{at}\|u(t)\|_\alpha &\leq M\|x\|_\alpha + \varepsilon M \int_0^t (t-s)^{-\alpha} e^{-\theta(t-s)} e^{as} \|u(s)\|_\alpha ds \\
&\leq M\|x\|_\alpha + \varepsilon M\theta^{\alpha-1}\Gamma(1-\alpha) \sup_{0\leq s\leq t} e^{as}\|u(s)\|_\alpha.
\end{aligned}$$

Hence

$$\sup_{0\leq s\leq t} e^{as}\|u(s)\|_\alpha \leq \frac{M\|x\|_\alpha}{1 - \varepsilon M\theta^{\alpha-1}\Gamma(1-\alpha)}$$

and therefore

$$\|u(t)\|_\alpha \leq \frac{M\|x\|_\alpha e^{-at}}{1 - \varepsilon M\theta^{\alpha-1}\Gamma(1-\alpha)}.$$

This, the fact that $\|x\|_\alpha < \mu$ and (6.19) imply that $\|u(t)\|_\alpha < \delta/2$ for $t \in [0, \tau')$. Continuity of u and the definition of τ' imply that $\tau' = \tau$. This implies that $\|F(t, u(t))\| < \varepsilon\delta/2$ for $t \in [0, \tau)$ and that u stays away from the boundary of $\mathcal{U}$. Thus, $\tau = \infty$ by Theorem 6.4.10. □

If the spectrum of A contains a point with a negative real part (the point does not have to be isolated) and if F is genuinely nonlinear at 0, then we have instability:

Theorem 6.4.12 *Suppose that* $\operatorname{Re} z < 0$ *for some* $z \in \sigma(A)$, $T = \infty$ *and that there exist* $\delta, \theta, M \in (0, \infty)$ *such that*

$$x \in \mathcal{U} \text{ and } \|F(t,x)\| \le M\|x\|_\alpha^{1+\theta} \text{ whenever } x \in X^\alpha,\ \|x\|_\alpha < \delta \text{ and } t \ge 0. \tag{6.20}$$

Then there exists $c > 0$ *such that for each* $\varepsilon > 0$ *there exists* $u \in \mathcal{S}(\tau)$ *for which*

$$\|u(0)\|_\alpha < \varepsilon, \quad \|u(t)\|_\alpha > c \text{ for some } t \in (0,\tau), \quad u(0) \in \cap_{n=1}^\infty \mathcal{D}(A^n).$$

PROOF Define $z_0 = -\inf_{\zeta\in\sigma(A)} \operatorname{Re}\zeta$ and choose λ, μ so that

$$0 < \mu < z_0 < \lambda \quad \text{and} \quad \mu(1+\theta) > \lambda.$$

Note that Theorem 6.1.15 implies that

$$\|(A-a)^\alpha e^{-At}\| \le M_1\, t^{-\alpha} e^{\lambda t} \quad \text{for all} \quad t > 0. \tag{6.21}$$

Since there exists $\zeta \in \sigma(A)$ such that $\operatorname{Re}\zeta < -\mu$, we have, by Corollary 4.3.3, that $\sup_{t>0} \|e^{-At}x_1\|e^{-\mu t} = \infty$ for some $x_1 \in X$. Define

$$x_2 = (A-a)^{-\alpha} e^{-A1} x_1, \quad x = \|x_2\|_\alpha^{-1} x_2 \in \cap_{n=1}^\infty \mathcal{D}(A^n),$$

$$a(t) = \|e^{-At}x\|_\alpha e^{-\mu t}, \quad b(t) = \sup_{0\le s\le t} a(s).$$

Note that a is continuous, $\sup_{t>0} a(t) = \infty$, $a(0) = 1$, $\|e^{-At}x\|_\alpha = a(t)e^{\mu t}$.

Pick $c \in (0, \delta/3)$ so that

$$M_2 \equiv \Gamma(1-\alpha)(\mu(1+\theta)-\lambda)^{\alpha-1} 3MM_1(3c)^\theta < 1. \tag{6.22}$$

Let $\varepsilon > 0$ be given.

Choose $t_0 > 0$ such that $\varepsilon a(t_0)e^{\mu t_0} > 2c$. Find the smallest $t_1 > 0$ such that $a(t_1) = 1 + b(t_0)$. Note that

$$t_1 > t_0, \quad a(t) < a(t_1) \text{ for } t \in [0, t_1), \quad a(t_1) = b(t_1).$$

Choose ε_0 so that $\varepsilon_0 a(t_1)e^{\mu t_1} = 2c$ and note that $\varepsilon_0 < \varepsilon$, $\varepsilon_0 < 2c < 2\delta/3$.

Theorem 6.4.10 gives $u \in \mathcal{S}(\tau)$ such that $u(0) = \varepsilon_0 x$ and τ is maximal. Note that $\|u(0)\|_\alpha = \varepsilon_0 < \varepsilon$. Continuity implies that there exists $\tau_1 \in (0, \tau]$ such that

$$\|u(t)\|_\alpha \le 1.5\varepsilon_0 b(t) e^{\mu t} \quad \text{for} \quad t \in [0, \tau_1).$$

Let τ_2 be the maximal of such τ_1. Note that

$$\|u(t)\|_\alpha \le 1.5\varepsilon_0 b(t)e^{\mu t} \le 1.5\varepsilon_0 b(t_1)e^{\mu t_1} = 3c < \delta \tag{6.23}$$

for $t \in [0, \tau_2) \cap [0, t_1]$. Hence, (6.20) and (6.21) imply

$$\left\| \int_0^t e^{-A(t-s)} F(s, u(s)) ds \right\|_\alpha \leq M_1 M \int_0^t (t-s)^{-\alpha} e^{\lambda(t-s)} \|u(s)\|_\alpha^{1+\theta} ds$$

$$\begin{aligned}
&\leq M_1 M \int_0^t (t-s)^{-\alpha} e^{\lambda(t-s)} (1.5\varepsilon_0 b(s) e^{\mu s})^{1+\theta} ds \\
&\leq M_1 M (1.5\varepsilon_0 b(t))^{1+\theta} \int_0^t (t-s)^{-\alpha} e^{\lambda(t-s)} e^{\mu(1+\theta)s} ds \\
&\leq M_1 M (1.5\varepsilon_0 b(t) e^{\mu t})^{1+\theta} \Gamma(1-\alpha)(\mu(1+\theta) - \lambda)^{\alpha-1} \\
&\leq 0.5\varepsilon_0 b(t) e^{\mu t} 3 M_1 M (3\mathrm{c})^\theta \Gamma(1-\alpha)(\mu(1+\theta) - \lambda)^{\alpha-1} \\
&= \varepsilon_0 b(t) e^{\mu t} M_2/2. \qquad (6.24)
\end{aligned}$$

Therefore (6.15) implies that

$$\|u(t)\|_\alpha \leq (1 + M_2/2)\varepsilon_0 b(t) e^{\mu t} \quad \text{for} \quad t \in [0, \tau_2) \cap [0, t_1]. \qquad (6.25)$$

If $\tau_2 \leq t_1$, then the continuity, the fact that $M_2 < 1$ and the definition of τ_2 would imply that $\tau_2 = \tau$. However, in this case (6.23) and (6.20) would imply both boundedness of F on $[0, \tau)$ and that u stays away from the boundary of $\mathcal{U}$ - which is not possible according to Theorem 6.4.10. Therefore $\tau_2 > t_1$. By construction, $\|e^{-At_1} u(0)\|_\alpha = 2\mathrm{c}$. Hence (6.15) and (6.24) imply

$$2\mathrm{c} < \|u(t_1)\|_\alpha + 0.5\varepsilon_0 b(t_1) e^{\mu t_1} = \|u(t_1)\|_\alpha + \mathrm{c}$$

and $\|u(t_1)\|_\alpha > \mathrm{c}$. □

EXAMPLE 6.4.13 Since $-\Delta_\ell$ is a sectorial operator (Example 4.5.4) in the complex C_ℓ, we see that the nonlinear heat equation (Section 5.3) is a semilinear parabolic equation ($\alpha = 0$) provided that we define $F(t, u) = f(\mathrm{Re}\, u)$ and modify $\mathcal{U}$ in the obvious way.

Since the solution $u \in \mathcal{S}_m(\tau)$, given by (5.10), is also the mild solution, we have, by Theorem 6.4.3, that $u \in \mathcal{S}(\tau)$ and hence

$$u'(t) = \Delta_\ell u(t) + f(u(t)) \quad \text{for} \quad t \in (0, \tau).$$

If $f(\mathrm{c}) = 0$, $f'(\mathrm{c}) > 0$ and f' is Hölder continuous near c, then $\sigma(-\Delta_\ell - f'(\mathrm{c})) = [-f'(\mathrm{c}), \infty)$. Hence Theorem 6.4.12 implies that the constant solution c is unstable.

6.5 Example: Navier-Stokes Equations

Velocity of a fluid $u = (u_1, u_2, u_3)$ and a scalar pressure p satisfy

$$u_t + (u \cdot \nabla)u = -\nabla p + \Delta u. \qquad (6.26)$$

The incompressibility of the fluid is specified by

$$\nabla \cdot u = 0. \tag{6.27}$$

These are the Navier-Stokes equations. We shall apply the results of Section 6.4 to the case of a flow that is periodic in space and has 0 mean velocity.

The periodicity cell is taken to be $\Omega = (-\pi, \pi)^3$. Define

$$H_0 = \{u \in L^2(\Omega) \mid \textstyle\int_\Omega u = 0\}, \quad H = H_0 \times H_0 \times H_0.$$

H is a Hilbert space with an inner product

$$(u, v) = (2\pi)^{-3} \int_\Omega u_1 \overline{v_1} + u_2 \overline{v_2} + u_3 \overline{v_3}$$

where $u = (u_1, u_2, u_3) \in H$ and $v = (v_1, v_2, v_3) \in H$. For $k \in \mathbb{Z}^3$ and $u \in H$ define

$$\ell_k(u) = (2\pi)^{-3} \int_\Omega u(x) e^{-ik\cdot x} dx \in \mathbb{C}^3.$$

Completeness of the Fourier series (Example 2.1.7) and Theorem 2.1.10 imply that for every $u \in H$,

$$u(x) = \sum \ell_k(u) e^{ik\cdot x} \tag{6.28}$$

where $\sum$ denotes, throughout this section, the sum over all $k \in \mathbb{Z}^3 \backslash \{0\}$ with convergence in H and not necessarily pointwise. Note also that

$$(u, v) = \sum \ell_k(u) \cdot \overline{\ell_k(v)} \quad \text{for} \quad u, v \in H \tag{6.29}$$

and that u is real valued a.e. iff $\overline{\ell_k(u)} = \ell_{-k}(u)$ for all $k \in \mathbb{Z}^3$.

A formal differentiation of (6.28) implies that in order for u to satisfy (6.27) we should have

$$k \cdot \ell_k(u) = 0 \quad \text{for all} \quad k \in \mathbb{Z}^3. \tag{6.30}$$

(6.30) makes sense, unlike (6.27), for every $u \in H$. Hence, the incompressibility condition (6.27) will be replaced with (6.30). Let X consist of all $u \in H$ which satisfy (6.30). Note that X is a closed subspace of H and hence a Hilbert space.

(6.30) suggests the following decomposition of a given $v \in H$. For $k \in \mathbb{Z}^3 \backslash \{0\}$ define $p_k = -i\ell_k(v) \cdot k/k \cdot k$ and note that

$$\ell_k(v) - ip_k k \perp k, \quad |\ell_k(v) - ip_k k|^2 + |p_k k|^2 = |\ell_k(v)|^2.$$

Lemma 2.1.4 implies that we can define $u \in H$ and $p \in H_0$ by

$$u(x) = \sum (\ell_k(v) - ip_k k) e^{ik\cdot x} \quad \text{and} \quad p(x) = \sum p_k e^{ik\cdot x}. \tag{6.31}$$

Note that u satisfies (6.30) and that

$$v = u + \nabla p$$

where $\nabla p = i \sum k p_k e^{ik\cdot x}$. Define $Q \in \mathfrak{B}(H)$ and $\Pi \in \mathfrak{B}(H, H_0)$ by

$$Qv = u \quad \text{and} \quad \Pi v = p.$$

Observe that Q is a projection whose range is X, $\|Q\| = 1$, $Q(\nabla p) = 0$ and that u, p are real valued if v is real valued. Using (6.31) and (6.29) it is also easy to see that Q is self-adjoint.

A formal differentiation of (6.28) implies that

$$(\Delta u)(x) = -\sum |k|^2 \ell_k(u) e^{ik\cdot x}$$

and that Δu satisfies the incompressibility condition (6.30) when u satisfies it. This suggests that we define the linear operator A in X by

$$\left.\begin{aligned} \mathcal{D}(A) &= \{u \in X \mid \textstyle\sum |k|^4 |\ell_k(u)|^2 < \infty\} \\ (Au)(x) &= \textstyle\sum |k|^2 \ell_k(u) e^{ik\cdot x} \quad \text{for} \quad u \in \mathcal{D}(A). \end{aligned}\right\} \tag{6.32}$$

Using (6.29) it is easy to see that A is symmetric. Theorem 2.6.3 implies that A is self-adjoint. Theorem 2.6.6 implies that $\sigma(A) \subset [1, \infty)$. By rearranging (6.32), so that it fits Examples 6.1.2 and 6.1.7, we see that for all $\beta > 0$ we have

$$\left.\begin{aligned} \mathcal{D}(A^\beta) &= \{u \in X \mid \textstyle\sum |k|^{4\beta} |\ell_k(u)|^2 < \infty\} \\ (A^\beta u)(x) &= \textstyle\sum |k|^{2\beta} \ell_k(u) e^{ik\cdot x} \quad \text{for} \quad u \in \mathcal{D}(A^\beta). \end{aligned}\right\} \tag{6.33}$$

Let us now analyze the $(u \cdot \nabla)u$ term in (6.26). Note that

$$((v \cdot \nabla)u)_n = v_1 \frac{\partial u_n}{\partial x_1} + v_2 \frac{\partial u_n}{\partial x_2} + v_3 \frac{\partial u_n}{\partial x_3} \quad \text{for} \quad n = 1, 2, 3 \tag{6.34}$$

where $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$. As is usual in this section, we take derivatives by formal differentiation of (6.28) and then we require that the resulting series converges in H, i.e.

$$\frac{\partial u_n}{\partial x_m}(x) = \sum i k_m \ell_k(u)_n e^{ik\cdot x}, \quad \sum |k_m \ell_k(u)_n|^2 < \infty.$$

Thus, in order that all $\partial u_n / \partial x_m$ exist in H we should have $\sum |k|^2 |\ell_k(u)|^2 < \infty$. Hence (6.33) implies that we should have $u \in \mathcal{D}(A^{1/2})$. (6.33) and (6.29) then imply

$$\int_\Omega \sum_{n,m=1}^{3} \left| \frac{\partial u_n}{\partial x_m} \right|^2 = (2\pi)^3 \|A^{1/2} u\|^2 \quad \text{for} \quad u \in \mathcal{D}(A^{1/2}). \tag{6.35}$$

(6.28) implies

$$|v(x)| \le \sum |\ell_k(v)e^{ik\cdot x}| = \sum |k|^{-2\beta}|k|^{2\beta}|\ell_k(v)|$$

and since $C_\beta \equiv (\sum |k|^{-4\beta})^{1/2} < \infty$ for $\beta > 3/4$ we have that

$$|v(x)|^2 \le C_\beta^2 \sum |k|^{4\beta}|\ell_k(v)|^2 = C_\beta^2 \|A^\beta v\|^2$$

$$|v(x)| \le C_\beta \|A^\beta v\| \quad \text{for} \quad v \in \mathcal{D}(A^\beta),\ \beta > 3/4,\ x \in \Omega. \tag{6.36}$$

(6.34), (6.35) and (6.36) imply that $(v\cdot\nabla)u \in L^2(\Omega)^3$ and since

$$\int_\Omega (v\cdot\nabla)u = \int_\Omega i\sum (v(x)\cdot k)\ell_k(u)e^{ik\cdot x}dx = i(2\pi)^3 \sum (\ell_{-k}(v)\cdot k)\ell_k(u) = 0$$

we have that $(v\cdot\nabla)u \in H$. We also have the bound

$$\|(v\cdot\nabla)u\| \le C_\beta \|A^\beta v\| \|A^{1/2}u\|.$$

Since (6.33) and (6.29) imply $\|A^{1/2}u\| \le \|A^\beta u\|$ we also have

$$\|(v\cdot\nabla)u\| \le C_\beta \|A^\beta v\| \|A^\beta u\| \quad \text{for} \quad u, v \in \mathcal{D}(A^\beta),\ \beta > 3/4. \tag{6.37}$$

Fix now any $\alpha \in (3/4, 1)$ and let X^α be $\mathcal{D}(A^\alpha)$ with the usual norm. Define $F : X^\alpha \to X$ by

$$F(u) = -Q((u\cdot\nabla)u).$$

Note that (6.37) implies

$$\|F(u)\| \le C_\alpha \|u\|_\alpha^2 \quad \text{for} \quad u \in X^\alpha \tag{6.38}$$

$$\|F(u) - F(v)\| \le C_\alpha (\|u\|_\alpha + \|v\|_\alpha)\|u - v\|_\alpha \quad \text{for} \quad u, v \in X^\alpha. \tag{6.39}$$

Applying Q to (6.26) yields the following semilinear parabolic equation in X,

$$u' + Au = F(u). \tag{6.40}$$

Let the initial velocity $u_0 \in X^\alpha$ be given. Theorem 6.4.10 implies that (6.40) has a solution $u \in \mathcal{S}(\tau)$, on some maximal interval $[0, \tau)$, such that $u(0) = u_0$. The solution is unique by Theorem 6.4.6. If pressure is given by

$$p = -\Pi((u\cdot\nabla)u),$$

then this pair u, p satisfies the Navier-Stokes equations.

Since $\sigma(A) \subset [1, \infty)$, Theorem 6.4.11 implies that if $\|u_0\|_\alpha$ is small enough, then the solution u exists for all time ($\tau = \infty$) and that $\|u(t)\|_\alpha \to 0$ exponentially.

If u_0 is real valued, then it is easy to see that Theorem 6.4.7 applies and thus $u(t)$ is real valued a.e. for all $t \in [0,\tau)$. Let us show now that when u_0 is real valued

$$\|u(t)\| \leq e^{-t}\|u_0\| \quad \textit{for} \quad t \in [0,\tau) \tag{6.41}$$

independently of the size of the initial $\|u_0\|_\alpha$ (it is an open question whether or not $\tau = \infty$ also for large $\|u_0\|_\alpha$). (6.41) is called the energy bound. The following procedure for obtaining it is called the energy method and it can be adapted to many other evolution equations. Note that

$$\frac{d}{dt}\|u\|^2 = 2\mathrm{Re}\,(u',u) = -2\mathrm{Re}\,(Au,u) + 2\mathrm{Re}\,R \quad \text{on} \quad (0,\tau) \tag{6.42}$$

where $R = (F(u),u) = -(Q((u\cdot\nabla)u),u)$. Since Q is a self-adjoint projection we have that $R = -((u\cdot\nabla)u,u)$. Using (6.28) gives

$$R = -(2\pi)^3 i \sum_{k,k'} (k\cdot \ell_{k'-k}(u))\ell_k(u)\cdot\overline{\ell_{k'}(u)}.$$

This, (6.30) and the fact that u is real valued imply that $\mathrm{Re}\,R = 0$ and hence (6.42) implies that

$$\frac{d}{dt}\|u\|^2 = -2\mathrm{Re}\,(Au,u) \quad \text{on} \quad (0,\tau).$$

(6.32) and (6.29) imply that $(Au,u) \geq \|u\|^2$. Hence

$$\frac{d}{dt}(e^{2t}\|u(t)\|^2) \leq 0$$

which implies the energy bound (6.41).

6.6 Example: A Stability Problem

There are many journals devoted exclusively to studies of fluid flows. In their articles, the governing equations are usually derived from the Navier-Stokes equations and one topic frequently addressed is stability of a certain solution. In the preceding Section 6.5 it was shown that the zero solution of the Navier-Stokes equations is stable. Stability analysis of nontrivial flows is usually more demanding and in unbounded domains an extra complication occurs which will be highlighted next.

Studying impulsive stretching of a surface in a viscous fluid leads to

$$\phi_t(x,t) - \phi_{xx}(x,t) + \phi^2(x,t) = \phi_x(x,t)\int_0^x \phi(s,t)ds \quad \text{for} \quad t>0,\ x>0. \tag{6.43}$$

One solution of this equation is Crane's solution e^{-x}. We want to show that this solution is stable for small perturbations.

If $g(x,t) = \phi(x,t) - \mathrm{e}^{-x}$, then (6.43) becomes

$$g_t - g_{xx} - g_x + \mathrm{e}^{-x}\left(g_x + 2g + \int_0^x g\right) = -g^2 + g_x \int_0^x g. \tag{6.44}$$

It is required that $g(0,t) = 0$ and that $g(x,t)$ decays as x increases.

To apply stability Theorem 6.4.11 or instability Theorem 6.4.12 one needs to have the sharp lower bound on the spectrum of the linear operator different from 0. The following argument suggests that this may not be the case for equation (6.44). The spectrum of the following operator in $\mathcal{H} = L^2(0,\infty)$

$$A_0 u = -u_{xx} - u_x \quad \text{for} \quad u \in \mathcal{D}(A_0) \equiv W^2(0,\infty) \cap W_0^1(0,\infty) \tag{6.45}$$

consists of those $z \in \mathbb{C}$ such that $(\operatorname{Im} z)^2 \leq \operatorname{Re} z$ (Exercise 12). So, for this operator Theorems 6.4.11 and 6.4.12 do not apply. The actual linear operator in (6.44) consists of A_0 and of a lower order perturbation of A_0. So it does not seem very likely that the introduction of lower order terms would move the spectrum to the right - which is needed to prove stability. In such situations specifying a different growth rate at infinity can sometimes move the spectrum. After some experimentation, we find that the following requirements will solve the problem

$$g(x,t) = \mathrm{e}^{-x/2} u(x,t) \quad \text{and} \quad u(\cdot,t) \in L^2(0,\infty). \tag{6.46}$$

Thus, (6.44) will in effect be studied in a weighted L^2. Using (6.46) in (6.44) gives

$$u_t + Au = f(u) \tag{6.47}$$

where

$$Au = -u_{xx} + \mathrm{e}^{-x} u_x + \left(\tfrac{1}{4} + \tfrac{3}{2}\mathrm{e}^{-x}\right) u + \mathrm{e}^{-x/2} \int_0^x \mathrm{e}^{-s/2} u \tag{6.48}$$

$$f(u) = -\mathrm{e}^{-x/2} u^2 + (u_x - u/2) \int_0^x \mathrm{e}^{-s/2} u. \tag{6.49}$$

Let $\mathcal{V} = W_0^1(0,\infty)$. Define $\mathfrak{F} : \mathcal{V} \times \mathcal{V} \to \mathbb{C}$ by

$$\mathfrak{F}(v,w) = \int_0^\infty \left(v'\overline{w'} + v'\overline{w}\mathrm{e}^{-x} + v\overline{w}\left(\tfrac{1}{4} + \tfrac{3}{2}\mathrm{e}^{-x}\right) + \overline{w}\mathrm{e}^{-x/2} \int_0^x \mathrm{e}^{-s/2} v(s)\,ds \right) dx.$$

Observe that if $v \in \mathcal{V}$ and $g(x) = \int_0^x \mathrm{e}^{-s/2} v(s)\,ds$, then

$$\begin{aligned}
\operatorname{Re}\mathfrak{F}(v,v) &= \operatorname{Re} \int_0^\infty \left(|v'|^2 + v'\overline{v}\mathrm{e}^{-x} + |v|^2\left(\tfrac{1}{4} + \tfrac{3}{2}\mathrm{e}^{-x}\right) + \overline{g}\hat{g}' \right) dx \\
&= \tfrac{1}{2}|g(\infty)|^2 + \operatorname{Re} \int_0^\infty \left(|v'|^2 + |v|^2\left(\tfrac{1}{4} + \tfrac{3}{2}\mathrm{e}^{-x}\right) + v'\overline{v}\mathrm{e}^{-x} \right) dx \\
&= \tfrac{1}{2}|g(\infty)|^2 + \int_0^\infty \left(|v'|^2 + |v|^2\left(\tfrac{1}{4} + 2\mathrm{e}^{-x}\right) \right) dx \\
&\geq \tfrac{1}{4}\|v\|_{1,2}^2.
\end{aligned}$$

Therefore, $\mathfrak{F}$ is a sectorial form on $\mathcal{V}$ ($M_3 = 1/4$, $M_1 = 1$, $a = 0$). Lemma 2.4.5 implies that the linear operator associated with $\mathfrak{F}$ is the operator A, given by (6.48), with the domain $\mathcal{D}(A) = W^2(0,\infty) \cap W_0^1(0,\infty)$. *Theorem 2.8.2 implies that A is a sectorial operator and that its spectrum is contained in the half-plane* $\operatorname{Re} z \geq 1/4$.

Integration by parts implies that we can express $\mathfrak{F}(v,w) - \overline{\mathfrak{F}(w,v)}$ as

$$\int_0^\infty \left(2v'\overline{w}e^{-x} - v\overline{w}e^{-x} + \overline{w}e^{-x/2} \int_0^x e^{-s/2} v(s)ds - ve^{-x/2} \int_0^x e^{-s/2}\overline{w(s)}ds \right) dx$$

which implies that

$$|\mathfrak{F}(v,w) - \overline{\mathfrak{F}(w,v)}| \leq 4\|w\|_2\|v\|_{1,2} \quad \text{for all} \quad w,v \in \mathcal{V}.$$

Therefore Theorems 2.8.6, 2.8.12 and Corollary 6.1.14 imply that $\mathcal{D}(A^{1/2}) = \mathcal{V}$ and that for some $c \in (0,\infty)$ we have

$$c^{-1}\|v\|_{1,2} \leq \|v\|_{1/2} \leq c\|v\|_{1,2} \quad \text{for} \quad v \in \mathcal{V}.$$

If $v \in V$, then

$$|v(x)|^2 = 2\operatorname{Re} \int_0^x v'\overline{v} \leq 2\|v\|_2\|v'\|_2 \leq \|v\|_{1,2}^2.$$

Hence $\|v\|_\infty \leq \|v\|_{1,2}$ which implies that the nonlinearity $f : \mathcal{V} \to \mathcal{H}$, given by (6.49), is well defined

$$\|f(v)\|_2 \leq 3\|v\|_{1,2}\|v\|_2 \leq 3c\|v\|_{1/2}^2 \quad \text{for} \quad v \in \mathcal{V}$$

and that

$$\|f(w) - f(v)\|_2 \leq 3\|w - v\|_{1,2}(\|w\|_{1,2} + \|v\|_{1,2}) \quad \text{for} \quad w,v \in \mathcal{V},$$

implying that f is locally Lipschitz in $\mathcal{V}$. Thus, *(6.47) is a semilinear parabolic equation ($X = \mathcal{H}$, $\alpha = 1/2$, $\mathcal{U} = \mathcal{V}$) and therefore Theorem 6.4.11 implies stability of the 0 solution of (6.47) and hence stability of Crane's solution.*

6.7 Example: A Classical Solution

We shall now study, in more detail, the following initial value problem:

$$u_t(x,t) = u_{xx}(x,t) + g(u(x,t)) \quad \text{for} \quad 0 \leq x \leq \ell,\ 0 < t < \tau, \tag{6.50}$$

$$u(0,t) = u(\ell,t) = 0 \quad \text{for} \quad 0 \leq t < \tau, \tag{6.51}$$

$$u(x,0) = u_0(x) \quad \text{for} \quad 0 \leq x \leq \ell, \tag{6.52}$$

while assuming that $\ell \in (0,\infty)$, $u_0 \in W_0^1(0,\ell)$ is real valued and $g \in C^1(\mathbb{R},\mathbb{R})$.

In $X \equiv L^2(0,\ell)$ define the linear operator A by

$$Av = -v'' \quad \text{for} \quad v \in \mathcal{D}(A) \equiv W^2(0,\ell) \cap W_0^1(0,\ell).$$

Choose $\alpha \in (1/4, 1/2)$ and define $\mathcal{U} = \mathcal{D}(A^\alpha) = X^\alpha$. In view of Exercise 7, there exists $c < \infty$ such that $\|v\|_\infty \le c\|v\|_\alpha$ for $v \in \mathcal{U}$. Define $F(v) = g(\operatorname{Re} v)$ for $v \in \mathcal{U}$ and note that

$$\|F(v_2) - F(v_1)\| \le \|v_2 - v_1\| \max_{|r| \le mc} |g'(r)|$$

whenever $v_i \in \mathcal{U}$ and $\|v_i\|_\alpha < m$. Note also that $W_0^1(0,\ell) = \mathcal{D}(A^{1/2}) \subset \mathcal{U}$ and that $\|v\|_{1/2} \equiv \|A^{1/2}v\| = \|v'\|$ (Exercise 7). Thus, all assumptions of Section 6.4 are satisfied. Hence Theorem 6.4.10 implies existence of the solution $u \in \mathcal{S}(\tau)$ on some maximal interval $[0,\tau)$. The solution is unique and real valued (Theorems 6.4.6, 6.4.7).

u is actually the solution of the abstract problem

$$u' + Au = F(u), \quad u(0) = u_0.$$

It does satisfy the original PDE with a proper generalization of derivatives. In many applications we do not really need to elaborate further on this point. However, there are many important techniques that require more precise pointwise estimates of solutions. Use of a maximum principle, described below, is an example of such a technique. Let us demonstrate, through this example, how to obtain pointwise regularity of the solution.

Since Theorem 6.4.3 implies that $u(t) \in W_0^1(0,\ell)$, we may assume that $u(\cdot,t)$ is continuous on $[0,\ell]$ for all $t \in [0,\tau)$.

Theorem 6.7.1 *$u \in C([0,\ell] \times [0,\tau), \mathbb{R})$ and $u_t, u_{xx} \in C([0,\ell] \times (0,\tau), \mathbb{R})$. Moreover (6.50), (6.51) and (6.52) are satisfied pointwise.*

PROOF Choose $0 \le x \le \ell$, $0 \le t < \tau$ and $\varepsilon > 0$. Theorem 6.4.3 implies that $u \in C([0,\tau), X^{1/2})$. Hence $u \in C([0,\tau), C_B(0,\ell))$ (Exercise 7) and therefore there exists $\delta > 0$ such that

$$\max_{0 \le r \le \ell} |u(r,t) - u(r,s)| < \varepsilon/2 \quad \text{if} \quad s \in [0,\tau),\ |t-s| < \delta.$$

This implies

$$\begin{aligned} |u(x,t) - u(y,s)| &\le |u(x,t) - u(y,t)| + |u(y,t) - u(y,s)| \\ &< |u(x,t) - u(y,t)| + \varepsilon/2. \end{aligned}$$

Hence the continuity of $u(\cdot,t)$ implies that $u \in C([0,\ell] \times [0,\tau), \mathbb{R})$.

Theorem 6.4.3 implies that $u : (0,\tau) \to X^\eta$ is differentiable for $\eta \in [0, 1-\alpha)$. In particular, we may choose $\eta = 1/2$. This implies (Exercise 7) that

$$\lim_{h\to 0} \max_{0\le r\le \ell} \left| \frac{u(r,t+h) - u(r,t)}{h} - u'(r,t) \right| = 0$$

and hence u_t exists on $[0,\ell]\times(0,\tau)$ in the classical sense. Since we have that $u_t \in C((0,\tau), X^{1/2})$, we can use exactly the same argument used to show continuity of u to show that $u_t \in C([0,\ell] \times (0,\tau), \mathbb{R})$.

Observation $Au = g(u) - u_t \in C([0,\ell] \times [0,\tau), \mathbb{R})$ implies the rest. □

One can learn a lot about second order PDEs by pursuing the consequences of a solution having a local maximum or a local minimum. There are many recipes about how to do this, all of which usually go under the name **maximum principle**. One version is represented by Theorem 6.7.2. Note that, in (2) of the Theorem, it is assumed that only the indicated partial derivatives exist at the specified points and that f is an arbitrary real valued function satisfying (5). If $f_r(x,t,r,0)$ is bounded from above for $x \in (0,\ell)$, $t \in (0,T)$ and $r \in (-\varepsilon, 0]$, then (5) is satisfied provided that $f(x,t,0,0) \ge 0$.

Theorem 6.7.2 *Suppose*

(1) $\ell, T \in (0,\infty)$, $v \in C([0,\ell] \times [0,T], \mathbb{R})$, $D : (0,\ell) \times (0,T) \to [0,\infty)$

(2) $v_t(x,t) \ge D(x,t)v_{xx}(x,t) + f(x,t,v(x,t),v_x(x,t))$ *for* $0 < x < \ell$, $0 < t < T$

(3) $v(0,t) \ge 0$ *and* $v(\ell,t) \ge 0$ *for* $t \in [0,T]$

(4) $v(x,0) \ge 0$ *for* $x \in [0,\ell]$

(5) for some $\varepsilon > 0$, $L \in [0,\infty)$ *we have that* $f(x,t,r,0) > Lr$ *for* $r \in (-\varepsilon, 0)$, $x \in (0,\ell)$ *and* $t \in (0,T)$.

Then, $v(x,t) \ge 0$ *for all* $x \in [0,\ell]$ *and* $t \in [0,T]$.

PROOF Let us consider the case $L = 0$ first. Assume that $v(x_0,t_0) < 0$ for some $x_0 \in [0,\ell]$ and $t_0 \in [0,T]$. Define

$$M(t) = \min_{0\le x\le \ell, 0\le s\le t} v(x,s) \quad \text{for} \quad t \in [0,T].$$

Note that M is continuous, nonincreasing, $M(0) \ge 0$ and $M(t_0) < 0$. Hence we can choose $t_1 \in (0,T)$ so that $M(t_1) \in (-\varepsilon, 0)$. Pick now $x_2 \in (0,\ell)$ and $t_2 \in (0,t_1]$ so that $v(x_2,t_2) = M(t_1)$. Since $v(\cdot,t_2)$ has a minimum at x_2 we must have $v_{xx}(x_2,t_2) \ge 0$. Hence $v_t(x_2,t_2) \ge f(x_2,t_2,v(x_2,t_2),0) > 0$ which

implies that $v(x_2,t_3) < v(x_2,t_2)$ for some $t_3 < t_2$. Therefore, $v(x_2,t_3) < M(t_1)$ which is a contradiction. This proves the Theorem when $L = 0$.

In general, define $w(x,t) = v(x,t)e^{-Lt}$ and note that

$$w_t(x,t) \geq Dw_{xx}(x,t) + f_1(x,t,w(x,t),w_x(x,t)) \quad \text{for} \quad 0 < x < \ell,\ 0 < t < T$$

where

$$f_1(x,t,r,s) = (f(x,t,re^{Lt},se^{Lt}) - Lre^{Lt})e^{-Lt}.$$

Since $f_1(x,t,r,0) > 0$ if $r \in (-\varepsilon e^{-LT},0)$ this reduces the problem to the case $L = 0$. □

The proof of the following Corollary shows that uniqueness of the classical solution of (6.50) - (6.52) follows immediately from the above Theorem 6.7.2.

Corollary 6.7.3 *If $T \in (0,\tau)$ and $w \in C([0,\ell] \times [0,T],\mathbb{R})$ is such that*

$$w_t(x,t) = w_{xx}(x,t) + g(w(x,t)) \quad \textit{for} \quad 0 < x < \ell,\ 0 < t < T,$$

$$w(0,t) = w(\ell,t) = 0 \quad \textit{for} \quad 0 \leq t \leq T,$$

$$w(x,0) = u_0(x) \quad \textit{for} \quad 0 \leq x \leq \ell,$$

then $w = u$ on $[0,\ell] \times [0,T]$.

PROOF If $v = u - w$ and $f(x,t,v) = g(w(x,t) + v) - g(w(x,t))$, then the assumptions of Theorem 6.7.2 are satisfied. Hence $u - w \geq 0$.

If $v = w - u$ and $f(x,t,v) = g(u(x,t) + v) - g(u(x,t))$, then the assumptions of Theorem 6.7.2 are satisfied. Hence $w - u \geq 0$. □

An easy exercise, left to the reader, gives

Corollary 6.7.4 *If $a \in (-\infty,0]$, $g(a) \geq 0$ and $u_0 \geq a$, then $u \geq a$.*

Corollary 6.7.5 *If $b \in [0,\infty)$, $g(b) \leq 0$ and $u_0 \leq b$, then $u \leq b$.*

6.8 Dynamical Systems

Let M be a complete nonempty metric space with metric d and suppose that a family of maps $R_t \in C(M,M)$, $t \geq 0$, satisfies

(a) R_0 is the identity map

(b) $R_t(R_s(x)) = R_{t+s}(x)$ for all $x \in M$ and $t,s \geq 0$

(c) the map $t \to R_t(x)$ belongs to $C([0,\infty), M)$ for every $x \in M$.

Such families of maps are called **dynamical systems** on M.

A strongly continuous semigroup of linear operators on a Banach space is an example of a dynamical system. Here is a nonlinear example:

EXAMPLE 6.8.1 Suppose that $g \in C^1(\mathbb{R}, \mathbb{R})$ is such that $g(a) \geq 0$ and $g(b) \leq 0$ for some $-\infty < a < 0 < b < \infty$. Let M be the set of all $v \in W_0^1(0,1)$ such that $a \leq v(x) \leq b$ for all $x \in [0,1]$. M is a complete metric space with metric given by the $W_0^1(0,1)$ norm. Corollaries 6.7.4 and 6.7.5 and Theorems 6.4.10 and 6.7.1 imply that for each $u_0 \in M$ there exists $u \in C([0,\infty), W_0^1(0,1))$ such that

$$u_t(x,t) = u_{xx}(x,t) + g(u(x,t)) \quad \text{for} \quad 0 \leq x \leq 1,\ 0 < t < \infty$$

$$u(0,t) = u(1,t) = 0 \quad \text{for} \quad 0 \leq t < \infty,$$

$$u(x,0) = u_0(x) \quad \text{for} \quad 0 \leq x \leq 1.$$

Define $R_t(u_0) = u(\cdot, t)$. Continuous dependence Theorem 6.4.9 implies that $R_t \in C(M,M)$ for $t \geq 0$. Uniqueness Theorem 6.4.6 implies (b) and hence R_t, $t \geq 0$, is a dynamical system on M.

We shall now introduce some basic notions used to study dynamical systems. The **orbit** of $x \in M$ is defined by

$$\gamma(x) = \{R_t(x) \mid t \geq 0\}.$$

x is said to be an **equilibrium point** if $\gamma(x) = \{x\}$. Define $\omega(x)$, the **ω-limit set** for $x \in M$, to be the set of those $y \in M$ for which there exist $t_n \in (0,\infty)$, $n \geq 1$, such that

$$\lim_{n\to\infty} t_n = \infty \quad \text{and} \quad \lim_{n\to\infty} R_{t_n}(x) = y.$$

A set $A \subset M$ is said to be **positively invariant** if $R_t(A) \subset A$ for $t \geq 0$. A set $K \subset M$ is said to be **invariant** if for each $x \in K$ there exists $u \in C(\mathbb{R}, K)$ such that

$$u(0) = x \quad \text{and} \quad R_t(u(s)) = u(t+s) \quad \text{for} \quad t \geq 0,\ s \in \mathbb{R}.$$

Theorem 6.8.2 *If the orbit of $x \in M$ is relatively compact, then $\omega(x)$ is nonempty, compact, connected, invariant and $\lim_{t\to\infty} \operatorname{dist}(R_t(x), \omega(x)) = 0$.*

PROOF Relative compactness of $\gamma(x)$ implies that $\omega(u_0)$ is not empty.

Suppose $z_n \in \omega(x)$ for $n \geq 1$. Pick $t_n > n$ such that $d(R_{t_n}(x), z_n) < 1/n$. Relative compactness of $\gamma(x)$ implies that a subsequence $\{R_{t_{n_i}}(x)\}$ converges to some $z \in \omega(x)$. Clearly, $z_{n_i} \to z$. This proves compactness of $\omega(x)$.

Suppose there exist disjoint open sets A, B in M such that both $\omega(x) \cap A$ and $\omega(x) \cap B$ are nonempty and that $\omega(x) \subset A \cup B$. Then, for each $n \geq 1$ we

can find $t_{2n-1} > n$ and $t_{2n} > t_{2n-1}$ such that $R_{t_{2n-1}}(x) \in A$ and $R_{t_{2n}}(x) \in B$. Continuity of $R_\cdot(x)$ implies that $\{R_t(x)\}_{t_{2n-1} \le t \le t_{2n}}$ is connected. Hence there exists $s_n \in (t_{2n-1}, t_{2n})$ such that $R_{s_n}(x) \notin A \cup B$. Relative compactness of $\gamma(x)$ implies that a subsequence of $\{R_{s_n}(x)\}_{n \ge 1}$ converges to some y. Now, this y should belong to both the closed set $(A \cup B)^c$ and $\omega(x)$ - which is not possible. So, such sets A, B cannot exist and therefore $\omega(x)$ is connected.

Pick $y_0 \in \omega(x)$ and a sequence $\{t_n\}_{n \ge 1}$ such that $t_n \to \infty$ and $R_{t_n}(x) \to y_0$ as $n \to \infty$. Let $A_0 = \mathbb{N}$. For $m \ge 1$ let A_m be an infinite subset of A_{m-1} such that $\{R_{t_n - m}(x)\}_{n \in A_m}$ converges to some $y_m \in \omega(x)$. Let $n_0 = 1$ and for $m \ge 1$ pick $n_m \in A_m$ so that $n_m > n_{m-1}$. Note that $n_i \in A_m$ for $i \ge m$ and hence

$$y_m = \lim_{i \to \infty} R_{t_{n_i} - m}(x) \quad \text{for all} \quad m \ge 0.$$

Continuity of R_{t+m}, R_{t+k} and the semigroup property imply

$$R_{t+m}(y_m) = \lim_{i \to \infty} R_{t+m}(R_{t_{n_i} - m}(x)) = \lim_{i \to \infty} R_{t+k}(R_{t_{n_i} - k}(x)) = R_{t+k}(y_k)$$

whenever $t \ge -m$ and $t \ge -k$. This enables us to define $u \in C(\mathbb{R}, \omega(x))$ by $u(t) = R_{t+m}(y_m)$ for $t \ge -m$, $m \ge 0$. This u has the required properties.

If $R_t(x)$ would not converge to $\omega(x)$, then there would exist $c > 0$ such that for each n there exists $t_n > n$ for which $\operatorname{dist}(R_{t_n}(x), \omega(x)) > c$. This is not possible because relative compactness of $\gamma(x)$ would then imply that $\{R_{t_n}(x)\}$ would have a subsequence converging to some $y \in \omega(x)$. Therefore, $R_t(x)$ converges to $\omega(x)$ as $t \to \infty$. □

We say that $V \in C(M, \mathbb{R})$ is a **Liapunov function** for R_t, $t \ge 0$, if

(i) the mapping $t \to V(R_t(x))$ is nonincreasing on $[0, \infty)$ for every $x \in M$

(ii) $\inf_{t \ge 0} V(R_t(x)) > -\infty$ for each $x \in M$

(iii) $x \in M$ is such that $V(R_t(x)) = V(x)$ for $t \ge 0$ then x is an equilibrium point.

The conditions ii) and iii) are sometimes omitted in the literature. However, one can sometimes find additional conditions attached.

Theorem 6.8.3 *If R_t, $t \ge 0$, has a Liapunov function, then every nonempty ω-limit set consists of equilibrium points only.*

PROOF Let $\omega(x)$ be nonempty and let V be the Liapunov function. Note that there exists $c \in \mathbb{R}$ such that $V(y) = c$ for all $y \in \omega(x)$. Pick $y_0 \in \omega(x)$ and a sequence $\{t_n\}_{n \ge 1}$ such that $t_n \to \infty$ and $R_{t_n}(x) \to y_0$ as $n \to \infty$. Since $R_{t+t_n}(x) = R_t(R_{t_n}(x)) \to R_t(y_0)$, we see that $R_t(y_0) \in \omega(x)$ and hence $V(R_t(y_0)) = c$ for all $t \ge 0$. Thus, y_0 is an equilibrium point. □

$R_t \in C(M,M)$, $t \geq 0$, is said to be a **gradient system** on M if it has a Liapunov function and if the orbit of every $x \in M$ is relatively compact.

Various evolution equations generate gradient systems. Webb [32] has shown that a damped nonlinear wave equation is one such equation. A detailed analysis of this and various other gradient systems has been done by Hale in [10]. Here we shall now give an example of a semilinear parabolic equation that generates a gradient system.

6.9 Example: The Chafee-Infante Problem

We shall study the following PDE

$$u_t(x,t) = u_{xx}(x,t) + g(u(x,t)) \quad \text{for} \quad 0 \leq x \leq \ell,\ 0 < t < \tau,$$

$$u(0,t) = u(\ell,t) = 0 \quad \text{for} \quad 0 \leq t < \tau,$$

$$u(x,0) = u_0(x) \quad \text{for} \quad 0 \leq x \leq \ell,$$

while assuming that $u_0 \in W_0^1(0,\ell)$ is real valued, $g \in C^1(\mathbb{R},\mathbb{R})$ and that there exist $a \in [0, \pi^2\ell^{-2})$ and $b \in \mathbb{R}$ such that

$$G(x) \equiv 2\int_0^x g(s)ds \leq ax^2 + b \quad \text{for all} \quad x \in \mathbb{R}. \tag{6.53}$$

As in Section 6.7, we are going to set the PDE as a semilinear parabolic equation in $X \equiv L^2(0,\ell)$, with the linear operator A given by

$$Av = -v'' \quad \text{for} \quad v \in \mathcal{D}(A) \equiv W^2(0,\ell) \cap W_0^1(0,\ell),$$

$\alpha \in (1/4, 1/2)$ and $\mathcal{U} = \mathcal{D}(A^\alpha) = X^\alpha$. Theorem 6.4.10 implies existence of the solution $u \in \mathcal{S}(\tau)$ on some maximal interval $[0,\tau)$. The solution is unique and real valued (see Theorems 6.4.6, 6.4.7, 6.7.1 and Corollary 6.7.3).

For real valued $v \in W_0^1(0,\ell)$, define

$$W(v) = \int_0^\ell \big((v'(x))^2 - G(v(x))\big)\,dx. \tag{6.54}$$

Lemma 6.9.1 *$W(u(\cdot)) \in C([0,\tau),\mathbb{R})$ and $\dfrac{d}{dt}W(u(t)) = -2\|u'(t)\|^2$ for $t \in (0,\tau)$.*

PROOF Choose $\theta \in (0,\tau)$ and $t_1, t_2 \in [0,\theta]$.

Let $h(t) = (Au(t), u(t)) = \int_0^\ell (u_x(x,t))^2 dx$. Since A is self-adjoint we have

$$h(t_2) - h(t_1) = (u(t_2) - u(t_1), Au(t_2)) + (Au(t_1), u(t_2) - u(t_1)).$$

Hence the differentiability of u and the continuity of Au (Theorem 6.4.3) imply that $h' = 2(u', Au)$ on $(0,\tau)$. $u \in C([0,\tau), W_0^1(0,\ell))$ implies $h \in C([0,\tau),\mathbb{R})$.

Define $k(t) = \int_0^\ell G(u(x,t))dx$ and note that $k(t_2) - k(t_1)$ equals

$$\int_0^\ell \int_0^1 2(u(x,t_2) - u(x,t_1))g(u(x,t_1) + \theta(u(x,t_2) - u(x,t_1)))d\theta\, dx.$$

Hence

$$k(t_2) - k(t_1) = 2(u(t_2) - u(t_1), g(u(t_1)) + etc).$$

To estimate the etc term define $c = \max_{0\le t\le\theta} \|u(t)\|_\infty$ and note that

$$|etc| \le |u(x,t_2) - u(x,t_1)| \max_{|s|\le c} |g'(s)|/2.$$

Thus, $k' = 2(u', g(u))$ on $(0,\tau)$ and $k \in C([0,\tau),\mathbb{R})$. □

Theorem 6.9.2 *$\tau = \infty$, $\sup_{t\ge0} \|u(t)\|_{1/2} < \infty$ and $\inf_{t\ge0} W(u(t)) \ge -b\ell$.*

PROOF (6.53) and Example 2.8.8 imply

$$\int_0^\ell G(u(x,t))dx \le b\ell + a\int_0^\ell u(x,t)^2dx \le b\ell + a\ell^2\pi^{-2}\int_0^\ell u_x(x,t)^2dx.$$

Hence

$$(1 - a\ell^2\pi^{-2})\int_0^\ell u_x(x,t)^2dx \le b\ell + W(u(t)) \quad \text{for} \quad t \in [0,\tau).$$

Since Lemma 6.9.1 implies that $W(u(t)) \le W(u_0)$, we have that $\|u(t)\|_{1/2}$ is bounded and hence $g(u(t))$ is bounded for $t \in [0,\tau)$. Theorem 6.4.10 implies that $\tau = \infty$. □

Thus, if we define $R_t(u_0) = u(t)$, then $R_t \in C(W_0^1(0,\ell), W_0^1(0,\ell))$, $t \ge 0$, is a dynamical system on $W_0^1(0,\ell)$. Moreover, W is a Liapunov function for R_t.

Lemma 6.9.3 *The orbit $\{u(t)\}_{t\ge0}$ is a relatively compact subset of $W_0^1(0,\ell)$.*

PROOF Choose $\beta \in (1/2, 1)$. Boundedness of $g(u)$ (Theorem 6.9.2), Theorem 6.1.15 and (6.15) imply that there exists $c < \infty$ such that

$$\|A^\beta u(t)\| \le ct^{-\beta}e^{-rt} + c\int_0^t (t-s)^{-\beta}e^{-r(t-s)}ds \quad \text{for} \quad t > 0$$

where $r \in (0, \pi^2\ell^{-2})$ (recall that $\sigma(A) \subset [\pi^2\ell^{-2}, \infty)$). Thus, $\{\|A^\beta u(t)\|\}_{t\ge1}$ is bounded. Compactness of A^{-1} (Example 2.8.3) implies that $\{A^{1/2-\beta}A^\beta u(t)\}_{t\ge1}$ is relatively compact in X (Exercise 5). Since the continuity implies that $\{A^{1/2}u(t)\}_{0\le t\le1}$ is compact in X, we are done. □

Therefore, R_t, $t \geq 0$ is a gradient system on $W_0^1(0,\ell)$ and Theorems 6.8.2 and 6.8.3 imply

Theorem 6.9.4 *$\omega(u_0)$, the ω-limit set for u_0, is a nonempty, compact, connected subset of $W_0^1(0,\ell)$. Moreover,*

$$\lim_{t\to\infty} \inf_{v\in\omega(u_0)} \|u(t)-v\|_{1/2} = 0$$

and if $v \in \omega(u_0)$, then $v \in \mathcal{D}(A)$ and $Av = g(v)$.

Thus, to study the limiting behavior of u we need to study solutions of the steady state equation

$$v''(x) + g(v(x)) = 0 \quad \text{for} \quad x \in [0,\ell], \quad v(0) = v(\ell) = 0. \tag{6.55}$$

Observe that if v satisfies (6.55) and if $v'(x_0) = 0$ at some $x_0 \in (0,\ell)$, then v is symmetric around x_0, i.e. if $w(x) = v(2x_0 - x)$ then $w = v$ near x_0 by the uniqueness of the initial value problem at x_0. Suppose now that the solution of (6.55) has a maximum $M > 0$. Choose the smallest $x_0 \in (0,\ell)$ such that $v(x_0) = M$. By continuity there should exist $x_1 \in (0,x_0)$ such that $v > 0$ on $[x_1,x_0]$ and, by symmetry, $v' > 0$ on $[x_1,x_0)$. Let x_2 be the infimum of such x_1. Note that $v(x_2) = 0$ and $v' > 0$ on (x_2,x_0). Multiplying (6.55) by v' and integrating gives

$$(v')^2 + G(v) = G(M). \tag{6.56}$$

Hence $v'(x) = \sqrt{G(M) - G(v(x))} > 0$ and

$$\int_0^{v(x)} \frac{ds}{\sqrt{G(M)-G(s)}} = x - x_2 \quad \text{for} \quad x \in (x_2,x_0).$$

By symmetry, this determines v on $[x_2, 2x_0 - x_2] \subset [0,\ell]$. Note that $v'(x_2) \geq 0$. When $v'(x_2) = 0$ the symmetry implies that v is periodic with period $2(x_0 - x_2)$. Hence $\ell = 2n(x_0 - x_2)$ for some $n \in \mathbb{N}$. If $x_2 > 0$ and $v'(x_2) > 0$, then $G(0) < G(M)$ by (6.56) and v has a negative arch which can be analyzed similarly. The point $2x_0 - x_2$ can also be analyzed similarly. Let us gather some of these observations in the following:

Theorem 6.9.5 *Suppose that v is a solution of (6.55) and*

$$M = \max_{0\leq x\leq \ell} v(x), \quad m = \min_{0\leq x\leq \ell} v(x).$$

If $M > 0$, then $G(s) < G(M)$ for $s \in (0,M)$ and $\ell_+(M) \equiv \displaystyle\int_0^M \frac{ds}{\sqrt{G(M)-G(s)}} < \infty$.

If $m < 0$, then $G(s) < G(m)$ for $s \in (m,0)$ and $\ell_-(m) \equiv \displaystyle\int_m^0 \frac{ds}{\sqrt{G(m)-G(s)}} < \infty$.

Moreover,

(1) If $M > 0$ and $m = 0$, then $\ell = 2n\ell_+(M)$ for some $n \in \mathbb{N}$.

(2) If $M = 0$ and $m < 0$, then $\ell = 2n\ell_-(m)$ for some $n \in \mathbb{N}$.

(3) If $M > 0$ and $m < 0$, then $G(M) = G(m) > G(0)$ and

$$\ell = 2n_+\ell_+(M) + 2n_-\ell_-(m) \text{ for some } n_\pm \in \mathbb{N} \text{ with } |n_+ - n_-| \leq 1.$$

EXAMPLE 6.9.6 Let us take $g(x) = \lambda \sin x$ where $\lambda > 0$. Theorem 6.9.5 implies that any solution v of (6.55) satisfies $-\pi < v(x) < \pi$ for $x \in [0, \ell]$. Note that

$$\ell_+(A) = \ell_-(-A) = \int_0^A \frac{ds}{\sqrt{2\lambda(\cos s - \cos A)}} \quad \text{for} \quad A \in (0, \pi).$$

A change of variable $1 - \cos s = (1 - \cos A)\cos^2 \varphi$ gives that

$$\ell_+(A) = \frac{1}{\sqrt{\lambda}} \int_0^{\pi/2} \frac{d\varphi}{\sqrt{1 - \sin^2(A/2)\cos^2 \varphi}} \quad \text{for} \quad A \in (0, \pi).$$

Hence, ℓ_+ increases form $\pi/\sqrt{4\lambda}$ at $A = 0$ to $+\infty$ at $A = \pi$. If $v \not\equiv 0$ satisfies (6.55) and $A = \max_x |v(x)|$, then Theorem 6.9.5 implies that

$$\ell = 2n\ell_+(A) > n\pi/\sqrt{\lambda} \quad \text{for some} \quad n \in \mathbb{N}.$$

Let $N \geq 0$ be the largest integer such that $N < \ell\sqrt{\lambda}/\pi$. If $n \in \mathbb{N}$ and $n \leq N$, then one can find a unique $A \in (0, \pi)$ such that $\ell = 2n\ell_+(A)$ and one can construct, as outlined above, two corresponding solutions of (6.55): one with $v'(0) > 0$ and one with $v'(0) < 0$. Since 0 is also a solution, we see that *(6.55) has exactly $2N + 1$ solutions* for this g. Since ω-limit sets have to be connected (Theorem 6.9.4), we see that each consists of a single point.

One can make a conjecture, based on the above Theorem 6.9.5 and Example 6.9.6, that $\omega(u_0)$ always consists of a single point, which would imply that the u always tends to a solution of the steady state equation (6.55). This is actually a popular method among engineers for solving (6.55). This conjecture was proven under very mild assumptions on g. For more details and generalizations to higher dimensions see Henry [11] and Hale [10].

6.10 Exercises

1. Suppose that $T - \lambda$ is an m-accretive operator on a Hilbert space for some $\lambda > 0$. Show that $T^{1/2}$ can be defined via Definition 6.1.6. Show that $T^{1/2}$ is equal to S as given in Theorem 2.3.5.

2. For $A \in \mathfrak{G}(X)$, $\lambda \in \mathbb{R}$, $\alpha \in [0, 1)$ and $x \in \mathcal{D}(A^\alpha)$ show that

$$\lim_{t \to 0+} t^{-\alpha} \|x - e^{\lambda t} Q(t)x\| = 0.$$

3. Show that for every $A \in \mathfrak{G}(X)$, $x \in X$ and $\alpha \geq 0$ we have

$$\lim_{\lambda \to +\infty} \lambda^{\alpha}(A+\lambda)^{-\alpha}x = x.$$

4. Show that for $A \in \mathfrak{G}(X)$, $\alpha, \beta, \gamma \in \mathbb{R}$, $\alpha < \beta < \gamma$ there exists $c < \infty$ such that

$$\|A^{\beta}x\| \leq c\|A^{\gamma}x\|^{\frac{\beta-\alpha}{\gamma-\alpha}}\|A^{\alpha}x\|^{\frac{\gamma-\beta}{\gamma-\alpha}} \quad \text{for all} \quad x \in \mathcal{D}(A^{\gamma}).$$

5. Show that if $A \in \mathfrak{G}(X)$ and $A^{-\alpha}$ is compact for some $\alpha > 0$, then $A^{-\beta}$ is compact for all $\beta > 0$.

6. Suppose that $A - \delta$ is an m-accretive operator in a Hilbert space for some $\delta > 0$. Show that $A^{\alpha} - \delta^{\alpha}$ is an m-accretive operator for all $\alpha \in [0,1]$.

7. Let A be a linear operator in $L^2(0,\ell)$, $\ell \in (0,\infty)$, given by

$$Au = -u'' \quad \text{for} \quad u \in \mathcal{D}(A) \equiv W^2(0,\ell) \cap W_0^1(0,\ell).$$

Show that

(a) $\mathcal{D}(A^{1/2}) = W_0^1(0,\ell)$ and that $\|A^{1/2}u\| = \|u'\|$ for $u \in W_0^1(0,\ell)$

(b) for $\alpha > 1/4$ there exists $c < \infty$ such that $\|u\|_{\infty} \leq c\|A^{\alpha}u\|$ for $u \in \mathcal{D}(A^{\alpha})$.

8. Let A be a sectorial operator in a Banach space X, $\alpha \in (0,1)$ and $a < \operatorname{Re}\lambda$ for all $\lambda \in \sigma(A)$. Show that for each $u_0 \in X$ there exists a unique $u \in C([0,\infty), X)$ such that $u(0) = u_0$ and

$$u'(t) + Au(t) = (A-a)^{\alpha}u(t) \quad \text{for all} \quad t \in (0,\infty).$$

9. Verify the details in Example 6.4.1 and construct an actual example for each $0 < \alpha < 1$, $p > 1/\alpha$. (*Hint:* try using multiplication operators.)

10. Show that

$$u(x,t) = \left(c + \frac{\pi^2}{2}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{e^{-2(n^2+m^2)t}}{n^2+m^2}\right)^{-1/4}\sum_{k=1}^{\infty}\frac{\sin kx}{k}e^{-k^2t}$$

satisfies, for every $c \geq 0$,

$$u_t(x,t) = u_{xx}(x,t) + \left(\int_0^{\pi}|u_x(s,t)|^2 ds\right)^2 u(x,t) \qquad \text{for} \quad t > 0,\ 0 \leq x \leq \pi$$

$$u(0,t) = u(\pi,t) = 0 \qquad \text{for} \quad t \geq 0$$

$$\lim_{t \to 0+}\sup_{0 \leq x \leq \pi}|u(x,t)| = 0.$$

(*Hint:* evaluate the Fourier sine series of $\pi - x$ and see Exercise 16 in Chapter 4.) Show that the PDE can be set as a semilinear parabolic equation in $L^2(0,\pi)$, with $\alpha = 1/2$ (see also Exercise 7) and that the uniqueness fails in this case because

$$\int_0^1 \left(\int_0^\pi |u_x(x,t)|^2 dx \right)^2 dt = \infty.$$

11. Show that Galerkin approximations of a solution of the Navier-Stokes equations exist globally, for any real valued initial $u_0 \in X$, and that they satisfy (6.41) on $[0,\infty)$. Prove that a subsequence converges weakly in $L^2((0,\infty) \times \Omega)^3$.

12. Find the spectrum of the operator A_0 in $L^2(0,\infty)$ given by (6.45).

13. Prove Corollaries 6.7.4 and 6.7.5.

14. Suppose that g, in Section 6.7, is such that for some $a, b \in \mathbb{R}$ we have

$$g(x) \leq ax + b \quad \text{for} \quad x \geq 0$$

and $g(0) \geq 0$. Show that if $u_0 \geq 0$, then the solution of (6.50) - (6.52) exists for all time ($\tau = \infty$).

15. Let $R_t \in C(M,M)$, $t \geq 0$, be a dynamical system on M. Show that

$$\omega(x) = \cap_{\tau>0} \overline{\gamma(R_\tau(x))} \quad \text{for all} \quad x \in M.$$

16. Suppose that $u_0 \in C_B(\mathbb{R})$. Show that if

$$v(x,t) = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-s)^2}{4t} - \frac{1}{2}\int_0^s u_0(r)dr} ds$$

and $u = -2v_x/v$, then u solves Burger's equation:

$$u_t(x,t) + u(x,t)u_x(x,t) = u_{xx}(x,t) \quad \text{for} \quad x \in \mathbb{R},\ t > 0,$$

$$u(x,0) = u_0(x) \quad \text{for} \quad x \in \mathbb{R}.$$

This is known as Hopf-Cole substitution.

Bibliography

[1] R. A. ADAMS, *Sobolev Spaces*, Academic Press, 1975.

[2] S. N. CHOW AND K. LU, *Invariant manifolds for flows in Banach spaces*, J. Differential Equations, **74**(1988), 285–317.

[3] S. N. CHOW, D. R. DUNNINGER AND M. MIKLAVČIČ, *Galerkin Approximations for Singular Linear Elliptic and Semilinear Parabolic Problems*, Applicable Analysis, **40**(1990), 41–52.

[4] P. CONSTANTIN AND C. FOIAS, *Navier-Stokes Equation*, The University of Chicago Press, 1988.

[5] L. C. EVANS, *Partial Differential Equations*, Berkely Mathematics Lecture Notes, v. 3ab, 1994.

[6] C. FOIAS, G. R. SELL AND E. S. TITI, *Exponential tracking and approximation of inertial manifolds for dissipative nonlinear equations*, J. Dynamics and Diff. Eq., **1**(1989), 199–244.

[7] G. B. FOLLAND, *Fourier Analysis and Its Applications*, Wadsworth & Brooks/Cole, 1992.

[8] AVNER FRIEDMAN, *Partial Differential Equations*, Krieger, 1983.

[9] J. A. GOLDSTEIN, *Semigroups of Linear Operators and Applications*, Oxford University Press, 1985.

[10] J. HALE, *Asymptotic Behavior of Dissipative Systems*, Mathematical Surveys and Monographs, No. 25, AMS, 1988.

[11] D. HENRY, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Mathematics, No. 840, Springer, 1981.

[12] L. HÖRMANDER, *The Analysis of Linear Partial Differential Operators I*, Springer, 1983.

[13] T. KATO, *Perturbation Theory for Linear Operators*, 2nd edition, Springer, 1980.

[14] H. KOMATSU, *Fractional powers of operators*, Pacific J. Math. **19**(1966), 285–346.

[15] O. A. LADYZHENSKAYA, *The Boundary Value Problems of Mathematical Physics*, Springer, 1985.

[16] M. MIKLAVČIČ, *Nonlinear stability of asymptotic suction*, AMS Transactions, **281**(1984), 215–231.

[17] M. MIKLAVČIČ, *Stability for semilinear parabolic equations with noninvertible linear operator*, Pacific J. Math. **118**(1985), 199–214.

[18] M. MIKLAVČIČ, *Approximations for Weakly Nonlinear Evolution Equations*, Math. Computations, **53**(1989), 471–484.

[19] M. MIKLAVČIČ, *A Sharp Condition for Existence of an Inertial Manifold*, J. Dynamics and Diff. Eq., **3**(1991), 437–456.

[20] A. PAZY, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, 1983.

[21] M. H. PROTTER AND H. F. WEINBERGER, *Maximum Principles in Differential Equations*, Springer, 1984.

[22] J. RAUCH, *Partial Differential Equations*, Springer, 1991.

[23] M. REED AND B. SIMON, *Methods of Modern Mathematical Physics*, v. I-III, Academic Press, 1972.

[24] M. RENARDY AND R. C. ROGERS, *An Introduction to Partial Differential Equations*, Springer, 1993.

[25] W. RUDIN, *Functional Analysis*, McGraw Hill, 1973.

[26] W. RUDIN, *Real & Complex Analysis*, 3rd edition, McGraw Hill, 1987.

[27] M. SCHECHTER, *Principles of Functional Analysis*, Academic Press, 1971.

[28] M. SCHECHTER, *Modern Methods in Partial Differential Equations*, McGraw-Hill, 1977.

[29] R. E. SHOWALTER, *Hilbert Space Methods for Partial Differential Equations*, Pitman, London, 1977.

[30] R. TEMAM, *Navier-Stokes Equations and Nonlinear Functional Analysis*, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, 1983.

[31] R. TEMAM, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Springer-Verlag, New York, 1988.

[32] G. WEBB, *A Bifurcation Problem for a Nonlinear Hyperbolic Partial Differential Equation*, SIAM J. Math. Anal. **10**(1979), 922–932.

[33] J. WLOKA, *Partial Differential Equations*, Cambridge University Press, 1987.

[34] K. YOSIDA, *Functional Analysis*, 6th edition, Springer, 1980.

List of Symbols

$\mathbb{R}$ real numbers	
$\mathbb{C}$ complex numbers	
$\mathbb{K}$ scalars ($\mathbb{R}$ or $\mathbb{C}$)	
$\mathbb{Z} = \{\ldots, -1, 0, 1, 2, \ldots\}$	
$\mathbb{N} = \{1, 2, \ldots\}$	
$\emptyset$ empty set	
$\overline{A}$ closure or conjugate	
dist	1
span$\{\cdot\}$	7
$B(x, r)$ open ball	1
$C(M, N)$, $C_B(M, N)$	4
$C^\nu(M, N)$, $C_H(M, N)$	4
$C(M)$	7
$C_0(M)$, supp$(\cdot)$	7
$C^m(\Omega)$, $C^\infty(\Omega)$	7
$C_0^m(\Omega)$, $C_0^\infty(\Omega)$	8
$C_0^\infty(a, b)$	8
$C_B^m(\Omega)$, $\|\cdot\|_{m,\infty}$	9
$C_B(\Omega) = C_B^0(\Omega)$	
$C_u^m(\Omega) = C^m(\overline{\Omega})$	10
$C_u(\Omega) = C(\overline{\Omega}) = C_u^0(\Omega)$	
C_ℓ, $C_{\ell 0}$	10
L^p, L_{loc}^p, $\|\cdot\|_p$	11
ℓ^p, c_0	11
$\mathcal{D}(\cdot)$, $\mathcal{N}(\cdot)$, $\mathcal{R}(\cdot)$	12
$\mathfrak{B}(X, Y)$, $\mathfrak{B}(X)$	13
$AC[a, b]$	15
$X^\star = \mathfrak{B}(X, \mathbb{K})$	16
$T^\star$ adjoint of T	21
T^* Hilbert space adjoint	56
$\rho(T)$, $\sigma(T)$, $\sigma_p(T)$	23
Δ	92
$\overline{\Delta}$	185
Δ_ℓ	193
J_ν Bessel function	95
DCT	78
J_ε mollifier	110
$\mathcal{F}$ Fourier transform	115
$\mathcal{S}$ Schwartz space	114
$\mathcal{D}(\Omega)$, $\mathcal{D}'(\Omega)$	122
D^α, $D_i = \frac{\partial}{\partial x_i}$	7
D^α weak derivative	127
$\|\cdot\|_{m,p}$	133
$W_{loc}^{m,p}$, W_{loc}^m, $W^{m,p}$, W^m	135
$(\cdot,\cdot)_m$, $\|\cdot\|_{m,p}$	135
$L(S, X)$	172
$\mathfrak{A}(a, M, \theta, X)$	193
e^{-Az}	196
$\mathcal{S}_m(\tau)$	218
$\mathcal{S}(\tau)$	260
$\mathfrak{G}(X)$	247
A^α	249
X^α, $\|\cdot\|_\alpha$	254
$\omega(\cdot), \gamma(\cdot)$	277

Index

Abel's integral equation, **45**
accretive, **28**, **58**, 73, 185
addition, **6**
adjoint, **21**, 56, 63
adjoint problem, **40**
adjoint space, **16**
almost separably-valued, **171**
analytic, **27**, **117**, 169, 193
arcwise-connected, **5**
Arzela-Ascoli Theorem, 5, 17, 31, 34, 90, 94, 150

Baire Theorem, 2
Banach space, **9**
Banach-Steinhaus Theorem, 14
basis, **7**
Bessel function, **95**
biharmonic operator, **156**
Bochner integrable, **172**
Bochner integral, **172**
boundary value problem, 39, 120
bounded, **12**
bounded linear functionals, **16**, 55
bounded set, **9**
Burger's equation, 284

Cauchy sequence, **2**
Čebyšev polynomials, **52**
closable, **15**
closed, **2**, **15**
Closed Graph Theorem, 16
closure, **2**, **16**
compact, **2**, **31**
compact imbedding, 82, 83, 152, 156
compact resolvent, **37**, 74, 78, 82, 83, 87, 90, 91, 94, 154, 155, 157
complement, **2**
complete, **2**, **48**, 50, 74, 75, 78, 91, 92, 104, 163
complex vector space, **6**
cone property, **141**, 155
conjugate gradient method, **105**
conjugate space, **16**
connected, **5**, 130, 155
continuous, **4**
continuous at the point, **4**
contraction, **3**
Contraction Mapping Theorem, 3, 25, 216, 259
contraction semigroup, **184**
converge, **2**
convex, **7**
convolution, **109**, **123**
cycle, **206**

defined in, **12**
defined on, **12**
delay equation, 257
delta function, **123**
dense, **2**
densely defined, **12**
derivative of the distribution, **122**
derivative of the product, **8**
  weak, 131
dimension, **7**
Dirichlet problem, **153**
discretization, 224
dissipative, **28**

distance, **1**
distribution, **122**, 127
domain, **11**
dual space, **16**
dynamical systems, **277**

eigenvalues, **23**
eigenvectors, **23**
elliptic equation, **70**, 96, 102, 118, 153
energy method, 271
entire, **117**
equilibrium point, **277**
equivalent norms, **9**
Euler method, **183**, 224, 235
extension, **12**, 13, 18, 61, 99, 104

finite difference, 129, 234
finite element, **96**, 107, 236, 237, 239
fixed-point, **3**
formal adjoint, **63**
Fourier cosine series, **52**, 53
Fourier series, **52**, 75, 104
Fourier sine series, **52**, 79
Fourier transform, **114**, **115**, 120, 121, 129, 137, 159
fractional powers of operators, 58, 249
Fredholm alternative, **38**, 42, 57, 83, 156, 157
Friedrichs extension, **99**, 243
fundamental solution, **123**

Galerkin approximation, **98**, 99, 103, 236, 239, 243
Gårding's inequality, 103, 156, 157
generator, **168**
gradient system, **279**
Gram-Schmidt orthogonalization, **48**
Gronwall Theorem, 216, 258

Hahn-Banach Theorem, 18
Hamiltonian, 88, 187
harmonic functions, 120
hat function, **98**
Hausdorff Maximality Theorem, 18
heat equation, 222, 260, 267
Heisenberg inequality, **117**
Hermite functions, **91**, 163
Hilbert space, **48**
Hilbert space adjoint, **56**
Hilbert-Schmidt Theorem, 74
Hille-Yosida Theorem, 180
Hölder continuous, 4
Hopf-Cole substitution, 284
Hörmander Theorem, 70
hyperbolic equation, 66, 187

imbedding compact, 82, 83, 152, 156
infinitesimal generator, **168**
initial value problem, 120, 121
inner product, **47**
inner product space, **47**
instability, 201, 265
intermediate derivatives, 132
interpolation inequalities, 132, 135
invariant, **277**
invariant subspace, **204**
inverse Fourier transform, **115**

Laplace operator, **70**
Laplace transform, **116**
Laplacian, **70**, 92, 118, 153, 155, 163, 185, 193
Legendre polynomials, **48**, 104
Liapunov function, **278**, 280
linear operator from X to Y, **11**
linearly independent, **7**
Lipschitz continuous, **4**

m-accretive, **58**, 61, 73, 87, 186, 283
Malgrange-Ehrenpreis Theorem, 125
maximal totally ordered, **18**
maximum principle, **275**
Maxwell's equations, 164
metric, **1**
metric space, **1**
mild solution, **210**, 234, 242

mild solutions, 218
mollifier, **110**

Neumann problem, **155**, 165
norm, **8**
normalized, **47**
normalized tangent functional, **20**, 172
normed space, **8**
null space, **12**
numerical range, **28**, 80, 99, 192

ω-limit set, **277**
open, **2**
open ball, **1**
Open Mapping Theorem, 14
orbit, **277**
origin, **6**
orthogonal, **47**
orthogonal projection, **55**
orthonormal, **48**
orthonormal basis, **48**

parabolic equation, 66, 120, 169, 185, **192**, **200**, 222, 233, 236, 237, 243
Parseval equality, 50, 52
partially ordered, **18**
Partition of Unity Theorem, 110
perturbation of a generator, 186
Pettis Theorem, 171
point spectrum, **23**
Ponicare inequality, **138**, 154, 165
positive definite, 67
positively invariant, **277**
projection, 55, **205**
projection associated with the spectral set, **206**

range, **12**
rapidly decreasing functions, **114**
real subspace, **7**, 174, 205, 220, 263
real vector space, **6**
reflexive, **20**
regularity of weak solutions, 64, 160
relatively compact, **2**
resolvent, **23**
resolvent identity, **23**
resolvent set, **23**
restriction, **12**, 61
restriction of A to Y, 204
retarded functional differential equation, 257
retraction map, **217**
Riesz Lemma, 55

scalar, **6**
scalar field, **6**
scalar multiplication, **6**
scalar product, **47**
Schrödinger equation, 187
Schwartz space, **114**
second dual space, **20**
sectorial operator, **193**, 201, 253, 254, 258, 259, 283
sectorial sesquilinear forms, **80**, 89, 93, 97, 99, 153, 155–157, 186, 187, 193, 236, 239, 253, 273
self-adjoint, **57**, 71, 73–75, 78, 83, 84, 88, 91, 93, 94, 99, 105, 107, 154–156
semigroup, **167**
  analytic, **192**
semilinear parabolic equations, **260**
seminorm, **9**
separable, **2**, 5, 11, 17
simple functions, **173**
singular differential equation, 103
smoothing operator, 170, 201
span, **7**
spectral representation, 74
spectral set, **206**, 213
spectrum, **23**
square root, **58**, 87
stability, 201, 264, 271
stability condition, 225

star-shaped, **138**, 165
strong ellipticity condition, **156**, 157, 200
strongly continuous semigroup, **167**
strongly elliptic operator, **157**
strongly measurable, **171**
Sturm-Liouville problem, **76**
subspace, **7**
support, **7**
symmetric, **57**, 70
symmetric hyperbolic system, **66**

Taylor formula, **8**
totally ordered, **18**
Trotter Theorem, 226

uncertainty principle, **118**
uniform boundedness principle, **14**, 17
uniformly continuous, **4**
uniqueness, 260, 283

variational formulation, **85**
vector space over , **6**
Volterra integral equation, 26

wave equation, **68**, 121, **189**, 239, 243, 246, 279
  generalized, 170, **187**, 190
weak derivative, **127**
weak limit, **18**
weak solution, **63**
weakly bounded, **44**
weakly convergent, **18**
weakly measurable, **171**
weakly nonlinear evolution equations, **215**

Young Lemma, 112

zero vector, **6**