Abstract. This manuscript provides a brief introduction to Real and (linear and nonlinear) Functional Analysis. It covers basic Hilbert and Banach space theory as well as basic measure theory including Lebesgue spaces and the Fourier transform.

Keywords and phrases. Functional Analysis, Banach space, Hilbert space, Measure theory, Lebesgue spaces, Fourier transform, Mapping degree, fixed-point theorems, differential equations, Navier–Stokes equation.
Contents

Preface vii

Part 1. Functional Analysis 3

Chapter 0. Introduction 3

§0.1. Linear partial differential equations 3

Chapter 1. A first look at Banach and Hilbert spaces 9

§1.1. Warm up: Metric and topological spaces 9

§1.2. The Banach space of continuous functions 24

§1.3. The geometry of Hilbert spaces 32

§1.4. Completeness 38

§1.5. Bounded operators 39

§1.6. Sums and quotients of Banach spaces 44

§1.7. Spaces of continuous and differentiable functions 45

Chapter 2. Hilbert spaces 49

§2.1. Orthonormal bases 49

§2.2. The projection theorem and the Riesz lemma 55

§2.3. Operators defined via forms 57

§2.4. Orthogonal sums and tensor products 61

§2.5. Applications to Fourier series 63

Chapter 3. Compact operators 69

§3.1. Compact operators 69

Chapter 4. Unbounded operators 73

§4.1. The Hille-Yosida theorem 73

§4.2. Semigroups of operators 79

§4.3. Spectral theory 85

Chapter 5. Spectral theory 91

§5.1. Basic spectral theorems 91

§5.2. The spectral theorem 97

§5.3. Applications 103

Chapter 6. Applications 109

§6.1. Maximal dissipative operators 109

§6.2. The linear Schrödinger operator 115

§6.3. Self-Adjoint Extensions 121

§6.4. The Huygens principle 127

§6.5. Conservation laws 133

§6.6. Some applications of the spectral theory 139

Chapter 7. Perturbation theory 143

§7.1. One-parameter perturbations 143

§7.2. Spectral shift functions 151

§7.3. Applications 157

Chapter 8. Compact perturbations of the Laplacian 163

§8.1. The compactness of the Laplacian 163

§8.2. The compactness of the Laplacian on manifolds 171

§8.3. Applications 177

Chapter 9. The fundamental solution and its applications 183

§9.1. The fundamental solution 183

§9.2. The fundamental solution of the Laplacian 189

§9.3. The fundamental solution of the heat equation 195

§9.4. Applications 201

Chapter 10. Scattering theory 207

§10.1. The scattering matrix 207

§10.2. The scattering matrix for the Schrödinger operator 215

§10.3. Applications 221
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>§3.2.</td>
<td>The spectral theorem for compact symmetric operators</td>
<td>73</td>
</tr>
<tr>
<td>§3.3.</td>
<td>Applications to Sturm–Liouville operators</td>
<td>77</td>
</tr>
<tr>
<td>§3.4.</td>
<td>Estimating eigenvalues</td>
<td>80</td>
</tr>
<tr>
<td>Chapter 4</td>
<td>The main theorems about Banach spaces</td>
<td>85</td>
</tr>
<tr>
<td>§4.1.</td>
<td>The Baire theorem and its consequences</td>
<td>85</td>
</tr>
<tr>
<td>§4.2.</td>
<td>The Hahn–Banach theorem and its consequences</td>
<td>93</td>
</tr>
<tr>
<td>§4.3.</td>
<td>The adjoint operator</td>
<td>99</td>
</tr>
<tr>
<td>§4.4.</td>
<td>The geometric Hahn–Banach theorem</td>
<td>102</td>
</tr>
<tr>
<td>§4.5.</td>
<td>Weak convergence</td>
<td>105</td>
</tr>
<tr>
<td>§4.6.</td>
<td>Weak topologies</td>
<td>114</td>
</tr>
<tr>
<td>§4.7.</td>
<td>Beyond Banach spaces: Locally convex spaces</td>
<td>116</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>More on compact operators</td>
<td>125</td>
</tr>
<tr>
<td>§5.1.</td>
<td>Canonical form of compact operators</td>
<td>125</td>
</tr>
<tr>
<td>§5.2.</td>
<td>Hilbert–Schmidt and trace class operators</td>
<td>129</td>
</tr>
<tr>
<td>§5.3.</td>
<td>Fredholm theory for compact operators</td>
<td>134</td>
</tr>
<tr>
<td>Chapter 6</td>
<td>Bounded linear operators</td>
<td>137</td>
</tr>
<tr>
<td>§6.1.</td>
<td>Banach algebras</td>
<td>137</td>
</tr>
<tr>
<td>§6.2.</td>
<td>The $C^*$ algebra of operators and the spectral theorem</td>
<td>143</td>
</tr>
<tr>
<td>§6.3.</td>
<td>Spectral measures</td>
<td>146</td>
</tr>
<tr>
<td>§6.4.</td>
<td>The Stone–Weierstraß theorem</td>
<td>149</td>
</tr>
<tr>
<td>Part 2. Real Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chapter 7</td>
<td>Almost everything about Lebesgue integration</td>
<td>155</td>
</tr>
<tr>
<td>§7.1.</td>
<td>Borel measures in a nut shell</td>
<td>155</td>
</tr>
<tr>
<td>§7.2.</td>
<td>Extending a premeasure to a measure</td>
<td>163</td>
</tr>
<tr>
<td>§7.3.</td>
<td>Measurable functions</td>
<td>167</td>
</tr>
<tr>
<td>§7.4.</td>
<td>How wild are measurable objects</td>
<td>169</td>
</tr>
<tr>
<td>§7.5.</td>
<td>Integration — Sum me up, Henri</td>
<td>172</td>
</tr>
<tr>
<td>§7.6.</td>
<td>Product measures</td>
<td>179</td>
</tr>
<tr>
<td>§7.7.</td>
<td>Transformation of measures and integrals</td>
<td>182</td>
</tr>
<tr>
<td>§7.8.</td>
<td>Appendix: Transformation of Lebesgue–Stieltjes integrals</td>
<td>188</td>
</tr>
<tr>
<td>§7.9.</td>
<td>Appendix: The connection with the Riemann integral</td>
<td>192</td>
</tr>
<tr>
<td>Chapter 8</td>
<td>The Lebesgue spaces $L^p$</td>
<td>197</td>
</tr>
<tr>
<td>§8.1.</td>
<td>Functions almost everywhere</td>
<td>197</td>
</tr>
</tbody>
</table>
Contents

§8.2. Jensen ≤ Hölder ≤ Minkowski 199
§8.3. Nothing missing in $L^p$ 204
§8.4. Approximation by nicer functions 205
§8.5. Integral operators 210

Chapter 9. More measure theory 213
§9.1. Decomposition of measures 213
§9.2. Derivatives of measures 216
§9.3. Complex measures 222
§9.4. Appendix: Functions of bounded variation and absolutely continuous functions 228

Chapter 10. The dual of $L^p$ 235
§10.1. The dual of $L^p$, $p < \infty$ 235
§10.2. The dual of $L^\infty$ and the Riesz representation theorem 236

Chapter 11. The Fourier transform 241
§11.1. The Fourier transform on $L^1$ and $L^2$ 241
§11.2. Applications to linear partial differential equations 249
§11.3. Sobolev spaces 254
§11.4. Applications to evolution equations 258
§11.5. Tempered distributions 266

Chapter 12. Interpolation 275
§12.1. Interpolation and the Fourier transform on $L^p$ 275
§12.2. The Marcinkiewicz interpolation theorem 279

Part 3. Nonlinear Functional Analysis

Chapter 13. Analysis in Banach spaces 287
§13.1. Differentiation and integration in Banach spaces 287
§13.2. Contraction principles 291
§13.3. Ordinary differential equations 293

Chapter 14. The Brouwer mapping degree 297
§14.1. Introduction 297
§14.2. Definition of the mapping degree and the determinant formula 299
§14.3. Extension of the determinant formula 303
§14.4. The Brouwer fixed-point theorem 309
§14.5. Kakutani’s fixed-point theorem and applications to game theory 312
§14.6. Further properties of the degree 315
§14.7. The Jordan curve theorem 317

Chapter 15. The Leray–Schauder mapping degree 319
§15.1. The mapping degree on finite dimensional Banach spaces 319
§15.2. Compact maps 320
§15.3. The Leray–Schauder mapping degree 321
§15.4. The Leray–Schauder principle and the Schauder fixed-point theorem 323
§15.5. Applications to integral and differential equations 324

Chapter 16. The stationary Navier–Stokes equation 327
§16.1. Introduction and motivation 327
§16.2. An insert on Sobolev spaces 328
§16.3. Existence and uniqueness of solutions 333

Chapter 17. Monotone maps 337
§17.1. Monotone maps 337
§17.2. The nonlinear Lax–Milgram theorem 339
§17.3. The main theorem of monotone maps 341

Appendix A. Some set theory 343

Bibliography 351

Glossary of notation 353

Index 357
Preface

The present manuscript was written for my course *Functional Analysis* given at the University of Vienna in winter 2004 and 2009. It was adapted and extended for a course *Real Analysis* given in summer 2011. The last part are the notes for my course *Nonlinear Functional Analysis* held at the University of Vienna in Summer 1998 and 2001. The three parts are essentially independent. In particular, the first part does not assume any knowledge from measure theory (at the expense of not mentioning $L^p$ spaces).

It is updated whenever I find some errors and extended from time to time. Hence you might want to make sure that you have the most recent version, which is available from


Please do not redistribute this file or put a copy on your personal webpage but link to the page above.

*Acknowledgments*

I wish to thank my students, Kerstin Ammann, Phillip Bachler, Alexander Beigl, Peng Du, Christian Ekstrand, Damir Ferizović, Michael Fischer, Melanie Graf, Matthias Hammerl, Nobuya Kakehashi, Florian Kogelbauer, Helge Krüger, Reinhold Küstner, Oliver Leingang, Joris Mestdagh, Alice Mikikits-Leitner, Caroline Moosmüller, Piotr Owczarek, Martina Pflegpeter, Tobias Preinsterstorfer, Christian Schmid, Vincent Valmorin, Richard Welke, David Wimnesberger, Song Xiaojun, Rudolf Zeidler, and colleagues Nils C. Framstad, Heinz Hanßmann, Günther Hörmann, Aleksey Kostenko, Wallace Lam, Daniel Lenz, Johanna Michor, Viktor Qvarfordt, Alex Strohmaier, and Hendrik Vogt, who have pointed out several typos and made useful suggestions for improvements. I am also grateful to Volker Enß for making his
lecture notes on nonlinear Functional Analysis available to me.

Finally, no book is free of errors. So if you find one, or if you have comments or suggestions (no matter how small), please let me know.

Gerald Teschl

Vienna, Austria
March, 2015
Part 1

Functional Analysis
Chapter 0

Introduction

Functional analysis is an important tool in the investigation of all kind of problems in pure mathematics, physics, biology, economics, etc.. In fact, it is hard to find a branch in science where functional analysis is not used.

The main objects are (infinite dimensional) vector spaces with different concepts of convergence. The classical theory focuses on linear operators (i.e., functions) between these spaces but nonlinear operators are of course equally important. However, since one of the most important tools in investigating nonlinear mappings is linearization (differentiation), linear functional analysis will be our first topic in any case.

0.1. Linear partial differential equations

Rather than overwhelming you with a vast number of classical examples I want to focus on one: linear partial differential equations. We will use this example as a guide throughout our first chapter and will develop all necessary tools for a successful treatment of our particular problem.

In his investigation of heat conduction Fourier was lead to the (one dimensional) heat or diffusion equation

$$\frac{\partial}{\partial t} u(t, x) = \frac{\partial^2}{\partial x^2} u(t, x).$$

Here $u(t, x)$ is the temperature distribution at time $t$ at the point $x$. It is usually assumed, that the temperature at $x = 0$ and $x = 1$ is fixed, say $u(t, 0) = a$ and $u(t, 1) = b$. By considering $u(t, x) \to u(t, x) - a - (b - a)x$ it is clearly no restriction to assume $a = b = 0$. Moreover, the initial temperature distribution $u(0, x) = u_0(x)$ is assumed to be known as well.
Since finding the solution seems at first sight not possible, we could try to find at least some solutions of (0.1) first. We could, for example, make an ansatz for $u(t, x)$ as a product of two functions, each of which depends on only one variable, that is,

$$u(t, x) = w(t)y(x). \quad (0.2)$$

Accordingly, this ansatz is called separation of variables. Plugging everything into the heat equation and bringing all $t, x$ dependent terms to the left, right side, respectively, we obtain

$$\frac{\dot{w}(t)}{w(t)} = \frac{y''(x)}{y(x)}. \quad (0.3)$$

Here the dot refers to differentiation with respect to $t$ and the prime to differentiation with respect to $x$.

Now if this equation should hold for all $t$ and $x$, the quotients must be equal to a constant $-\lambda$ (we choose $-\lambda$ instead of $\lambda$ for convenience later on). That is, we are lead to the equations

$$-\dot{w}(t) = \lambda w(t) \quad (0.4)$$

and

$$-y''(x) = \lambda y(x), \quad y(0) = y(1) = 0, \quad (0.5)$$

which can easily be solved. The first one gives

$$w(t) = c_1 e^{-\lambda t} \quad (0.6)$$

and the second one

$$y(x) = c_2 \cos(\sqrt{\lambda}x) + c_3 \sin(\sqrt{\lambda}x). \quad (0.7)$$

However, $y(x)$ must also satisfy the boundary conditions $y(0) = y(1) = 0$. The first one $y(0) = 0$ is satisfied if $c_2 = 0$ and the second one yields ($c_3$ can be absorbed by $w(t)$)

$$\sin(\sqrt{\lambda}) = 0, \quad (0.8)$$

which holds if $\lambda = (\pi n)^2, \; n \in \mathbb{N}$. In summary, we obtain the solutions

$$u_n(t, x) = c_n e^{-(\pi n)^2 t} \sin(n\pi x), \quad n \in \mathbb{N}. \quad (0.9)$$

So we have found a large number of solutions, but we still have not dealt with our initial condition $u(0, x) = u_0(x)$. This can be done using the superposition principle which holds since our equation is linear. Hence any finite linear combination of the above solutions will again be a solution. Moreover, under suitable conditions on the coefficients we can even consider infinite linear combinations. In fact, choosing

$$u(t, x) = \sum_{n=1}^{\infty} c_n e^{-(\pi n)^2 t} \sin(n\pi x), \quad (0.10)$$
where the coefficients $c_n$ decay sufficiently fast, we obtain further solutions of our equation. Moreover, these solutions satisfy
\[ u(0, x) = \sum_{n=1}^{\infty} c_n \sin(n\pi x) \]  
(0.11)
and expanding the initial conditions into a Fourier series
\[ u_0(x) = \sum_{n=1}^{\infty} u_{0,n} \sin(n\pi x), \]  
(0.12)
we see that the solution of our original problem is given by (0.10) if we choose $c_n = u_{0,n}$.

Of course for this last statement to hold we need to ensure that the series in (0.10) converges and that we can interchange summation and differentiation. You are asked to do so in Problem 0.1.

In fact, many equations in physics can be solved in a similar way:

**Reaction-Diffusion equation:**
\[
\begin{align*}
\frac{\partial}{\partial t} u(t, x) - \frac{\partial^2}{\partial x^2} u(t, x) + q(x)u(t, x) &= 0, \\
u(0, x) &= u_0(x), \\
u(t, 0) &= u(t, 1) = 0.
\end{align*}
\]  
(0.13)
Here $u(t, x)$ could be the density of some gas in a pipe and $q(x) > 0$ describes that a certain amount per time is removed (e.g., by a chemical reaction).

**Wave equation:**
\[
\begin{align*}
\frac{\partial^2}{\partial t^2} u(t, x) - \frac{\partial^2}{\partial x^2} u(t, x) &= 0, \\
u(0, x) &= u_0(x), \\
\frac{\partial u}{\partial t}(0, x) &= v_0(x) \\
u(t, 0) &= u(t, 1) = 0.
\end{align*}
\]  
(0.14)
Here $u(t, x)$ is the displacement of a vibrating string which is fixed at $x = 0$ and $x = 1$. Since the equation is of second order in time, both the initial displacement $u_0(x)$ and the initial velocity $v_0(x)$ of the string need to be known.

**Schrödinger equation:**
\[
\begin{align*}
i \frac{\partial}{\partial t} u(t, x) &= -\frac{\partial^2}{\partial x^2} u(t, x) + q(x)u(t, x), \\
u(0, x) &= u_0(x), \\
u(t, 0) &= u(t, 1) = 0.
\end{align*}
\]  
(0.15)
Here $|u(t, x)|^2$ is the probability distribution of a particle trapped in a box $x \in [0, 1]$ and $q(x)$ is a given external potential which describes the forces acting on the particle.

All these problems (and many others) lead to the investigation of the following problem

$$Ly(x) = \lambda y(x), \quad L = -\frac{d^2}{dx^2} + q(x),$$

subject to the boundary conditions

$$y(a) = y(b) = 0.$$

Such a problem is called a Sturm–Liouville boundary value problem. Our example shows that we should prove the following facts about our Sturm–Liouville problems:

(i) The Sturm–Liouville problem has a countable number of eigenvalues $E_n$ with corresponding eigenfunctions $u_n(x)$, that is, $u_n(x)$ satisfies the boundary conditions and $Lu_n(x) = E_n u_n(x)$.

(ii) The eigenfunctions $u_n$ are complete, that is, any nice function $u(x)$ can be expanded into a generalized Fourier series

$$u(x) = \sum_{n=1}^{\infty} c_n u_n(x).$$

This problem is very similar to the eigenvalue problem of a matrix and we are looking for a generalization of the well-known fact that every symmetric matrix has an orthonormal basis of eigenvectors. However, our linear operator $L$ is now acting on some space of functions which is not finite dimensional and it is not at all clear what (e.g.) orthogonal should mean for functions. Moreover, since we need to handle infinite series, we need convergence and hence we need to define the distance of two functions as well.

Hence our program looks as follows:

- What is the distance of two functions? This automatically leads us to the problem of convergence and completeness.
- If we additionally require the concept of orthogonality, we are lead to Hilbert spaces which are the proper setting for our eigenvalue problem.
- Finally, the spectral theorem for compact symmetric operators will be the solution of our above problem.

**Problem 0.1.** Suppose $\sum_{n=1}^{\infty} |c_n| < \infty$. Show that (0.10) is continuous for $(t, x) \in [0, \infty) \times [0, 1]$ and solves the heat equation for $(t, x) \in (0, \infty) \times$
[0, 1]. (Hint: Weierstraß M-test. When can you interchange the order of summation and differentiation?)
A first look at Banach and Hilbert spaces

1.1. Warm up: Metric and topological spaces

Before we begin, I want to recall some basic facts from metric and topological spaces. I presume that you are familiar with these topics from your calculus course. As a general reference I can warmly recommend Kelly’s classical book [15] or the nice book by Kaplansky [14].

A metric space is a space $X$ together with a distance function $d : X \times X \to \mathbb{R}$ such that

(i) $d(x, y) \geq 0$,
(ii) $d(x, y) = 0$ if and only if $x = y$,
(iii) $d(x, y) = d(y, x)$,
(iv) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

If (ii) does not hold, $d$ is called a pseudometric. Moreover, it is straightforward to see the inverse triangle inequality (Problem 1.1)

$$|d(x, y) - d(z, y)| \leq d(x, z).$$

Example. Euclidean space $\mathbb{R}^n$ together with $d(x, y) = (\sum_{k=1}^{n} (x_k - y_k)^2)^{1/2}$ is a metric space and so is $\mathbb{C}^n$ together with $d(x, y) = (\sum_{k=1}^{n} |x_k - y_k|^2)^{1/2}$. 

The set

$$B_r(x) = \{y \in X | d(x, y) < r\}$$

is called an open ball around $x$ with radius $r > 0$. A point $x$ of some set $U$ is called an interior point of $U$ if $U$ contains some ball around $x$. If $x$
is an interior point of $U$, then $U$ is also called a **neighborhood** of $x$. A point $x$ is called a **limit point** of $U$ (also **accumulation** or **cluster point**) if $(B_r(x) \setminus \{x\}) \cap U \neq \emptyset$ for every ball around $x$. Note that a limit point $x$ need not lie in $U$, but $U$ must contain points arbitrarily close to $x$. A point $x$ is called an **isolated point** of $U$ if there exists a neighborhood of $x$ not containing any other points of $U$. A set which consists only of isolated points is called a **discrete set**. If any neighborhood of $x$ contains at least one point in $U$ and at least one point not in $U$, then $x$ is called a **boundary point** of $U$. The set of all boundary points of $U$ is called the **boundary** of $U$ and denoted by $\partial U$.

**Example.** Consider $\mathbb{R}$ with the usual metric and let $U = (-1,1)$. Then every point $x \in U$ is an interior point of $U$. The points $[-1,1]$ are limit points of $U$, and the points $\{-1,1\}$ are boundary points of $U$.  

A set consisting only of interior points is called **open**. The family of open sets $\mathcal{O}$ satisfies the properties

1. $\emptyset, X \in \mathcal{O}$,
2. $O_1, O_2 \in \mathcal{O}$ implies $O_1 \cap O_2 \in \mathcal{O}$,
3. $\{O_\alpha\} \subseteq \mathcal{O}$ implies $\bigcup_\alpha O_\alpha \in \mathcal{O}$.

That is, $\mathcal{O}$ is closed under finite intersections and arbitrary unions.

In general, a space $X$ together with a family of sets $\mathcal{O}$, the open sets, satisfying (i)–(iii), is called a **topological space**. The notions of interior point, limit point, and neighborhood carry over to topological spaces if we replace open ball by open set.

There are usually different choices for the topology. Two not too interesting examples are the **trivial topology** $\mathcal{O} = \{\emptyset, X\}$ and the **discrete topology** $\mathcal{O} = \mathcal{P}(X)$ (the powerset of $X$). Given two topologies $\mathcal{O}_1$ and $\mathcal{O}_2$ on $X$, $\mathcal{O}_1$ is called **weaker** (or **coarser**) than $\mathcal{O}_2$ if and only if $\mathcal{O}_1 \subseteq \mathcal{O}_2$.

**Example.** Note that different metrics can give rise to the same topology. For example, we can equip $\mathbb{R}^n$ (or $\mathbb{C}^n$) with the Euclidean distance $d(x,y)$ as before or we could also use

\[
\tilde{d}(x,y) = \sum_{k=1}^{n} |x_k - y_k|.
\] (1.3)

Then

\[
\frac{1}{\sqrt{n}} \sum_{k=1}^{n} |x_k| \leq \sqrt{\sum_{k=1}^{n} |x_k|^2} \leq \sum_{k=1}^{n} |x_k|
\] (1.4)

shows $B_{\sqrt{n}}(x) \subseteq \tilde{B}_r(x) \subseteq B_r(x)$, where $B, \tilde{B}$ are balls computed using $d, \tilde{d}$, respectively.  

\[\diamond\]
Example. We can always replace a metric $d$ by the bounded metric

$$
\tilde{d}(x, y) = \frac{d(x, y)}{1 + d(x, y)}
$$

without changing the topology (since the family of open balls does not change: $B_\delta(x) = B_{\delta/(1+\delta)}(x)$). To see that $\tilde{d}$ is again a metric, observe that $f(r) = \frac{r}{1+r}$ is concave and hence subadditive.

Every subspace $Y$ of a topological space $X$ becomes a topological space of its own if we call $O \subseteq Y$ open if there is some open set $\tilde{O} \subseteq X$ such that $O = \tilde{O} \cap Y$. This natural topology $\tilde{O} \cap Y$ is known as the relative topology (also subspace, trace or induced topology).

Example. The set $(0, 1] \subseteq \mathbb{R}$ is not open in the topology of $X = \mathbb{R}$, but it is open in the relative topology when considered as a subset of $Y = [-1, 1]$.

A family of open sets $\mathcal{B} \subseteq \mathcal{O}$ is called a base for the topology if for each $x$ and each neighborhood $U(x)$, there is some set $O \in \mathcal{B}$ with $x \in O \subseteq U(x)$. Since an open set $O$ is a neighborhood of every one of its points, it can be written as $O = \bigcup_{O \supseteq \tilde{O} \in \mathcal{B}} \tilde{O}$ and we have

**Lemma 1.1.** If $\mathcal{B} \subseteq \mathcal{O}$ is a base for the topology if and only if every open set can be written as a union of elements from $\mathcal{B}$.

**Proof.** To see the converse let $x, U(x)$ be given. Then $U(x)$ contains an open set $O$ containing $x$ which can be written as a union of elements from $\mathcal{B}$. One of the elements in this union must contain $x$ and this is the set we are looking for.

There is also a local version of the previous notions. A neighborhood base for a point $x$ is a collection of neighborhoods $\mathcal{B}(x)$ of $x$ such that for each neighborhood $U(x)$, there is some set $B \in \mathcal{B}(x)$ with $B \subseteq U(x)$. Note that the sets in a neighborhood base are not required to be open.

If every point has a countable neighborhood base, then $X$ is called first countable. If there exists a countable base, then $X$ is called second countable. Note that a second countable space is in particular first countable since for every base $\mathcal{B}$ the subset $\mathcal{B}(x) = \{O \in \mathcal{B} | x \in O\}$ is a neighborhood base for $x$.

Example. By construction, in a metric space the open balls $\{B_{1/m}(x)\}_{m \in \mathbb{N}}$ are a neighborhood base for $x$. Hence every metric space is first countable. Taking the union over all $x$, we obtain a base. In the case of $\mathbb{R}^n$ (or $\mathbb{C}^n$) it even suffices to take balls with rational center, and hence $\mathbb{R}^n$ (as well as $\mathbb{C}^n$) is second countable.
Given two topologies on $X$ their intersection will again be a topology on 
$X$. In fact, the intersection of an arbitrary collection of topologies is again a 
topology and hence given a collection $\mathcal{M}$ of subsets of $X$ we can define the 
topology generated by $\mathcal{M}$ as the smallest topology (i.e., the intersection of all 
topologies) containing $\mathcal{M}$. Note that if $\mathcal{M}$ is closed under finite intersections 
and $\emptyset, X \in \mathcal{M}$, then it will be a base for the topology generated by $\mathcal{M}$.

A topological space is called a **Hausdorff space** if for two different 
points there are always two disjoint neighborhoods.

**Example.** Any metric space is a Hausdorff space: Given two different 
points $x$ and $y$, the balls $B_{d/2}(x)$ and $B_{d/2}(y)$, where $d = d(x, y) > 0$, are 
disjoint neighborhoods (a pseudometric space will not be Hausdorff).

The complement of an open set is called a **closed set**. It follows from 
de Morgan’s rules that the family of closed sets $C$ satisfies

(i) $\emptyset, X \in C$,  
(ii) $C_1, C_2 \in C$ implies $C_1 \cup C_2 \in C$,  
(iii) $\{C_\alpha\} \subseteq C$ implies $\bigcap_\alpha C_\alpha \in C$.

That is, closed sets are closed under finite unions and arbitrary intersections.

The smallest closed set containing a given set $U$ is called the **closure**

$$
\overline{U} = \bigcap_{C \in \mathcal{C}, U \subseteq C} C, \quad (1.6)
$$

and the largest open set contained in a given set $U$ is called the **interior**

$$
U^\circ = \bigcup_{O \in \mathcal{O}, O \subseteq U} O. \quad (1.7)
$$

It is not hard to see that the closure satisfies the following axioms (**Kuratowski closure axioms**):

(i) $\overline{\emptyset} = \emptyset$,  
(ii) $U \subset \overline{U}$,  
(iii) $\overline{\overline{U}} = \overline{U}$,  
(iv) $\overline{U \cup V} = \overline{U} \cup \overline{V}$.

In fact, one can show that they can equivalently be used to define the topology 
by observing that the closed sets are precisely those which satisfy $\overline{A} = A$.

We can define interior and limit points as before by replacing the word 
ball by open set. Then it is straightforward to check

**Lemma 1.2.** Let $X$ be a topological space. Then the interior of $U$ is the set 
of all interior points of $U$, and the closure of $U$ is the union of $U$ with all 
limit points of $U$. 


1.1. Warm up: Metric and topological spaces

Example. The closed ball
\[ \overline{B}_r(x) = \{ y \in X | d(x, y) \leq r \} \]  \hspace{1cm} (1.8)
is a closed set (check that its complement is open). But in general we have only
\[ \overline{B}_r(x) \subseteq \overline{B}_r(x) \]  \hspace{1cm} (1.9)
since an isolated point \( y \) with \( d(x, y) = r \) will not be a limit point. In \( \mathbb{R}^n \)
or \( \mathbb{C}^n \) we have of course equality.

A sequence \( (x_n)_{n=1}^{\infty} \subseteq X \) is said to converge to some point \( x \in X \) if \( d(x, x_n) \to 0 \). We write \( \lim_{n \to \infty} x_n = x \) as usual in this case. Clearly the limit is unique if it exists (this is not true for a pseudometric).

Note that convergence can also be equivalently formulated in topological terms: A sequence \( x_n \) converges to \( x \) if and only if for every neighborhood \( U(x) \) of \( x \) there is some \( N \in \mathbb{N} \) such that \( x_n \in U(x) \) for \( n \geq N \). In a Hausdorff space the limit is unique. However, sequences usually do not suffice and in general definitions in terms of sequences are weaker. If one wants to get equivalent definitions one would need to use generalized sequences, so-called nets, where the index set \( \mathbb{N} \) is replaced by arbitrary directed sets.

Example. For example, we can call a set \( U \) sequentially closed if every convergent sequence from \( V \) also has its limit in \( U \). If \( U \) is closed then every point in the complement is an inner point of the complement thus no sequence from \( U \) can converge to such a point point. Hence every closed set is sequentially closed. In a metric space (or more generally in a first countable space) we can find a sequence for every limit point \( x \) by choosing a point (different from \( x \)) from every set in a neighborhood base. Hence the converse is also true in this case.

Every convergent sequence is a Cauchy sequence; that is, for every \( \varepsilon > 0 \) there is some \( N \in \mathbb{N} \) such that
\[ d(x_n, x_m) \leq \varepsilon, \quad n, m \geq N. \]  \hspace{1cm} (1.10)
If the converse is also true, that is, if every Cauchy sequence has a limit, then \( X \) is called complete. It is easy to see that a Cauchy sequence converges if and only if it has a convergent subsequence.

Example. Both \( \mathbb{R}^n \) and \( \mathbb{C}^n \) are complete metric spaces.

Example. The metric
\[ d(x, y) = |\arctan(x) - \arctan(y)| \]  \hspace{1cm} (1.11)
gives rise to the same topology on \( \mathbb{R} \) (since \( \arctan \) is bi-Lipschitz on every compact interval). However, \( x_n = n \) is a Cauchy sequence with respect to this metric but not with respect to the usual metric. Moreover, any sequence
1. A first look at Banach and Hilbert spaces

with $x_n \to \infty$ or $x_n \to -\infty$ will be Cauchy with respect to this metric and hence (show this) the for completion of $\mathbb{R}$ precisely the two new points $-\infty$ and $+\infty$ have to be added.

As noted before, in a metric space $x$ is a limit point of $U$ if and only if there exists a sequence $(x_n)_{n=1}^{\infty} \subseteq U \setminus \{x\}$ with $\lim_{n \to \infty} x_n = x$. Hence $U$ is closed if and only if for every convergent sequence the limit is in $U$. In particular,

**Lemma 1.3.** A closed subset of a complete metric space is again a complete metric space.

A set $U$ is called **dense** if its closure is all of $X$, that is, if $\overline{U} = X$. A space is called **separable** if it contains a countable dense set.

**Lemma 1.4.** A metric space is separable if and only if it is second countable as a topological space.

**Proof.** From every dense set we get a countable base by considering open balls with rational radii and centers in the dense set. Conversely, from every countable base we obtain a dense set by choosing an element from each set in the base. □

**Lemma 1.5.** Let $X$ be a separable metric space. Every subset $Y$ of $X$ is again separable.

**Proof.** Let $A = \{x_n\}_{n \in \mathbb{N}}$ be a dense set in $X$. The only problem is that $A \cap Y$ might contain no elements at all. However, some elements of $A$ must be at least arbitrarily close: Let $J \subseteq \mathbb{N}^2$ be the set of all pairs $(n,m)$ for which $B_{1/m}(x_n) \cap Y \neq \emptyset$ and choose some $y_{n,m} \in B_{1/m}(x_n) \cap Y$ for all $(n,m) \in J$. Then $B = \{y_{n,m}\}_{(n,m) \in J} \subseteq Y$ is countable. To see that $B$ is dense, choose $y \in Y$. Then there is some sequence $x_{n_k}$ with $d(x_{n_k}, y) < 1/k$. Hence $(n_k, k) \in J$ and $d(y_{n_k, k}, y) \leq d(y_{n_k, k}, x_{n_k}) + d(x_{n_k}, y) \leq 2/k \to 0$. □

Next, we come to functions $f : X \to Y$, $x \mapsto f(x)$. We use the usual conventions $f(U) = \{f(x) \mid x \in U\}$ for $U \subseteq X$ and $f^{-1}(V) = \{x \mid f(x) \in V\}$ for $V \subseteq Y$. The set $\text{Ran}(f) = f(X)$ is called the **range** of $f$, and $X$ is called the **domain** of $f$. A function $f$ is called **injective** if for each $y \in Y$ there is at most one $x \in X$ with $f(x) = y$ (i.e., $f^{-1}\{y\}$ contains at most one point) and **surjective** or **onto** if $\text{Ran}(f) = Y$. A function $f$ which is both injective and surjective is called **bijective**.

A function $f$ between metric spaces $X$ and $Y$ is called continuous at a point $x \in X$ if for every $\varepsilon > 0$ we can find a $\delta > 0$ such that

$$d_Y(f(x), f(y)) \leq \varepsilon \quad \text{if} \quad d_X(x, y) < \delta. \quad (1.12)$$

If $f$ is continuous at every point, it is called **continuous**.
Lemma 1.6. Let $X$ be a metric space. The following are equivalent:

(i) $f$ is continuous at $x$ (i.e., (1.12) holds).

(ii) $f(x_n) \to f(x)$ whenever $x_n \to x$.

(iii) For every neighborhood $V$ of $f(x)$, $f^{-1}(V)$ is a neighborhood of $x$.

Proof. (i) $\Rightarrow$ (ii) is obvious. (ii) $\Rightarrow$ (iii): If (iii) does not hold, there is a neighborhood $V$ of $f(x)$ such that $B_\delta(x) \not\subseteq f^{-1}(V)$ for every $\delta$. Hence we can choose a sequence $x_n \in B_{1/n}(x)$ such that $f(x_n) \not\in f^{-1}(V)$. Thus $x_n \to x$ but $f(x_n) \not\to f(x)$. (iii) $\Rightarrow$ (i): Choose $V = B_\varepsilon(f(x))$ and observe that by (iii), $B_\delta(x) \subseteq f^{-1}(V)$ for some $\delta$. □

The last item serves as a definition for topological spaces. In particular, it implies that $f$ is continuous if and only if the inverse image of every open set is again open (equivalently, the inverse image of every closed set is closed). If the image of every open set is open, then $f$ is called open. A bijection $f$ is called a homeomorphism if both $f$ and its inverse $f^{-1}$ are continuous. Note that if $f$ is a bijection, then $f^{-1}$ is continuous if and only if $f$ is open.

Again, in a general topological space continuity (iii) implies sequential continuity (ii) but the converse will not be true unless we assume (e.g.) that the space is first countable.

The support of a function $f : X \to \mathbb{C}^n$ is the closure of all points $x$ for which $f(x)$ does not vanish; that is,

$$\text{supp}(f) = \{x \in X | f(x) \neq 0\}. \quad (1.13)$$

If $X$ and $Y$ are metric spaces, then $X \times Y$ together with

$$d(\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle) = d_X(x_1, x_2) + d_Y(y_1, y_2) \quad (1.14)$$

is a metric space. A sequence $(x_n, y_n)$ converges to $(x, y)$ if and only if $x_n \to x$ and $y_n \to y$. In particular, the projections onto the first $(x, y) \mapsto x$, respectively, onto the second $(x, y) \mapsto y$, coordinate are continuous. Moreover, if $X$ and $Y$ are complete, so is $X \times Y$.

In particular, by the inverse triangle inequality (1.1),

$$|d(x_n, y_n) - d(x, y)| \leq d(x_n, x) + d(y_n, y), \quad (1.15)$$

we see that $d : X \times X \to \mathbb{R}$ is continuous.

Example. If we consider $\mathbb{R} \times \mathbb{R}$, we do not get the Euclidean distance of $\mathbb{R}^2$ unless we modify (1.14) as follows:

$$\tilde{d}(\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle) = \sqrt{d_X(x_1, x_2)^2 + d_Y(y_1, y_2)^2}. \quad (1.16)$$

As noted in our previous example, the topology (and thus also convergence/continuity) is independent of this choice. □
If $X$ and $Y$ are just topological spaces, the **product topology** is defined by calling $O \subseteq X \times Y$ open if for every point $(x, y) \in O$ there are open neighborhoods $U$ of $x$ and $V$ of $y$ such that $U \times V \subseteq O$. In other words, the products of open sets form a basis of the product topology. Again the projections onto the first and second component are continuous. In the case of metric spaces this clearly agrees with the topology defined via the product metric (1.14). There is also another way of constructing the product topology, namely, as the weakest topology which makes the projections continuous. In fact, this topology must contain all sets which are inverse images of open sets $U \subseteq X$, that is all sets of the form $U \times Y$ as well as all inverse images of open sets $V \subseteq Y$, that is all sets of the form $X \times V$. Adding finite intersections we obtain all sets of the form $U \times V$ and hence the same base as before. In particular, a sequence $(x_n, y_n)$ will converge if and only if both components converge.

Note that the product topology immediately extends to the product of an arbitrary number of spaces. In fact this is a special case of a more general construction which is often used. Let $\{f_\alpha\}_{\alpha \in A}$ be a collection of functions $f_\alpha : X \to Y_\alpha$, where $Y_\alpha$ are some topological spaces. Then we can equip $X$ with the weakest topology (known as the **initial topology**) which makes all $f_\alpha$ continuous. That is, we take the topology generated by sets of the forms $f_\alpha^{-1}(O_\alpha)$, where $O_\alpha \subseteq Y_\alpha$ is open. Finite intersections of such sets are hence a base for the topology and a sequence $x_n$ will converge to $x$ if and only if $f_\alpha(x_n) \to f_\alpha(x)$ for all $\alpha \in A$. It has the following characteristic property:

**Lemma 1.7.** Let $X$ have the initial topology from a collection of functions $\{f_\alpha\}_{\alpha \in A}$. A function $f : Z \to X$ is continuous (at $z$) if and only if $f_\alpha \circ f$ is continuous (at $z$) for all $\alpha \in A$.

**Proof.** If $f$ is continuous at $z$, then so is the composition $f_\alpha \circ f$. Conversely, let $U \subseteq X$ be a neighborhood of $f(z)$. Then $\bigcap_{j=1}^n f_\alpha^{-1}(O_\alpha_j) \subseteq U$ for some $\alpha_j$ and some open neighborhoods $O_\alpha_j$ of $f_\alpha_j(f(z))$. But then $f^{-1}(U)$ contains the neighborhood $f^{-1}(\bigcap_{j=1}^n f_\alpha_j^{-1}(O_\alpha_j)) = \bigcap_{j=1}^n (f_\alpha_j \circ f)^{-1}(O_\alpha_j)$ of $z$. \qed

A **cover** of a set $Y \subseteq X$ is a family of sets $\{U_\alpha\}$ such that $Y \subseteq \bigcup_\alpha U_\alpha$. A cover is called open if all $U_\alpha$ are open. Any subset of $\{U_\alpha\}$ which still covers $Y$ is called a **subcover**.

**Lemma 1.8** (Lindelöf). If $X$ is second countable, then every open cover has a countable subcover.

**Proof.** Let $\{U_\alpha\}$ be an open cover for $Y$, and let $\mathcal{B}$ be a countable base. Since every $U_\alpha$ can be written as a union of elements from $\mathcal{B}$, the set of all $B \in \mathcal{B}$ which satisfy $B \subseteq U_\alpha$ for some $\alpha$ form a countable open cover for $Y$. 

Moreover, for every $B_n$ in this set we can find an $\alpha_n$ such that $B_n \subseteq U_{\alpha_n}$. By construction, $\{U_{\alpha_n}\}$ is a countable subcover.

A refinement $\{V_\beta\}$ of a cover $\{U_\alpha\}$ is a cover such that for every $\beta$ there is some $\alpha$ with $V_\alpha \subseteq U_\alpha$. A cover is called locally finite if every point has a neighborhood that intersects only finitely many sets in the cover.

**Lemma 1.9** (Stone). In a metric space every countable open cover has a locally finite open refinement.

**Proof.** Denote the cover by $\{O_j\}_{j \in \mathbb{N}}$ and introduce the sets

$$\hat{O}_{j,n} = \bigcup_{x \in A_{j,n}} B_{2^{-n}}(x),$$

$$A_{j,n} = \{x \in O_j \setminus (O_1 \cup \cdots \cup O_{j-1}) | x \not\in \bigcup_{k \in \mathbb{N}, 1 \leq l < n} \hat{O}_{k,l} \text{ and } B_{3 \cdot 2^{-n}}(x) \subseteq O_j\}.$$

Then, by construction, $\hat{O}_{j,n}$ is open, $\hat{O}_{j,n} \subseteq O_j$, and it is a cover since for every $x$ there is a smallest $j$ such that $x \in O_j$ and a smallest $n$ such that $B_{3 \cdot 2^{-n}}(x) \subseteq O_j$ implying $x \in \hat{O}_{k,l}$ for some $l \leq n$.

To show that $\hat{O}_{j,n}$ is locally finite fix some $x$ and let $j$ be the smallest integer such that $x \in \hat{O}_{j,n}$ for some $n$. Moreover, choose $m$ such that $B_{2^{-m}}(x) \subseteq \hat{O}_{j,n}$. It suffices to show that:

(i) If $i \geq n + m$ then $B_{2^{-n-m}}(x)$ is disjoint from $\hat{O}_{k,i}$ for all $k$.

(ii) If $i < n + m$ then $B_{2^{-n-m}}(x)$ intersects $\hat{O}_{k,i}$ for at most one $k$.

To show (i) observe that since $i > n$ every ball $B_{2^{-i}}(y)$ used in the definition of $\hat{O}_{k,i}$ has its center outside of $\hat{O}_{j,n}$. Hence $d(x,y) \geq 2^{-m}$ and $B_{2^{-n-m}}(x) \cap B_{2^{-i}}(x) = \emptyset$ since $i \geq m + 1$ as well as $n + m \geq m + 1$.

To show (ii) let $y \in \hat{O}_{k,i}$ and $z \in \hat{O}_{k,i}$ with $j < k$. We will show $d(y,z) > 2^{-n-m+1}$. There are points $r$ and $s$ such that $y \in B_{2^{-i}}(r) \subseteq \hat{O}_{j,i}$ and $z \in B_{2^{-i}}(s) \subseteq \hat{O}_{k,i}$. Then by definition $B_{3 \cdot 2^{-i}}(r) \subseteq O_j$ but $s \not\in O_j$. So $d(r,s) \geq 3 \cdot 2^{-i}$ and $d(y,z) > 2^{-i} \geq 2^{-n-m+1}$.

A subset $K \subseteq X$ is called compact if every open cover has a finite subcover. A set is called relatively compact if its closure is compact.

**Lemma 1.10.** A topological space is compact if and only if it has the finite intersection property: The intersection of a family of closed sets is empty if and only if the intersection of some finite subfamily is empty.

**Proof.** By taking complements, to every family of open sets there is a corresponding family of closed sets and vice versa. Moreover, the open sets are a cover if and only if the corresponding closed sets have empty intersection.
Lemma 1.11. Let $X$ be a topological space.

(i) The continuous image of a compact set is compact.
(ii) Every closed subset of a compact set is compact.
(iii) If $X$ is Hausdorff, every compact set is closed.
(iv) The product of finitely many compact sets is compact.
(v) The finite union of compact sets is again compact.
(vi) If $X$ is Hausdorff, any intersection of compact sets is again compact.

Proof. (i) Observe that if $\{O_\alpha\}$ is an open cover for $f(Y)$, then $\{f^{-1}(O_\alpha)\}$ is one for $Y$.

(ii) Let $\{O_\alpha\}$ be an open cover for the closed subset $Y$ (in the induced topology). Then there are open sets $\tilde{O}_\alpha$ with $O_\alpha = \tilde{O}_\alpha \cap Y$ and $\{\tilde{O}_\alpha\} \cup \{X \setminus Y\}$ is an open cover for $X$ which has a finite subcover. This subcover induces a finite subcover for $Y$.

(iii) Let $Y \subseteq X$ be compact. We show that $X \setminus Y$ is open. Fix $x \in X \setminus Y$ (if $Y = X$, there is nothing to do). By the definition of Hausdorff, for every $y \in Y$ there are disjoint neighborhoods $V(y)$ of $y$ and $U_y(x)$ of $x$. By compactness of $Y$, there are $y_1, \ldots, y_n$ such that the $V(y_j)$ cover $Y$. But then $U(x) = \bigcap_{j=1}^n U_{y_j}(x)$ is a neighborhood of $x$ which does not intersect $Y$.

(iv) Let $\{O_\alpha\}$ be an open cover for $X \times Y$. For every $(x, y) \in X \times Y$ there is some $\alpha(x, y)$ such that $(x, y) \in O_{\alpha(x,y)}$. By definition of the product topology there is some open rectangle $U(x, y) \times V(x, y) \subseteq O_{\alpha(x,y)}$. Hence for fixed $x$, $\{V(x, y)\}_{y \in Y}$ is an open cover of $Y$. Hence there are finitely many points $y_k(x)$ such that the $V(x, y_k(x))$ cover $Y$. Set $U(x) = \bigcap_k U(x, y_k(x))$.

Since finite intersections of open sets are open, $\{U(x)\}_{x \in X}$ is an open cover and there are finitely many points $x_j$ such that the $U(x_j)$ cover $X$. By construction, the $U(x_j) \times V(x_j, y_k(x_j)) \subseteq O_{\alpha(x_j,y_k(x_j))}$ cover $X \times Y$.

(v) Note that a cover of the union is a cover for each individual set and the union of the individual subcovers is the subcover we are looking for.

(vi) Follows from (ii) and (iii) since an intersection of closed sets is closed. □

As a consequence we obtain a simple criterion when a continuous function is a homeomorphism.

Corollary 1.12. Let $X$ and $Y$ be topological spaces with $X$ compact and $Y$ Hausdorff. Then every continuous bijection $f : X \to Y$ is a homeomorphism.
Proof. It suffices to show that \( f \) maps closed sets to closed sets. By (ii) every closed set is compact, by (i) its image is also compact, and by (iii) it is also closed. □

Moreover, item (iv) generalizes to arbitrary products:

**Theorem 1.13** (Tychonoff). The product \( \prod_{\alpha \in A} K_{\alpha} \) of an arbitrary collection of compact topological spaces \( \{K_{\alpha}\}_{\alpha \in A} \) is compact with respect to the product topology.

**Proof.** We say that a family \( F \) of closed subsets of \( K \) has the finite intersection property if the intersection of every finite subfamily has nonempty intersection. The collection of all such families which contain \( F \) is partially ordered by inclusion and every chain has an upper bound (the union of all sets in the chain). Hence, by Zorn's lemma, there is a maximal family \( F_M \) (note that this family is closed under finite intersections).

Denote by \( \pi_{\alpha} : K \to K_{\alpha} \) the projection onto the \( \alpha \) component. Then the closed sets \( \{\pi_{\alpha}(F)\}_{F \in F_M} \) also have the finite intersection property and since \( K_{\alpha} \) is compact, there is some \( x_{\alpha} \in \bigcap_{F \in F_M} \pi_{\alpha}(F) \). Consequently, if \( F_{\alpha} \) is a closed neighborhood of \( x_{\alpha} \), then \( \pi_{\alpha}^{-1}(F_{\alpha}) \in F_M \) (otherwise there would be some \( F \in F_M \) with \( F \cap \pi_{\alpha}^{-1}(F_{\alpha}) = \emptyset \) contradicting \( F_{\alpha} \cap \pi_{\alpha}(F) \neq \emptyset \)). Furthermore, for every finite subset \( A_0 \subseteq A \) we have \( \bigcap_{\alpha \in A_0} \pi_{\alpha}^{-1}(F_{\alpha}) \in F_M \) and so every neighborhood of \( x = (x_{\alpha})_{\alpha \in A} \) intersects \( F \). Since \( F \) is closed, \( x \in F \) and hence \( x \in \bigcap_{F \in F_M} F \). □

A subset \( K \subset X \) is called **sequentially compact** if every sequence from \( K \) has a convergent subsequence whose limit is in \( K \). In a metric space, compact and sequentially compact are equivalent.

**Lemma 1.14.** Let \( X \) be a metric space. Then a subset is compact if and only if it is sequentially compact.

**Proof.** Without loss of generality we can assume the subset to be all of \( X \). Suppose \( X \) is compact and let \( x_n \) be a sequence which has no convergent subsequence. Then \( K = \{x_n\} \) has no limit points and is hence compact by Lemma 1.11 (ii). For every \( n \) there is a ball \( B_{\varepsilon_n}(x_n) \) which contains only finitely many elements of \( K \). However, finitely many suffice to cover \( K \), a contradiction.

Conversely, suppose \( X \) is sequentially compact and let \( \{O_{\alpha}\} \) be some open cover which has no finite subcover. For every \( x \in X \) we can choose some \( \alpha(x) \) such that if \( B_r(x) \) is the largest ball contained in \( O_{\alpha(x)} \), then either \( r \geq 1 \) or there is no \( \beta \) with \( B_{2r}(x) \subset O_{\beta} \) (show that this is possible). Now choose a sequence \( x_n \) such that \( x_n \not\in \bigcup_{m<n} O_{\alpha(x_m)} \). Note that by
construction the distance $d = d(x_m, x_n)$ to every successor of $x_m$ is either larger than 1 or the ball $B_{2d}(x_m)$ will not fit into any of the $O_\alpha$.

Now let $y$ be the limit of some convergent subsequence and fix some $r \in (0,1)$ such that $B_r(y) \subseteq O_\alpha(y)$. Then this subsequence must eventually be in $B_{r/5}(y)$, but this is impossible since if $d = d(x_{n_1}, x_{n_2})$ is the distance between two consecutive elements of this subsequence, then $B_{2d}(x_{n_1})$ cannot fit into $O_\alpha(y)$ by construction whereas on the other hand $B_{2d}(x_{n_1}) \subseteq B_{4r/5}(a) \subseteq O_\alpha(y)$.

If we drop the requirement that the limit must be in $K$, we obtain relatively compact sets:

**Corollary 1.15.** Let $X$ be a metric space and $K \subset X$. Then $K$ is relatively compact if and only if every sequence from $K$ has a convergent subsequence (the limit must not be in $K$).

**Proof.** For any sequence $x_n \in K$ we can find a nearby sequence $y_n \in K$ with $x_n - y_n \to 0$. If we can find a convergent subsequence of $y_n$ then the corresponding subsequence of $x_n$ will also converge (to the same limit) and $K$ is (sequentially) compact in this case. The converse is trivial.

As another simple consequence observe that

**Corollary 1.16.** A compact metric space $X$ is complete and separable.

**Proof.** Completeness is immediate from the previous lemma. To see that $X$ is separable note that, by compactness, for every $n \in \mathbb{N}$ there is a finite set $S_n \subseteq X$ such that the balls $\{B_{1/n}(x)\}_{x \in S_n}$ cover $X$. Then $\bigcup_{n \in \mathbb{N}} S_n$ is a countable dense set.

In a metric space, a set is called **bounded** if it is contained inside some ball. Note that compact sets are always bounded since Cauchy sequences are bounded (show this!). In $\mathbb{R}^n$ (or $\mathbb{C}^n$) the converse also holds.

**Theorem 1.17** (Heine–Borel). In $\mathbb{R}^n$ (or $\mathbb{C}^n$) a set is compact if and only if it is bounded and closed.

**Proof.** By Lemma 1.11 (ii), (iii), and (iv) it suffices to show that a closed interval in $I \subseteq \mathbb{R}$ is compact. Moreover, by Lemma 1.14, it suffices to show that every sequence in $I = [a, b]$ has a convergent subsequence. Let $x_n$ be our sequence and divide $I = [a, \frac{a+b}{2}] \cup [\frac{a+b}{2}, b]$. Then at least one of these two intervals, call it $I_1$, contains infinitely many elements of our sequence. Let $y_1 = x_{n_1}$ be the first one. Subdivide $I_1$ and pick $y_2 = x_{n_2}$, with $n_2 > n_1$ as before. Proceeding like this, we obtain a Cauchy sequence $y_n$ (note that by construction $I_{n+1} \subseteq I_n$ and hence $|y_n - y_m| \leq \frac{b-a}{2^n}$ for $m \geq n$).
By Lemma 1.14 this is equivalent to

**Theorem 1.18** (Bolzano–Weierstraß). Every bounded infinite subset of \( \mathbb{R}^n \) (or \( \mathbb{C}^n \)) has at least one limit point.

Combining Theorem 1.17 with Lemma 1.11 (i) we also obtain the extreme value theorem.

**Theorem 1.19** (Weierstraß). Let \( X \) be compact. Every continuous function \( f : X \to \mathbb{R} \) attains its maximum and minimum.

A metric space for which the Heine–Borel theorem holds is called proper. Lemma 1.11 (ii) shows that \( X \) is proper if and only if every closed ball is compact. Note that a proper metric space must be complete (since every Cauchy sequence is bounded). A topological space is called **locally compact** if every point has a compact neighborhood. Clearly a proper metric space is locally compact.

The **distance** between a point \( x \in X \) and a subset \( Y \subseteq X \) is

\[
\text{dist}(x, Y) = \inf_{y \in Y} d(x, y).
\]  

(1.17)

Note that \( x \) is a limit point of \( Y \) if and only if \( \text{dist}(x, Y) = 0 \).

**Lemma 1.20.** Let \( X \) be a metric space. Then

\[
|\text{dist}(x, Y) − \text{dist}(z, Y)| \leq d(x, z).
\]  

(1.18)

In particular, \( x \mapsto \text{dist}(x, Y) \) is continuous.

**Proof.** Taking the infimum in the triangle inequality \( d(x, y) \leq d(x, z) + d(z, y) \) shows \( \text{dist}(x, Y) \leq d(x, z) + \text{dist}(z, Y) \). Hence \( \text{dist}(x, Y) − \text{dist}(z, Y) \leq d(x, z) \). Interchanging \( x \) and \( z \) shows \( \text{dist}(z, Y) − \text{dist}(x, Y) \leq d(x, z) \). □

**Lemma 1.21** (Urysohn). Suppose \( C_1 \) and \( C_2 \) are disjoint closed subsets of a metric space \( X \). Then there is a continuous function \( f : X \to [0, 1] \) such that \( f \) is zero on \( C_2 \) and one on \( C_1 \).

If \( X \) is locally compact and \( C_1 \) is compact, one can choose \( f \) with compact support.

**Proof.** To prove the first claim, set \( f(x) = \frac{\text{dist}(x, C_2)}{\text{dist}(x, C_1) + \text{dist}(x, C_2)} \). For the second claim, observe that there is an open set \( O \) such that \( \overline{O} \) is compact and \( C_1 \subset O \subset \overline{O} \subset X \setminus C_2 \). In fact, for every \( x \in C_1 \), there is a ball \( B(x) \) such that \( B(x) \) is compact and \( B(x) \subset X \setminus C_2 \). Since \( C_1 \) is compact, finitely many of them cover \( C_1 \) and we can choose the union of those balls to be \( O \). Now replace \( C_2 \) by \( X \setminus O \). □
Note that Urysohn’s lemma implies that a metric space is normal; that is, for any two disjoint closed sets \( C_1 \) and \( C_2 \), there are disjoint open sets \( O_1 \) and \( O_2 \) such that \( C_j \subseteq O_j \), \( j = 1, 2 \). In fact, choose \( f \) as in Urysohn’s lemma and set \( O_1 = f^{-1}((0,1/2)) \), respectively, \( O_2 = f^{-1}((1/2,1]) \).

**Lemma 1.22.** Let \( X \) be a metric space and \( \{O_j\} \) a countable open cover. Then there is a continuous partition of unity subordinate to this cover; that is, there are continuous functions \( h_j : X \to [0,1] \) such that \( h_j \) has compact support contained in \( O_j \) and

\[
\sum_j h_j(x) = 1. \tag{1.19}
\]

Moreover, the partition of unity can be chosen locally finite; that is, every \( x \) has a neighborhood where all but a finite number of the functions \( h_j \) vanish.

**Proof.** For notational simplicity we assume \( j \in \mathbb{N} \). Now introduce \( f_n(x) = \min(1, \sup_{j \leq n} d(x, X\setminus O_j)) \) and \( g_n = f_n - f_{n-1} \) (with the convention \( f_0(x) = 0 \)). Since \( f_n \) is increasing we have \( 0 \leq g_n \leq 1 \). Moreover, \( g_n(x) > 0 \) implies \( d(x, X\setminus O_n) > 0 \) and thus \( \text{supp}(g_n) \subseteq O_n \). Next, by monotonicity \( f_\infty = \lim_{n \to \infty} f_n = \sum_n g_n \) exists and is everywhere positive since \( \{O_j\} \) is a cover. Finally, by

\[
|f_n(x) - f_n(y)| \leq \sup_{j \leq n} d(x, X\setminus O_j) - \sup_{j \leq n} d(y, X\setminus O_j) \\
\leq \sup_{j \leq n} d(x, X\setminus O_j) - d(y, X\setminus O_j) \leq d(x,y)
\]

we see that all \( f_n \) (and hence all \( g_n \)) are continuous. Moreover, the very same argument shows that \( f_\infty \) is continuous and thus we have found the required partition of unity \( h_j = g_j / f_\infty \).

Finally, by Lemma 1.9 we can replace the cover by a locally finite refinement and the resulting partition of unity for this refinement will be locally finite. \( \square \)

Another important result is the **Tietze extension theorem:**

**Theorem 1.23** (Tietze). Suppose \( C \) is a closed subset of a metric space \( X \). For every continuous function \( f : C \to [-1,1] \) there is a continuous extension \( \overline{f} : X \to [-1,1] \).

**Proof.** The idea is to construct a rough approximation using Urysohn’s lemma and then iteratively improve this approximation. To this end we set \( C_1 = f^{-1}([\frac{1}{3},1]) \) and \( C_2 = f^{-1}([-1,-\frac{1}{3}]) \) and let \( g \) be the function from Urysohn’s lemma. Then \( f_1 = \frac{2g-1}{3} \) satisfies \( |f(x) - f_1(x)| \leq \frac{2}{3} \) for \( x \in C \) as well as \( |f_1(x)| \leq \frac{1}{3} \) for all \( x \in X \). Applying this same procedure to \( f-f_1 \) we obtain a function \( f_2 \) such that \( |f(x) - f_1(x) - f_2(x)| \leq \left( \frac{2}{3} \right)^2 \) for \( x \in C \) and
|f_2(x)| \leq \frac{1}{3} \left( \frac{2}{3} \right)^n. \quad \text{Continuing this process we arrive at a sequence of functions} \ f_n \ \text{such that} \ \ |f(x) - \sum_{j=1}^{n} f_j(x)| \leq \left( \frac{2}{3} \right)^n \ \text{for} \ x \in C \ \text{and} \ |f_n(x)| \leq \frac{1}{3} \left( \frac{2}{3} \right)^{n-1}.

\text{By construction the corresponding series converges uniformly to the desired extension} \ \overline{f} = \sum_{j=1}^{\infty} f_j. \ \square

**Problem 1.1.** Show that \(|d(x, y) - d(z, y)| \leq d(x, z)|.

**Problem 1.2.** Show the quadrangle inequality \(|d(x, y) - d(x', y')| \leq d(x, x') + d(y, y').

**Problem 1.3.** Show that if \(f : [0, \infty) \to \mathbb{R}\) is concave, \(f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y)\) for \(\lambda \in [0, 1]\), and satisfies \(f(0) = 0\), then it is subadditive, \(f(x + y) \leq f(x) + f(y)\). (Hint: begin by showing \(f(\lambda x) \geq \lambda f(x)\).)

**Problem 1.4.** Show that the closure satisfies the Kuratowski closure axioms.

**Problem 1.5.** Show that the closure and interior operators are dual in the sense that
\[ X \setminus \overline{A} = (X \setminus A)^{\circ} \quad \text{and} \quad X \setminus A^{\circ} = (X \setminus A). \]

(Hint: De Morgan’s laws.)

**Problem 1.6.** Let \(U \subseteq V\) be subsets of a metric space \(X\). Show that if \(U\) is dense in \(V\) and \(V\) is dense in \(X\), then \(U\) is dense in \(X\).

**Problem 1.7.** Show that every open set \(O \subseteq \mathbb{R}\) can be written as a countable union of disjoint intervals. (Hint: Consider the set \(\{I_\alpha\}\) of all maximal open subintervals of \(O\); that is, \(I_\alpha \subseteq O\) and there is no other subinterval of \(O\) which contains \(I_\alpha\).

**Problem 1.8.** Show that the boundary of \(A\) is given by \(\partial A = \overline{A} \setminus A^{\circ}\).

**Problem 1.9.** Let \(X\) be a topological space and \(f : X \to \mathbb{R}\). Let \(x \in X_0\) and let \(B(x_0)\) be a neighborhood base for \(x_0\). Define
\[ \liminf_{x \to x_0} f(x) = \sup_{U \in B(x_0)} \inf_{U(x_0)} f, \quad \limsup_{x \to x_0} f(x) = \inf_{U \in B(x_0)} \sup_{U(x_0)} f. \]

Show that both are independent of the neighborhood base and satisfy

(i) \(\liminf_{x \to x_0} (-f(x)) = -\limsup_{x \to x_0} f(x)\).

(ii) \(\liminf_{x \to x_0} (\alpha f(x)) = \alpha \liminf_{x \to x_0} f(x), \ \alpha \geq 0\).

(iii) \(\liminf_{x \to x_0} (f(x) + g(x)) \geq \liminf_{x \to x_0} f(x) + \liminf_{x \to x_0} g(x)\).

Moreover, show that
\[ \liminf_{n \to \infty} f(x_n) \geq \liminf_{x \to x_0} f(x), \quad \limsup_{n \to \infty} f(x_n) \leq \limsup_{x \to x_0} f(x) \]

for every sequence \(x_n \to x_0\) and there exists a sequence attaining equality if \(X\) is a metric space.
1.2. The Banach space of continuous functions

Now let us have a first look at Banach spaces by investigating the set of continuous functions \( C(I) \) on a compact interval \( I = [a, b] \subset \mathbb{R} \). Since we want to handle complex models, we will always consider complex-valued functions!

One way of declaring a distance, well-known from calculus, is the maximum norm:

\[
\|f\|_\infty = \max_{x \in I} |f(x)|.
\]  

(1.20)

It is not hard to see that with this definition \( C(I) \) becomes a normed vector space:

A normed vector space \( X \) is a vector space \( X \) over \( \mathbb{C} \) (or \( \mathbb{R} \)) with a nonnegative function (the norm) \( \|\cdot\| \) such that

- \( \|f\| > 0 \) for \( f \neq 0 \) (positive definiteness),
- \( \|\alpha f\| = |\alpha| \|f\| \) for all \( \alpha \in \mathbb{C}, f \in X \) (positive homogeneity), and
- \( \|f + g\| \leq \|f\| + \|g\| \) for all \( f, g \in X \) (triangle inequality).

If positive definiteness is dropped from the requirements, one calls \( \|\cdot\| \) a seminorm.

From the triangle inequality we also get the inverse triangle inequality (Problem 1.10)

\[
\|f\| - \|g\| \leq \|f - g\|,
\]  

(1.21)

which shows that the norm is continuous.

Next, recall that a subset \( C \subseteq X \) is convex if for every \( x, y \in C \) we also have \( \lambda x + (1 - \lambda)y \in C \) whenever \( \lambda \in (0, 1) \). Moreover, a mapping \( f : C \to \mathbb{R} \) is called convex if \( f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \) and in our case it is not hard to check that every norm is convex:

\[
\|\lambda x + (1 - \lambda)y\| \leq \lambda \|x\| + (1 - \lambda)\|y\|, \quad \lambda \in [0, 1].
\]  

(1.22)

Moreover, choosing \( \lambda = \frac{1}{2} \) we get back the triangle inequality upon using homogeneity. In particular, the triangle inequality could be replaced by convexity in the definition.

Once we have a norm, we have a distance \( d(f, g) = \|f - g\| \) and hence we know when a sequence of vectors \( f_n \) converges to a vector \( f \). We will write \( f_n \to f \) or \( \lim_{n \to \infty} f_n = f \), as usual, in this case. Moreover, a mapping \( F : X \to Y \) between two normed spaces is called continuous if \( f_n \to f \) implies \( F(f_n) \to F(f) \). In fact, the norm, vector addition, and multiplication by scalars are continuous (Problem 1.11). In particular, a normed vector space is an example of a a topological vector space, that
is, a vector space which carries a topology such that both vector addition \( X \times X \to X \) and scalar multiplication \( \mathbb{C} \times X \to X \) are continuous mappings.

In addition to the concept of convergence we have also the concept of a **Cauchy sequence** and hence the concept of completeness: A normed space is called **complete** if every Cauchy sequence has a limit. A complete normed space is called a **Banach space**.

**Example.** The space \( \ell^1(\mathbb{N}) \) of all complex-valued sequences \( a = (a_j)_{j=1}^{\infty} \) for which the norm

\[
\|a\|_1 = \sum_{j=1}^{\infty} |a_j|
\]

is finite is a Banach space.

To show this, we need to verify three things: (i) \( \ell^1(\mathbb{N}) \) is a vector space that is closed under addition and scalar multiplication, (ii) \( \| \cdot \|_1 \) satisfies the three requirements for a norm, and (iii) \( \ell^1(\mathbb{N}) \) is complete.

First of all, observe

\[
\sum_{j=1}^{k} |a_j + b_j| \leq \sum_{j=1}^{k} |a_j| + \sum_{j=1}^{k} |b_j| \leq \|a\|_1 + \|b\|_1
\]

for every finite \( k \). Letting \( k \to \infty \), we conclude that \( \ell^1(\mathbb{N}) \) is closed under addition and that the triangle inequality holds. That \( \ell^1(\mathbb{N}) \) is closed under scalar multiplication together with homogeneity as well as definiteness are straightforward. It remains to show that \( \ell^1(\mathbb{N}) \) is complete. Let \( a^n = (a_j^n)_{j=1}^{\infty} \) be a Cauchy sequence; that is, for given \( \varepsilon > 0 \) we can find an \( N_\varepsilon \) such that \( \|a^n - a^m\|_1 \leq \varepsilon \) for \( m, n \geq N_\varepsilon \). This implies, in particular, \( |a_j^m - a_j^n| \leq \varepsilon \) for every fixed \( j \). Thus \( a_j^n \) is a Cauchy sequence for fixed \( j \) and, by completeness of \( \mathbb{C} \), it has a limit: \( \lim_{n \to \infty} a_j^n = a_j \). Now consider

\[
\sum_{j=1}^{k} |a_j^n - a_j^m| \leq \varepsilon
\]

(1.25)

and take \( m \to \infty \):

\[
\sum_{j=1}^{k} |a_j - a_j^n| \leq \varepsilon.
\]

(1.26)

Since this holds for all finite \( k \), we even have \( \|a - a^n\|_1 \leq \varepsilon \). Hence \( (a - a^n) \in \ell^1(\mathbb{N}) \) and since \( a^n \in \ell^1(\mathbb{N}) \), we finally conclude \( a = a^n + (a - a^n) \in \ell^1(\mathbb{N}) \). By our estimate \( \|a - a^n\|_1 \leq \varepsilon \), our candidate \( a \) is indeed the limit of \( a^n \). \( \diamond \)
Example. The previous example can be generalized by considering the space $\ell^p(N)$ of all complex-valued sequences $a = (a_j)_{j=1}^\infty$ for which the norm

$$
\|a\|_p = \left(\sum_{j=1}^\infty |a_j|^p\right)^{1/p}, \quad p \in [1, \infty),
$$

is finite. By $|a_j + b_j|^p \leq 2^p \max(|a_j|, |b_j|)^p = 2^p \max(|a_j|^p, |b_j|^p) \leq 2^p(|a_j|^p + |b_j|^p)$ it is a vector space, but the triangle inequality is only easy to see in the case $p = 1$. (It is also not hard to see that it fails for $p < 1$, which explains our requirement $p \geq 1$. See also Problem 1.20.)

To prove it we need the elementary inequality (Problem 1.15)

$$
\alpha^{1/p} \beta^{1/q} \leq \frac{1}{p} \alpha + \frac{1}{q} \beta, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad \alpha, \beta \geq 0,
$$

which implies Hölder’s inequality

$$
\|ab\|_1 \leq \|a\|_p \|b\|_q
$$

for $x \in \ell^p(N)$, $y \in \ell^q(N)$. In fact, by homogeneity of the norm it suffices to prove the case $\|a\|_p = \|b\|_q = 1$. But this case follows by choosing $\alpha = |a_j|^p$ and $\beta = |b_j|^q$ in (1.28) and summing over all $j$. (A different proof based on convexity will be given in Section 8.2.)

Now using $|a_j + b_j|^p \leq |a_j| |a_j + b_j|^{p-1} + |b_j| |a_j + b_j|^{p-1}$, we obtain from Hölder’s inequality (note $(p-1)q = p$)

$$
\|a + b\|_p \leq \|a\|_p \|a + b\|^{p-1}_q + \|b\|_p \|a + b\|^{p-1}_p
$$

$$
= (\|a\|_p + \|b\|_p) \|a + b\|^{p-1}_p.
$$

Hence $\ell^p$ is a normed space. That it is complete can be shown as in the case $p = 1$ (Problem 1.16).

Example. The space $\ell^\infty(N)$ of all complex-valued bounded sequences $a = (a_j)_{j=1}^\infty$ together with the norm

$$
\|a\|_\infty = \sup_{j \in \mathbb{N}} |a_j|
$$

is a Banach space (Problem 1.17). Note that with this definition, Hölder’s inequality (1.29) remains true for the cases $p = 1$, $q = \infty$ and $p = \infty$, $q = 1$. The reason for the notation is explained in Problem 1.19.

Example. Every closed subspace of a Banach space is again a Banach space. For example, the space $c_0(N) \subset \ell^\infty(N)$ of all sequences converging to zero is a closed subspace. In fact, if $a \in \ell^\infty(N) \setminus c_0(N)$, then $\limsup_{j \to \infty} |a_j| = \varepsilon > 0$ and thus $\|a - b\|_\infty \geq \varepsilon$ for every $b \in c_0(N)$. 
1.2. The Banach space of continuous functions

Now what about completeness of \( C(I) \)? A sequence of functions \( f_n(x) \) converges to \( f \) if and only if

\[
\lim_{n \to \infty} \|f - f_n\|_\infty = \lim_{n \to \infty} \sup_{x \in I} |f_n(x) - f(x)| = 0. \quad (1.31)
\]

That is, in the language of real analysis, \( f_n \) converges uniformly to \( f \). Now let us look at the case where \( f_n \) is only a Cauchy sequence. Then \( f_n(x) \) is clearly a Cauchy sequence of real numbers for every fixed \( x \in I \). In particular, by completeness of \( \mathbb{C} \), there is a limit \( f(x) \) for each \( x \). Thus we get a limiting function \( f(x) \). Moreover, letting \( m \to \infty \) in

\[
|f_m(x) - f_n(x)| \leq \varepsilon \quad \forall m, n > N_\varepsilon, x \in I, \quad (1.32)
\]

we see

\[
|f(x) - f_n(x)| \leq \varepsilon \quad \forall n > N_\varepsilon, x \in I; \quad (1.33)
\]

that is, \( f_n(x) \) converges uniformly to \( f(x) \). However, up to this point we do not know whether it is in our vector space \( C(I) \), that is, whether it is continuous. Fortunately, there is a well-known result from real analysis which tells us that the uniform limit of continuous functions is again continuous: Fix \( x \in I \) and \( \varepsilon > 0 \). To show that \( f \) is continuous we need to find a \( \delta \) such that \( |x - y| < \delta \) implies \( |f(x) - f(y)| < \varepsilon \). Pick \( n \) so that \( \|f_n - f\|_\infty < \varepsilon/3 \) and \( \delta \) so that \( |x - y| < \delta \) implies \( |f_n(x) - f_n(y)| < \varepsilon/3 \). Then \( |x - y| < \delta \) implies

\[
|f(x) - f(y)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon
\]

as required. Hence \( f(x) \in C(I) \) and thus every Cauchy sequence in \( C(I) \) converges. Or, in other words,

**Theorem 1.24.** \( C(I) \) with the maximum norm is a Banach space.

Next we want to look at countable bases. To this end we introduce a few definitions first.

The set of all finite linear combinations of a set of vectors \( \{u_n\}_{n \in \mathbb{N}} \subset X \) is called the span of \( \{u_n\}_{n \in \mathbb{N}} \) and denoted by

\[
\text{span}\{u_n\}_{n \in \mathbb{N}} = \left\{ \sum_{j=1}^{m} c_j u_{n_j} \mid n_j \in \mathbb{N}, c_j \in \mathbb{C}, m \in \mathbb{N} \right\}. \quad (1.34)
\]

A set of vectors \( \{u_n\}_{n \in \mathbb{N}} \subset X \) is called linearly independent if every finite subset is. If \( \{u_n\}_{n=1}^{N} \subset X, \ N \in \mathbb{N} \cup \{\infty\}, \) is countable, we can throw away all elements which can be expressed as linear combinations of the previous ones to obtain a subset of linearly independent vectors which have the same span.
We will call a countable set of vectors \( (u_n)_{n=1}^{\infty} \subset X \) a **Schauder basis** if every element \( f \in X \) can be uniquely written as a countable linear combination of the basis elements:

\[
f = \sum_{n=1}^{\infty} c_n u_n, \quad c_n = c_n(f) \in \mathbb{C},
\]

where the sum has to be understood as a limit if \( N = \infty \) (the sum is not required to converge unconditionally and hence the order of the basis elements is important). Since we have assumed the coefficients \( c_n(f) \) to be uniquely determined, the vectors are necessarily linearly independent.

**Example.** The set of vectors \( \delta^n \), with \( \delta^n_n = 1 \) and \( \delta^n_m = 0 \), \( n \neq m \), is a Schauder basis for the Banach space \( \ell^p(\mathbb{N}) \), \( 1 \leq p < \infty \).

Let \( a = (a_j)_{j=1}^{\infty} \in \ell^p(\mathbb{N}) \) be given and set \( a^m = \sum_{n=1}^{m} a_n \delta^n \). Then

\[
\|a - a^m\|_1 = \left( \sum_{j=m+1}^{\infty} |a_j|^p \right)^{1/p} \to 0
\]

since \( a_j^m = a_j \) for \( 1 \leq j \leq m \) and \( a_j^m = 0 \) for \( j > m \). Hence

\[
a = \sum_{n=1}^{\infty} a_n \delta^n
\]

and \( \{\delta^n\}_{n=1}^{\infty} \) is a Schauder basis (uniqueness of the coefficients is left as an exercise).

Note that \( \{\delta^n\}_{n=1}^{\infty} \) is also Schauder basis for \( c_0(\mathbb{N}) \) but not for \( \ell^\infty(\mathbb{N}) \). \( \diamond \)

A set whose span is dense is called **total**, and if we have a countable total set, we also have a countable dense set (consider only linear combinations with rational coefficients — show this). A normed vector space containing a countable dense set is called **separable**.

**Example.** Every Schauder basis is total and thus every Banach space with a Schauder basis is separable (the converse is not true). In particular, the Banach space \( \ell^p(\mathbb{N}) \) is separable. \( \diamond \)

While we will not give a Schauder basis for \( C(I) \) (Problem 1.21), we will at least show that it is separable. In order to prove this, we need a lemma first.

**Lemma 1.25** (Smoothing). Let \( u_n(x) \) be a sequence of nonnegative continuous functions on \([-1, 1]\) such that

\[
\int_{|x| \leq 1} u_n(x) \, dx = 1 \quad \text{and} \quad \int_{\delta \leq |x| \leq 1} u_n(x) \, dx \to 0, \quad \delta > 0.
\]

(In other words, \( u_n \) has mass one and concentrates near \( x = 0 \) as \( n \to \infty \).)
Then for every \( f \in C\left[\frac{-1}{2}, \frac{1}{2}\right] \) which vanishes at the endpoints, \( f\left(\frac{-1}{2}\right) = f\left(\frac{1}{2}\right) = 0 \), we have that

\[
f_n(x) = \int_{-1/2}^{1/2} u_n(x - y)f(y)dy
\]

converges uniformly to \( f(x) \).

**Proof.** Since \( f \) is uniformly continuous, for given \( \varepsilon \) we can find a \( \delta < 1/2 \) (independent of \( x \)) such that \( |f(x) - f(y)| \leq \varepsilon \) whenever \( |x - y| \leq \delta \). Moreover, we can choose \( n \) such that \( \int_{|y| \leq 1} u_n(y)dy \leq \varepsilon \). Now abbreviate \( M = \max_{x \in [-1/2, 1/2]} \{1, |f(x)|\} \) and note

\[
|f(x) - \int_{-1/2}^{1/2} u_n(x - y)f(y)dy| = |f(x)| - \int_{-1/2}^{1/2} u_n(x - y)dy |\leq M\varepsilon.
\]

In fact, either the distance of \( x \) to one of the boundary points \( \pm 1/2 \) is smaller than \( \delta \) and hence \( |f(x)| \leq \varepsilon \) or otherwise \( [-\delta, \delta] \subset [x-1/2, x+1/2] \) and the difference between one and the integral is smaller than \( \varepsilon \).

Using this, we have

\[
|f_n(x) - f(x)| \leq \int_{-1/2}^{1/2} u_n(x - y)|f(y) - f(x)|dy + M\varepsilon
\]

\[
= \int_{|y| \leq 1/2, |x-y| \leq \delta} u_n(x - y)|f(y) - f(x)|dy
\]

\[
+ \int_{|y| \leq 1/2, |x-y| \geq \delta} u_n(x - y)|f(y) - f(x)|dy + M\varepsilon
\]

\[
\leq \varepsilon + 2M\varepsilon + M\varepsilon = (1 + 3M)\varepsilon,
\]

which proves the claim. \( \square \)

Note that \( f_n \) will be as smooth as \( u_n \), hence the title smoothing lemma. Moreover, \( f_n \) will be a polynomial if \( u_n \) is. The same idea is used to approximate noncontinuous functions by smooth ones (of course the convergence will no longer be uniform in this case).

Now we are ready to show:

**Theorem 1.26 (Weierstraß).** Let \( I \) be a compact interval. Then the set of polynomials is dense in \( C(I) \).

**Proof.** Let \( f(x) \in C(I) \) be given. By considering \( f(x) - f(a) - \frac{f(b) - f(a)}{b-a}(x-a) \) it is no loss to assume that \( f \) vanishes at the boundary points. Moreover, without restriction, we only consider \( I = [-1/2, 1/2] \) (why?).
Now the claim follows from Lemma 1.25 using
\[ u_n(x) = \frac{1}{I_n}(1 - x^2)^n, \]
where (using integration by parts)
\[ I_n = \int_{-1}^{1} (1 - x^2)^n dx = \frac{n}{n + 1} \int_{-1}^{1} (1 - x)^{n-1}(1 + x)^{n+1} dx \]
\[ = \cdots = \frac{n!}{(n + 1) \cdots (2n + 1)} 2^{2n+1} = \frac{(n!)^2 2^{2n+1}}{(2n + 1)!} = \frac{n!}{\left(\frac{1}{2} + 1\right) \cdots (\frac{1}{2} + n)}. \]
Indeed, the first part of (1.36) holds by construction, and the second part follows from the elementary estimate
\[ \frac{1}{\frac{1}{2} + n} < I_n < 2, \]
which shows \( \int_{\delta \leq |x| \leq 1} u_n(x) dx \leq 2u_n(\delta) < (2n + 1)(1 - \delta^2)^n \to 0. \) \( \square \)

**Corollary 1.27.** \( C(I) \) is separable.

However, \( \ell^\infty(\mathbb{N}) \) is not separable (Problem 1.18)!

Note that while the proof of Theorem 1.26 provides an explicit way of constructing a sequence of polynomials \( f_n(x) \) which will converge uniformly to \( f(x) \) this method still has a few drawbacks: Suppose we have approximated \( f \) by a polynomial of degree \( n \) but our approximation turns out to be insufficient for a certain purpose. First of all, our polynomial will not be optimal in general, that is, there could be another polynomial of the same degree giving a better approximation. Moreover, if we ignore this fact, and simply increase the degree, all coefficients will change. This is in contradiction to a Schauder basis where the old coefficients will remain the same if we add one new element from the basis (hence it suffices to compute one new coefficient).

We will see in the next section that the concept of orthogonality resolves these problems.

**Problem 1.10.** Show that \( \|f\| - \|g\| \leq \|f - g\| \).

**Problem 1.11.** Let \( X \) be a Banach space. Show that the norm, vector addition, and multiplication by scalars are continuous. That is, if \( f_n \to f \), \( g_n \to g \), and \( \alpha_n \to \alpha \), then \( \|f_n\| \to \|f\| \), \( f_n + g_n \to f + g \), and \( \alpha_n g_n \to \alpha g \).

**Problem 1.12.** Let \( X \) be a Banach space. Show that \( \sum_{j=1}^{\infty} \|f_j\| < \infty \) implies that
\[ \sum_{j=1}^{\infty} f_j = \lim_{n \to \infty} \sum_{j=1}^{n} f_j \]
exists. The series is called **absolutely convergent** in this case.
Problem 1.13. While $\ell^1(N)$ is separable, it still has room for an uncountable set of linearly independent vectors. Show this by considering vectors of the form
$$ a^\alpha = (1, \alpha, \alpha^2, \ldots), \quad \alpha \in (0, 1). $$
(Hint: Recall the Vandermonde determinant.)

Problem 1.14. A Hamel basis is a maximal set of linearly independent vectors. Show that every vector space $X$ has a Hamel basis $\{u_\alpha\}_{\alpha \in A}$. Show that given a Hamel basis, every $x \in X$ can be written as a finite linear combination $x = \sum c_j u_{\alpha_j}$, where the vectors $u_{\alpha_j}$ and the constants $c_j$ are uniquely determined. (Hint: Use the Zorn’s lemma, see Appendix A, to show existence.)

Problem 1.15. Prove (1.28). (Hint: Take logarithms on both sides.)

Problem 1.16. Show that $\ell^p(N)$ is complete.

Problem 1.17. Show that $\ell^\infty(N)$ is a Banach space.

Problem 1.18. Show that $\ell^\infty(N)$ is not separable. (Hint: Consider sequences which take only the value one and zero. How many are there? What is the distance between two such sequences?)

Problem 1.19. Show that $p_0 \leq p$ implies $\ell^{p_0}(N) \subseteq \ell^p(N)$ and $\|a\|_p \leq \|a\|_{p_0}$. Moreover, show
$$\lim_{p \to \infty} \|a\|_p = \|a\|_\infty.$$  

Problem 1.20. Formally extend the definition of $\ell^p(N)$ to $p \in (0, 1)$. Show that $\|a\|_p$ does not satisfy the triangle inequality. However, show that it is a quasinormed space, that is, it satisfies all requirements for a normed space except for the triangle inequality which is replaced by
$$\|a + b\| \leq K(\|a\| + \|b\|)$$
with some constant $K \geq 1$. Show, in fact,
$$\|a + b\|_p \leq 2^{1/p-1}(\|a\|_p + \|b\|_p), \quad p \in (0, 1).$$
Moreover, show that $\|a\|_p^p$ satisfies the triangle inequality in this case, but of course it is no longer homogeneous (but at least you can get an honest metric $d(a, b) = \|a - b\|_p^p$ which gives rise to the same topology). (Hint: Show $\alpha + \beta \leq (\alpha^p + \beta^p)^{1/p} \leq 2^{1/p-1}(\alpha + \beta)$ for $0 < p < 1$ and $\alpha, \beta \geq 0$.)

Problem 1.21. Show that the following set of functions is a Schauder basis for $C[0,1]$: We start with $u_1(t) = t$, $u_2(t) = 1 - t$ and then split $[0,1]$ into $2^n$ intervals of equal length and let $u_{2^n+k+1}(t)$, for $1 \leq k \leq 2^n$ be a piecewise linear peak of height 1 supported in the $k$'th subinterval: $u_{2^n+k+1}(t) = \max(0, 1 - |2^{n+1}t - 2k + 1|)$ for $n \in \mathbb{N}_0$ and $1 \leq k \leq 2^n$. 
1.3. The geometry of Hilbert spaces

So it looks like $C(I)$ has all the properties we want. However, there is still one thing missing: How should we define orthogonality in $C(I)$? In Euclidean space, two vectors are called orthogonal if their scalar product vanishes, so we would need a scalar product:

Suppose $H$ is a vector space. A map $\langle \cdot,\cdot \rangle : H \times H \to \mathbb{C}$ is called a sesquilinear form if it is conjugate linear in the first argument and linear in the second; that is,

$$\langle \alpha_1 f_1 + \alpha_2 f_2, g \rangle = \alpha_1^* \langle f_1, g \rangle + \alpha_2^* \langle f_2, g \rangle,$$
$$\langle f, \alpha_1 g_1 + \alpha_2 g_2 \rangle = \alpha_1 \langle f, g_1 \rangle + \alpha_2 \langle f, g_2 \rangle,$$ \hspace{1cm} (1.39)

where ‘$*$’ denotes complex conjugation. A symmetric

$$\langle f, g \rangle = (g, f)^*$$ \hspace{1cm} (symmetry)

sesquilinear form is also called a Hermitian form and a positive definite

$$\langle f, f \rangle > 0 \text{ for } f \neq 0$$ \hspace{1cm} (positive definite),

Hermitian form is called an inner product or scalar product. Associated with every scalar product is a norm

$$\|f\| = \sqrt{\langle f, f \rangle}.$$ \hspace{1cm} (1.40)

Only the triangle inequality is nontrivial. It will follow from the Cauchy–Schwarz inequality below. Until then, just regard (1.40) as a convenient short hand notation.

The pair $(H, \langle \cdot,\cdot \rangle)$ is called an inner product space. If $H$ is complete (with respect to the norm (1.40)), it is called a Hilbert space.

**Example.** Clearly, $\mathbb{C}^n$ with the usual scalar product

$$\langle a, b \rangle = \sum_{j=1}^{n} a_j^* b_j$$ \hspace{1cm} (1.41)

is a (finite dimensional) Hilbert space. \hspace{1cm} ⋄

**Example.** A somewhat more interesting example is the Hilbert space $\ell^2(\mathbb{N})$, that is, the set of all complex-valued sequences

$$\left\{ (a_j)_{j=1}^{\infty} \mid \sum_{j=1}^{\infty} |a_j|^2 < \infty \right\}$$ \hspace{1cm} (1.42)

with scalar product

$$\langle a, b \rangle = \sum_{j=1}^{\infty} a_j^* b_j.$$ \hspace{1cm} (1.43)
1.3. The geometry of Hilbert spaces

By the Cauchy–Schwarz inequality for \( \mathbb{C}^n \) we infer
\[
\left| \sum_{j=1}^{n} a_j^* b_j \right|^2 \leq \left( \sum_{j=1}^{n} |a_j|^2 \right) \left( \sum_{j=1}^{n} |b_j|^2 \right) \leq \sum_{j=1}^{\infty} |a_j|^2 \sum_{j=1}^{\infty} |b_j|^2
\]
that the sum in the definition of the scalar product is absolutely convergent (an thus well-defined) for \( a, b \in \ell^2(\mathbb{N}) \). Observe that the norm \( \|a\| = \sqrt{\langle a, a \rangle} \) is identical to the norm \( \|a\|_2 \) defined in the previous section. In particular, \( \ell^2(\mathbb{N}) \) is complete and thus indeed a Hilbert space.

A vector \( f \in \mathfrak{H} \) is called normalized or a unit vector if \( \|f\| = 1 \). Two vectors \( f, g \in \mathfrak{H} \) are called orthogonal or perpendicular \( (f \perp g) \) if \( \langle f, g \rangle = 0 \) and parallel if one is a multiple of the other.

If \( f \) and \( g \) are orthogonal, we have the Pythagorean theorem:
\[
\|f + g\|^2 = \|f\|^2 + \|g\|^2, \quad f \perp g, \tag{1.44}
\]
which is one line of computation (do it!).

Suppose \( u \) is a unit vector. Then the projection of \( f \) in the direction of \( u \) is given by
\[
f_\parallel = \langle u, f \rangle u, \tag{1.45}
\]
and \( f_\perp \), defined via
\[
f_\perp = f - \langle u, f \rangle u, \tag{1.46}
\]
is perpendicular to \( u \) since \( \langle u, f_\perp \rangle = \langle u, f - \langle u, f \rangle u \rangle = \langle u, f \rangle - \langle u, f \rangle \langle u, u \rangle = 0 \).

Taking any other vector parallel to \( u \), we obtain from (1.44)
\[
\|f - \alpha u\|^2 = \|f_\perp + (f_\parallel - \alpha u)\|^2 = \|f_\perp\|^2 + |\langle u, f \rangle - \alpha|^2 \tag{1.47}
\]
and hence \( f_\parallel = \langle u, f \rangle u \) is the unique vector parallel to \( u \) which is closest to \( f \).

As a first consequence we obtain the Cauchy–Schwarz–Bunjakowski inequality:

**Theorem 1.28** (Cauchy–Schwarz–Bunjakowski). Let \( \mathfrak{H}_0 \) be an inner product space. Then for every \( f, g \in \mathfrak{H}_0 \) we have
\[
|\langle f, g \rangle| \leq \|f\| \|g\| \tag{1.48}
\]
with equality if and only if \( f \) and \( g \) are parallel.

**Proof.** It suffices to prove the case \( \|f\| = 1 \). But then the claim follows from \( \|f\|^2 = |\langle g, f \rangle|^2 + \|f_\perp\|^2 \).

Note that the Cauchy–Schwarz inequality implies that the scalar product is continuous in both variables; that is, if \( f_n \to f \) and \( g_n \to g \), we have \( \langle f_n, g_n \rangle \to \langle f, g \rangle \).

As another consequence we infer that the map \( \|\cdot\| \) is indeed a norm. In fact,

\[
\|f + g\|^2 = \|f\|^2 + \langle f, g \rangle + \langle g, f \rangle + \|g\|^2 \leq (\|f\| + \|g\|)^2.
\]

(1.49)

But let us return to \( C(I) \). Can we find a scalar product which has the maximum norm as associated norm? Unfortunately the answer is no! The reason is that the maximum norm does not satisfy the parallelogram law (Problem 1.24).

**Theorem 1.29** (Jordan–von Neumann). A norm is associated with a scalar product if and only if the **parallelogram law**

\[
\|f + g\|^2 + \|f - g\|^2 = 2\|f\|^2 + 2\|g\|^2
\]

holds.

In this case the scalar product can be recovered from its norm by virtue of the **polarization identity**

\[
\langle f, g \rangle = \frac{1}{4} \left( \|f + g\|^2 - \|f - g\|^2 + i\|f - ig\|^2 - i\|f + ig\|^2 \right).
\]

(1.51)

**Proof.** If an inner product space is given, verification of the parallelogram law and the polarization identity is straightforward (Problem 1.26).

To show the converse, we define

\[
s(f, g) = \frac{1}{4} \left( \|f + g\|^2 - \|f - g\|^2 + i\|f - ig\|^2 - i\|f + ig\|^2 \right).
\]

Then \( s(f, f) = \|f\|^2 \) and \( s(f, g) = s(g, f)^* \) are straightforward to check. Moreover, another straightforward computation using the parallelogram law shows

\[
s(f, g) + s(f, h) = 2s(f, \frac{g + h}{2}).
\]

Now choosing \( h = 0 \) (and using \( s(f, 0) = 0 \)) shows \( s(f, g) = 2s(f, \frac{g}{2}) \) and thus \( s(f, g) + s(f, h) = s(f, g + h) \). Furthermore, by induction we infer \( \frac{m}{2^n} s(f, g) = s(f, \frac{m}{2^n} g) \); that is, \( \alpha s(f, g) = s(f, \alpha g) \) for a dense set of positive rational numbers \( \alpha \). By continuity (which follows from continuity of the norm) this holds for all \( \alpha \geq 0 \) and \( s(f, -g) = -s(f, g) \), respectively, \( s(f, ig) = i s(f, g) \), finishes the proof.  \( \square \)
1.3. The geometry of Hilbert spaces

Note that the parallelogram law and the polarization identity even hold for sesquilinear forms (Problem 1.26).

But how do we define a scalar product on $\mathbb{C}^I$? One possibility is
\[
\langle f, g \rangle = \int_a^b f^*(x)g(x)dx.
\] (1.52)
The corresponding inner product space is denoted by $L^2_{\text{cont}}(I)$. Note that we have
\[
\|f\| \leq \sqrt{|b-a|}\|f\|_{\infty}
\] (1.53)
and hence the maximum norm is stronger than the $L^2_{\text{cont}}$ norm.

Suppose we have two norms \(\|\cdot\|_1\) and \(\|\cdot\|_2\) on a vector space $X$. Then \(\|\cdot\|_2\) is said to be stronger than \(\|\cdot\|_1\) if there is a constant $m > 0$ such that
\[
\|f\|_1 \leq m\|f\|_2.
\] (1.54)
It is straightforward to check the following.

**Lemma 1.30.** If \(\|\cdot\|_2\) is stronger than \(\|\cdot\|_1\), then every \(\|\cdot\|_2\) Cauchy sequence is also a \(\|\cdot\|_1\) Cauchy sequence.

Hence if a function $F : X \to Y$ is continuous in $(X, \|\cdot\|_1)$, it is also continuous in $(X, \|\cdot\|_2)$, and if a set is dense in $(X, \|\cdot\|_2)$, it is also dense in $(X, \|\cdot\|_1)$.

In particular, $L^2_{\text{cont}}$ is separable. But is it also complete? Unfortunately the answer is no:

**Example.** Take $I = [0, 2]$ and define
\[
f_n(x) = \begin{cases} 
0, & 0 \leq x \leq 1 - \frac{1}{n}, \\
1 + n(x - 1), & 1 - \frac{1}{n} \leq x \leq 1, \\
1, & 1 \leq x \leq 2.
\end{cases}
\] (1.55)
Then $f_n(x)$ is a Cauchy sequence in $L^2_{\text{cont}}$, but there is no limit in $L^2_{\text{cont}}$! Clearly, the limit should be the step function which is 0 for $0 \leq x < 1$ and 1 for $1 \leq x \leq 2$, but this step function is discontinuous (Problem 1.29)!

**Example.** The previous example indicates that we should consider (1.52) on a larger class of functions, for example on the class of Riemann integrable functions
\[
\mathcal{R}(I) = \{f : I \to \mathbb{C} | f \text{ is Riemann integrable}\}
\]
such that the integral makes sense. While this seems natural it implies another problem: Any function which vanishes outside a set which is negligible for the integral (e.g. finitely many points) has norm zero! That is, \(\|f\|_2 = (\int_I |f(x)|^2dx)^{1/2}\) is only a seminorm on $\mathcal{R}(I)$ (Problem 1.28). To get
a norm we consider \( \mathcal{N}(I) = \{ f \in \mathcal{R}(I) \mid \|f\|_2 = 0 \} \). By homogeneity and the
triangle inequality \( \mathcal{N}(I) \) is a subspace and we can consider
\[ L^2_{Ri} = \mathcal{R}(I)/\mathcal{N}(I). \]
Since \( \|f\|_2 = \|g\|_2 \) for \( f - g \in \mathcal{N}(I) \) we have a norm on \( L^2_{Ri} \). Moreover,
since this norm inherits the parallelogram law we even have an inner product space. However, this space will not be complete unless we replace the
Riemann by the Lebesgue integral. Hence will not pursue this further until we have the Lebesgue integral at our disposal.

This shows that in infinite dimensional vector spaces, different norms
will give rise to different convergent sequences! In fact, the key to solving
problems in infinite dimensional spaces is often finding the right norm! This
is something which cannot happen in the finite dimensional case.

**Theorem 1.31.** If \( X \) is a finite dimensional vector space, then all norms
are equivalent. That is, for any two given norms \( \|\cdot\|_1 \) and \( \|\cdot\|_2 \), there are
positive constants \( m_1 \) and \( m_2 \) such that
\[
\frac{1}{m_2} \|f\|_1 \leq \|f\|_2 \leq m_1 \|f\|_1.
\] (1.56)

**Proof.** Choose a basis \( \{u_j\}_{1 \leq j \leq n} \) such that every \( f \in X \) can be written
as \( f = \sum_j \alpha_j u_j \). Since equivalence of norms is an equivalence relation
(check this!), we can assume that \( \|\cdot\|_2 \) is the usual Euclidean norm: \( \|f\|_2 = \|
\sum_j \alpha_j u_j\|_2 = (\sum_j |\alpha_j|^2)^{1/2} \). Then by the triangle and Cauchy–Schwarz
inequalities,
\[
\|f\|_1 \leq \sum_j |\alpha_j| \|u_j\|_1 \leq \sqrt{\sum_j \|u_j\|_1^2} \|f\|_2
\]
and we can choose \( m_2 = \sqrt{\sum_j \|u_j\|_1^2} \).

In particular, if \( f_n \) is convergent with respect to \( \|\cdot\|_2 \), it is also convergent
with respect to \( \|\cdot\|_1 \). Thus \( \|\cdot\|_1 \) is continuous with respect to \( \|\cdot\|_2 \) and attains
its minimum \( m > 0 \) on the unit sphere \( S = \{u \mid \|u\|_2 = 1\} \) (which is compact
by the Heine–Borel theorem, Theorem 1.17). Now choose \( m_1 = 1/m \). \( \square \)

**Problem 1.22.** Show that the norm in a Hilbert space satisfies \( \|f + g\| = \|f\| + \|g\| \) if and only if \( f = \alpha g, \alpha \geq 0 \), or \( g = 0 \).

**Problem 1.23** (Generalized parallelogram law). Show that, in a Hilbert
space,
\[
\sum_{1 \leq j < k \leq n} \|x_j - x_k\|^2 + \| \sum_{1 \leq j \leq n} x_j \|^2 = n \sum_{1 \leq j \leq n} \|x_j\|^2.
\]
The case \( n = 2 \) is (1.50).
Problem 1.24. Show that the maximum norm on $C[0,1]$ does not satisfy the parallelogram law.

Problem 1.25. In a Banach space, the unit ball is convex by the triangle inequality. A Banach space $X$ is called uniformly convex if for every $\epsilon > 0$ there is some $\delta$ such that $\|x\| \leq 1$, $\|y\| \leq 1$, and $\frac{x+y}{2} \geq 1 - \delta$ imply $\|x - y\| \leq \epsilon$.

Geometrically this implies that if the average of two vectors inside the closed unit ball is close to the boundary, then they must be close to each other.

Show that a Hilbert space is uniformly convex and that one can choose $\delta(\epsilon) = 1 - \sqrt{1 - \frac{\epsilon}{2}}$. Draw the unit ball for $\mathbb{R}^2$ for the norms $\|x\|_1 = |x_1| + |x_2|$, $\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2}$, and $\|x\|_\infty = \max(|x_1|, |x_2|)$. Which of these norms makes $\mathbb{R}^2$ uniformly convex?

(Hint: For the first part, use the parallelogram law.)

Problem 1.26. Suppose $\Omega$ is a vector space. Let $s(f,g)$ be a sesquilinear form on $\Omega$ and $q(f) = s(f,f)$ the associated quadratic form. Prove the parallelogram law

$$q(f + g) + q(f - g) = 2q(f) + 2q(g)$$

and the polarization identity

$$s(f,g) = \frac{1}{4}(q(f + g) - q(f - g) + iq(f - ig) - iq(f + ig)).$$

Show that $s(f,g)$ is symmetric if and only if $q(f)$ is real-valued.

Problem 1.27. A sesquilinear form on an inner product space is called bounded if

$$\|s\| = \sup_{\|f\|=\|g\|=1} |s(f,g)|$$

is finite. Similarly, the associated quadratic form $q$ is bounded if

$$\|q\| = \sup_{\|f\|=1} |q(f)|$$

is finite. Show

$$\|q\| \leq \|s\| \leq 2\|q\|.$$  

(Hint: Use the the polarization identity from the previous problem.)

Problem 1.28. Suppose $\Omega$ is a vector space. Let $s(f,g)$ be a sesquilinear form on $\Omega$ and $q(f) = s(f,f)$ the associated quadratic form. Show that the Cauchy–Schwarz inequality

$$|s(f,g)| \leq q(f)^{1/2}q(g)^{1/2}$$

(1.59)
holds if \( q(f) \geq 0 \). In this case \( q(\cdot)^{1/2} \) satisfies the triangle inequality and hence is a seminorm.

(Hint: Consider \( 0 \leq q(f + \alpha g) = q(f) + 2\text{Re}(\alpha s(f,g)) + |\alpha|^2 q(g) \) and choose \( \alpha = t \frac{s(f,g)^*}{|s(f,g)|} \) with \( t \in \mathbb{R} \).)

Problem 1.29. Prove the claims made about \( f_n \), defined in (1.55), in the last example.

1.4. Completeness

Since \( \mathcal{L}^2_{\text{cont}} \) is not complete, how can we obtain a Hilbert space from it? Well, the answer is simple: take the completion.

If \( X \) is an (incomplete) normed space, consider the set of all Cauchy sequences \( \mathcal{X} \). Call two Cauchy sequences equivalent if their difference converges to zero and denote by \( \bar{X} \) the set of all equivalence classes. It is easy to see that \( \bar{X} \) (and \( \mathcal{X} \)) inherit the vector space structure from \( X \). Moreover,

**Lemma 1.32.** If \( x_n \) is a Cauchy sequence, then \( \|x_n\| \) is also a Cauchy sequence and thus converges.

Consequently, the norm of an equivalence class \( [(x_{n})_{n=1}^{\infty}] \) can be defined by \( \|[(x_{n})_{n=1}^{\infty}]\| = \lim_{n \to \infty} \|x_n\| \) and is independent of the representative (show this!). Thus \( \bar{X} \) is a normed space.

**Theorem 1.33.** \( \bar{X} \) is a Banach space containing \( X \) as a dense subspace if we identify \( x \in X \) with the equivalence class of all sequences converging to \( x \).

**Proof.** (Outline) It remains to show that \( \bar{X} \) is complete. Let \( \xi_n = [(x_{n,j})_{j=1}^{\infty}] \) be a Cauchy sequence in \( \bar{X} \). Then it is not hard to see that \( \xi = [(x_{j,j})_{j=1}^{\infty}] \) is its limit.

Let me remark that the completion \( \bar{X} \) is unique. More precisely, every other complete space which contains \( X \) as a dense subset is isomorphic to \( \bar{X} \). This can for example be seen by showing that the identity map on \( X \) has a unique extension to \( \bar{X} \) (compare Theorem 1.35 below).

In particular, it is no restriction to assume that a normed vector space or an inner product space is complete (note that by continuity of the norm the parallelogram law holds for \( \bar{X} \) if it holds for \( X \)).

**Example.** The completion of the space \( \mathcal{L}^2_{\text{cont}}(I) \) is denoted by \( L^2(I) \). While this defines \( L^2(I) \) uniquely (up to isomorphisms) it is often inconvenient to work with equivalence classes of Cauchy sequences. Hence we will give a different characterization as equivalence classes of square integrable (in the sense of Lebesgue) functions later.
Similarly, we define \( L^p(I), 1 \leq p < \infty \), as the completion of \( C(I) \) with respect to the norm
\[
\|f\|_p = \left( \int_a^b |f(x)|^p \right)^{1/p}.
\]
The only requirement for a norm which is not immediate is the triangle inequality (except for \( p = 1, 2 \)) but this can be shown as for \( \ell^p \) (cf. Problem 1.31).

**Problem 1.30.** For every \( f \in L^1(I) \) we can define its integral
\[
\int_c^d f(x)dx
\]
as the (unique) extension of the corresponding linear functional from \( C(I) \) to \( L^1(I) \) (by Theorem 1.35 below). Show that this integral is linear and satisfies
\[
\int_c^e f(x)dx = \int_c^d f(x)dx + \int_d^e f(x)dx, \quad \int_c^d |f(x)|dx \leq \int_c^d |f(x)|dx.
\]

**Problem 1.31.** Show the Hölder inequality
\[
\|fg\|_1 \leq \|f\|_p \|g\|_q, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad 1 \leq p, q \leq \infty,
\]
and conclude that \( \|\cdot\|_p \) is a norm on \( C(I) \). Also conclude that \( L^p(I) \subseteq L^1(I) \).

### 1.5. Bounded operators

A linear map \( A \) between two normed spaces \( X \) and \( Y \) will be called a (linear) operator
\[
A : \mathfrak{D}(A) \subseteq X \rightarrow Y.
\]
The linear subspace \( \mathfrak{D}(A) \) on which \( A \) is defined is called the domain of \( A \) and is usually required to be dense. The kernel (also null space)
\[
\text{Ker}(A) = \{f \in \mathfrak{D}(A) | Af = 0\} \subseteq X
\]
and range
\[
\text{Ran}(A) = \{Af | f \in \mathfrak{D}(A)\} = A\mathfrak{D}(A) \subseteq Y
\]
are defined as usual. Note that a linear map \( A \) will be continuous if and only if it is continuous at 0, that is, \( x_n \in \mathfrak{D}(A) \rightarrow 0 \) implies \( Ax_n \rightarrow 0 \).

The operator \( A \) is called **bounded** if the operator norm
\[
\|A\| = \sup_{f \in \mathfrak{D}(A), \|f\|_X = 1} \|Af\|_Y
\]
is finite. This says that \( A \) is bounded if the image of the closed unit ball \( B_1(0) \subset X \) is contained in some closed ball \( B_r(0) \subset Y \) of finite radius \( r \).
1. A first look at Banach and Hilbert spaces

(with the smallest radius being the operator norm). Hence \( A \) is bounded if and only if it maps bounded sets to bounded sets.

By construction, a bounded operator is Lipschitz continuous,

\[
\|Af\|_Y \leq \|A\|\|f\|_X, \quad f \in \mathcal{D}(A),
\]

and hence continuous. The converse is also true:

**Theorem 1.34.** An operator \( A \) is bounded if and only if it is continuous.

**Proof.** Suppose \( A \) is continuous but not bounded. Then there is a sequence of unit vectors \( u_n \) such that \( \|Au_n\| \geq n \). Then \( f_n = \frac{1}{n} u_n \) converges to 0 but \( \|Af_n\| \geq 1 \) does not converge to 0.

\( \square \)

In particular, if \( X \) is finite dimensional, then every operator is bounded. Note that in general one and the same operation might be bounded (i.e. continuous) or unbounded, depending on the norm chosen.

**Example.** Consider \( X = \ell^p(\mathbb{N}) \) and let \( a \in \ell^\infty(\mathbb{N}) \). Then the multiplication operator \( A : X \rightarrow X \) defined by

\[
(\lambda \cdot u)_j = \lambda u_j.
\]

Then \( |(\lambda \cdot u)_j| \leq \|\lambda\| \|u_j\| \) shows \( \|\lambda \| \leq \|\lambda\|_\infty \). In fact, we even have \( \|\lambda \| = \|\lambda\|_\infty \) (show this).

**Example.** Consider the vector space of differentiable functions \( X = C^1[0,1] \) and equip it with the norm (cf. Problem 1.34)

\[
\|f\|_{\infty,1} = \max_{x \in [0,1]} |f(x)| + \max_{x \in [0,1]} |f'(x)|.
\]

Let \( Y = C[0,1] \) and observe that the differential operator \( A = \frac{d}{dx} : X \rightarrow Y \) is bounded since

\[
\|Af\|_{\infty} = \max_{x \in [0,1]} |f'(x)| \leq \max_{x \in [0,1]} |f(x)| + \max_{x \in [0,1]} |f'(x)| = \|f\|_{\infty,1}.
\]

However, if we consider \( A = \frac{d}{dx} : \mathcal{D}(A) \subseteq Y \rightarrow Y \) defined on \( \mathcal{D}(A) = C^1[0,1] \), then we have an unbounded operator. Indeed, choose

\[
u_n(x) = \sin(n \pi x)
\]

which is normalized, \( \|\nu_n\|_{\infty} = 1 \), and observe that

\[
Au_n(x) = \nu'_n(x) = n \pi \cos(n \pi x)
\]

is unbounded, \( \|Au_n\|_{\infty} = n \pi \). Note that \( \mathcal{D}(A) \) contains the set of polynomials and thus is dense by the Weierstraß approximation theorem (Theorem 1.26).

\( \diamond \)

If \( A \) is bounded and densely defined, it is no restriction to assume that it is defined on all of \( X \).
Theorem 1.35 (B.L.T. theorem). Let $A : \mathcal{D}(A) \subseteq X \to Y$ be a bounded linear operator and let $Y$ be a Banach space. If $\mathcal{D}(A)$ is dense, there is a unique (continuous) extension of $A$ to $X$ which has the same operator norm.

**Proof.** Since a bounded operator maps Cauchy sequences to Cauchy sequences, this extension can only be given by

$$
\tilde{A} f = \lim_{n \to \infty} A f_n, \quad f_n \in \mathcal{D}(A), \quad f \in X.
$$

To show that this definition is independent of the sequence $f_n \to f$, let $g_n \to f$ be a second sequence and observe

$$
\|A f_n - A g_n\| = \|A(f_n - g_n)\| \leq \|A\| \|f_n - g_n\| \to 0.
$$

Since for $f \in \mathcal{D}(A)$ we can choose $f_n = f$, we see that $\tilde{A} f = A f$ in this case, that is, $\tilde{A}$ is indeed an extension. From continuity of vector addition and scalar multiplication it follows that $\tilde{A}$ is linear. Finally, from continuity of the norm we conclude that the operator norm does not increase. \qed

The set of all bounded linear operators from $X$ to $Y$ is denoted by $\mathfrak{L}(X,Y)$. If $X = Y$, we write $\mathfrak{L}(X,X) = \mathfrak{L}(X)$. An operator in $\mathfrak{L}(X,\mathbb{C})$ is called a **bounded linear functional**, and the space $X^* = \mathfrak{L}(X,\mathbb{C})$ is called the dual space of $X$.

**Example.** Let $X = \ell^p(\mathbb{N})$ and $b \in \ell^q(\mathbb{N})$ where $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$
\ell_b(a) = \sum_{j=1}^{\infty} b_j a_j
$$

is a bounded linear functional satisfying $\|\ell_b\| \leq \|b\|_q$ by Hölder’s inequality. In fact, we even have $\|\ell_b\| = \|b\|_q$ (Problem 4.8).

**Example.** Consider $X = C(I)$. Then for every $x_0 \in I$ the point evaluation $\ell_{x_0}(f) = f(x_0)$ is a bounded linear functional. In fact, $\|\ell_{x_0}\| = 1$ (show this).

However, note that $\ell_{x_0}$ is unbounded on $\mathcal{L}^2_{\text{cont}}(I)$! To see this take $f_n(x) = \sqrt{\frac{3n}{2}} \max(0, 1 - n|x - x_0|)$ which is a triangle shaped peak supported on $[x_0 - n^{-1}, x_0 + n^{-1}]$ and normalized according to $\|f_n\|_2 = 1$ for $n$ sufficiently large such that the support is contained in $I$. Then $\ell_{x_0}(f) = f_n(x_0) = \sqrt{\frac{3n}{2}} \to \infty$. This implies that $\ell_{x_0}$ cannot be extended to the completion of $\mathcal{L}^2_{\text{cont}}(I)$ in a natural way and reflects the fact that the integral cannot see individual points (changing the value of a function at one point does not change its integral). \qed
Example. Consider $X = C(I)$ and let $g$ be some (Riemann or Lebesgue) integrable function. Then
\[\ell_g(f) = \int_a^b g(x)f(x)\,dx\]
is a linear functional with norm
\[\|\ell_g\| = \|g\|_1.\]
First of all note that
\[|\ell_g(f)| \leq \int_a^b |g(x)f(x)|\,dx \leq \|f\|\infty \int_a^b |g(x)|\,dx\]
shows $\|\ell_g\| \leq \|g\|_1$. To see that we have equality consider $f_\varepsilon = g^*/(|g| + \varepsilon)$ and note
\[|\ell_g(f_\varepsilon)| = \int_a^b \frac{|g(x)|^2}{1 + \varepsilon|g(x)|^2}\,dx \geq \int_a^b \frac{|g(x)|^2 - \varepsilon^2}{|g(x)| + \varepsilon}\,dx = \|g\|_1 - (b - a)\varepsilon.
Since $\|f_\varepsilon\| \leq 1$ and $\varepsilon > 0$ is arbitrary this establishes the claim. \hfill \Box

Theorem 1.36. The space $\mathfrak{L}(X,Y)$ together with the operator norm (1.63) is a normed space. It is a Banach space if $Y$ is.

Proof. That (1.63) is indeed a norm is straightforward. If $Y$ is complete and $A_n$ is a Cauchy sequence of operators, then $A_n f$ converges to an element $g$ for every $f$. Define a new operator $A$ via $Af = g$. By continuity of the vector operations, $A$ is linear and by continuity of the norm $\|Af\| = \lim_{n \to \infty} \|A_n f\| \leq \lim_{n \to \infty} (\|A_n\|\|f\|)$, it is bounded. Furthermore, given $\varepsilon > 0$, there is some $N$ such that $\|A_n - A_m\| \leq \varepsilon$ for $n, m \geq N$ and thus $\|A_n f - A_m f\| \leq \varepsilon \|f\|$. Taking the limit $m \to \infty$, we see $\|A_n f - Af\| \leq \varepsilon \|f\|$; that is, $A_n \to A$. \hfill \Box

The Banach space of bounded linear operators $\mathfrak{L}(X)$ even has a multiplication given by composition. Clearly, this multiplication satisfies
\[(A + B)C = AC + BC, \quad A(B + C) = AB + BC, \quad A, B, C \in \mathfrak{L}(X)\] (1.65)
and
\[(AB)C = A(BC), \quad \alpha (AB) = (\alpha A)B = A (\alpha B), \quad \alpha \in \mathbb{C}.\] (1.66)
Moreover, it is easy to see that we have
\[\|AB\| \leq \|A\|\|B\|.\] (1.67)
In other words, $\mathfrak{L}(X)$ is a so-called Banach algebra. However, note that our multiplication is not commutative (unless $X$ is one-dimensional). We even have an identity, the identity operator $\mathbb{I}$, satisfying $\|\mathbb{I}\| = 1$. 
Problem 1.32. Consider $X = \mathbb{C}^n$ and let $A : X \to X$ be a matrix. Equip $X$ with the norm (show that this is a norm)

$$\|x\|_\infty = \max_{1 \leq j \leq n} |x_j|$$

and compute the operator norm $\|A\|$ with respect to this matrix in terms of the matrix entries. Do the same with respect to the norm

$$\|x\|_1 = \sum_{1 \leq j \leq n} |x_j|.$$ 

Problem 1.33. Show that the integral operator

$$(Kf)(x) = \int_0^1 K(x, y)f(y)dy,$$

where $K(x, y) \in C([0, 1] \times [0, 1])$, defined on $D(K) = C[0, 1]$, is a bounded operator both in $X = C[0, 1]$ (max norm) and $X = L^2_{\text{cont}}(0, 1)$. Show that the norm in the $X = C[0, 1]$ case is given by

$$\|K\| = \max_{x \in [0, 1]} \int_0^1 |K(x, y)|dy.$$ 

Problem 1.34. Let $I$ be a compact interval. Show that the set of differentiable functions $C^1(I)$ becomes a Banach space if we set $\|f\|_{\infty, 1} = \max_{x \in I} |f(x)| + \max_{x \in I} |f'(x)|$.

Problem 1.35. Show that $\|AB\| \leq \|A\|\|B\|$ for every $A, B \in \mathfrak{L}(X)$. Conclude that the multiplication is continuous: $A_n \to A$ and $B_n \to B$ imply $A_nB_n \to AB$.

Problem 1.36. Let $A \in \mathfrak{L}(X)$ be a bijection. Show

$$\|A^{-1}\|^{-1} = \inf_{f \in X, \|f\| = 1} \|Af\|.$$ 

Problem 1.37. Let

$$f(z) = \sum_{j=0}^{\infty} f_j z^j, \quad |z| < R,$$

be a convergent power series with convergence radius $R > 0$. Suppose $A$ is a bounded operator with $\|A\| < R$. Show that

$$f(A) = \sum_{j=0}^{\infty} f_j A^j$$

eexists and defines a bounded linear operator. Moreover, if $f$ and $g$ are two such functions and $\alpha \in \mathbb{C}$, then

$$(f + g)(A) = f(A) + g(A), \quad (\alpha f)(A) = \alpha f(A), \quad (fg)(A) = f(A)g(A).$$

(Hint: Problem 1.12.)
Problem 1.38. Show that a linear map $\ell : X \to \mathbb{C}$ is continuous if and only if its kernel is closed. (Hint: If $\ell$ is not continuous, we can find a sequence of normalized vectors $x_n$ with $|\ell(x_n)| \to \infty$ and a vector $x$ with $\ell(x) = 1$.)

1.6. Sums and quotients of Banach spaces

Given two Banach spaces $X_1$ and $X_2$ we can define their (direct) sum $X = X_1 \oplus X_2$ as the cartesian product $X_1 \times X_2$ together with the norm $\|(x_1, x_2)\| = \|x_1\| + \|x_2\|$. In fact, since all norms on $\mathbb{R}^2$ are equivalent (Theorem 1.31), we could as well take $\|(x_1, x_2)\| = (\|x_1\|^p + \|x_2\|^p)^{1/p}$ or $\|(x_1, x_2)\| = \max(\|x_1\|, \|x_2\|)$. In particular, in the case of Hilbert spaces the choice $p = 2$ will ensure that $X$ is again a Hilbert space. Note that $X_1$ and $X_2$ can be regarded as subspaces of $X_1 \times X_2$ by virtue of the obvious embeddings $x_1 \mapsto (x_1, 0)$ and $x_2 \mapsto (0, x_2)$. It is straightforward to show that $X$ is again a Banach space and to generalize this concept to finitely many spaces (Problem 1.39).

Moreover, given a closed subspace $M$ of a Banach space $X$ we can define the quotient space $X/M$ as the set of all equivalence classes $[x] = x + M$ with respect to the equivalence relation $x \equiv y$ if $x - y \in M$. It is straightforward to see that $X/M$ is a vector space when defining $[x] + [y] = [x + y]$ and $\alpha[x] = [\alpha x]$ (show that these definitions are independent of the representative of the equivalence class).

Lemma 1.37. Let $M$ be a closed subspace of a Banach space $X$. Then $X/M$ together with the norm

$$\|[x]\| = \inf_{y \in M} \|x + y\|. \quad (1.68)$$

is a Banach space.

Proof. First of all we need to show that (1.68) is indeed a norm. If $\|[x]\| = 0$ we must have a sequence $y_j \in M$ with $y_j \to -x$ and since $M$ is closed we conclude $x \in M$, that is $[x] = [0]$ as required. To see $\|\alpha[x]\| = |\alpha|\|[x]\|$ we use again the definition

$$\|\alpha[x]\| = \|\alpha x\| = \inf_{y \in M} \|\alpha x + y\| = \inf_{y \in M} \|\alpha x + \alpha y\|$$

$$= |\alpha| \inf_{y \in M} \|x + y\| = |\alpha|\|[x]\|. \quad (1.68)$$

The triangle inequality follows with a similar argument and is left as an exercise.

Thus (1.68) is a norm and it remains to show that $X/M$ is complete. To this end let $[x_n]$ be a Cauchy sequence. Since it suffices to show that some subsequence has a limit, we can assume $\|[x_{n+1}] - [x_n]\| < 2^{-n}$ without loss of generality. Moreover, by definition of (1.68) we can chose the representatives
1.7. Spaces of continuous and differentiable functions

$x_n$ such that $\|x_{n+1} - x_n\| < 2^{-n}$ (start with $x_1$ and then chose the remaining ones inductively). By construction $x_n$ is a Cauchy sequence which has a limit $x \in X$ since $X$ is complete. Moreover, by $\|x_n - x\| = \|x_{n+1} - x_n\| \leq 2^{-n}$ we see that $[x]$ is the limit of $[x_n]$. □

Note that by $\|\cdot\| \leq \|\cdot\|_\infty$ the quotient map $\pi : X \to X/M$, $x \mapsto [x]$ is bounded with norm at most one. As a small application we note:

**Corollary 1.38.** Let $X$ be a Banach space and let $M, N \subseteq X$ be two closed subspaces with one of them, say $N$, finite dimensional. Then $M + N$ is also closed.

**Proof.** If $\pi : X \to X/M$ denotes the quotient map, then $M + N = \pi^{-1}(\pi(N))$. Moreover, since $\pi(N)$ is finite dimensional it is closed and hence $\pi^{-1}(\pi(N))$ is closed by continuity. □

**Problem 1.39.** Let $X_j$, $j = 1, \ldots, n$, be Banach spaces. Let $X$ be the cartesian product $X_1 \times \cdots \times X_n$ together with the norm

$$\|(x_1, \ldots, x_n)\|_p = \begin{cases} \left( \sum_{j=1}^n \|x_j\|^p \right)^{1/p}, & 1 \leq p < \infty, \\ \max_{j=1,\ldots,n} \|x_j\|, & p = \infty. \end{cases}$$

Show that $X$ is a Banach space. Show that all norms are equivalent.

**Problem 1.40.** Suppose $A \in \mathcal{L}(X, Y)$. Show that $\text{Ker}(A)$ is closed. Show that $A$ is well defined on $X/\text{Ker}(A)$ and that this new operator is again bounded (with the same norm) and injective.

1.7. Spaces of continuous and differentiable functions

In this section we introduce a few further sets of continuous and differentiable functions which are of interest in applications.

First, for any set $U \subseteq \mathbb{R}^m$ the set of all bounded continuous functions $C_b(U)$ together with the sup norm

$$\|f\|_\infty = \sup_{x \in U} |f(x)|$$

is a Banach space and this can be shown as in Section 1.2. The space of continuous functions with compact support $C_c(U) \subseteq C_b(U)$ is in general not dense and its closure will be denoted by $C_0(U)$. If $U$ is open it can be interpreted as the functions in $C_b(U)$ which vanish at the boundary

$$C_0(U) = \{ f \in C(U) | \forall \varepsilon > 0, \exists K \text{ compact} : |f(x)| < \varepsilon, x \in X \setminus K \}. \tag{1.70}$$

Of course $\mathbb{R}^m$ could be replaced by any topological space up to this point.
Moreover, the above norm can be augmented to handle differentiable functions by considering the space $C_1^b(U)$ of all continuously differentiable functions for which the following norm is finite,

$$
\|f\|_{\infty, 1} = \|f\|_{\infty} + \sum_{j=1}^{m} \|\partial_j f\|_{\infty}
$$

(1.71)

finite, where $\partial_j = \frac{\partial}{\partial x_j}$. Note that $\|\partial_j f\|$ for one $j$ (or all $j$) is not sufficient as it is only a seminorm (it vanishes for every constant function). However, since the sum of seminorms is again a seminorm (Problem 1.41) the above expression defines indeed a norm. It is also not hard to see that $C_1^b(U, \mathbb{C}^n)$ is complete. In fact, let $f^k$ be a Cauchy sequence, then $f^k(x)$ converges uniformly to some continuous function $f(x)$ and the same is true for the partial derivatives $\partial_j f^k(x) \to g_j(x)$. Moreover, since $f^k(x) = f^k(c, x_2, \ldots, x_m) + \int_c^{x_1} \partial_j f^k(t, x_2, \ldots, x_m) dt \to f(x) = f(c, x_2, \ldots, x_m) + \int_c^{x_1} g_j(t, x_2, \ldots, x_m)$ we obtain $\partial_j f(x) = g_j(x)$. The remaining derivatives follow analogously and thus $f^k \to f$ in $C_1^b(U, \mathbb{C}^n)$.

To extend this approach to higher derivatives let $C^k(U)$ be the set of all complex-valued functions which have partial derivatives of order up to $k$. For $f \in C^k(U)$ and $\alpha \in \mathbb{N}_0^n$ we set

$$
\partial_\alpha f = \frac{\partial^{\alpha_1} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}, \quad |\alpha| = \alpha_1 + \cdots + \alpha_n.
$$

(1.72)

An element $\alpha \in \mathbb{N}_0^n$ is called a multi-index and $|\alpha|$ is called its order.

With this notation the above considerations can be easily generalized to higher order derivatives:

**Theorem 1.39.** Let $U \subseteq \mathbb{R}^m$ be open. The space $C^k_b(U)$ of all functions whose partial derivatives up to order $k$ are bounded and continuous form a Banach space with norm

$$
\|f\|_{\infty, k} = \sum_{|\alpha| \leq k} \sup_{x \in U} |\partial_\alpha f(x)|.
$$

(1.73)

An important subspace is $C^k_0(\mathbb{R}^m)$, the set of all functions in $C^k_b(\mathbb{R}^m)$ for which $\lim_{|x| \to \infty} |\partial_\alpha f(x)| = 0$ for all $|\alpha| \leq k$. For any function $f$ not in $C^k_0(\mathbb{R}^m)$ there must be a sequence $|x_j| \to \infty$ and some $\alpha$ such that $|\partial_\alpha f(x_j)| \geq \varepsilon > 0$. But then $\|f - g\|_{\infty, k} \geq \varepsilon$ for every $g$ in $C^k_0(\mathbb{R}^m)$ and thus $C^k_0(\mathbb{R}^m)$ is a closed subspace. In particular, it is a Banach space of its own.

Note that the space $C^k_b(U)$ could be further refined by requiring the highest derivatives to be Hölder continuous. Recall that a function $f : U \to \mathbb{C}$ is Hölder continuous with exponent $\beta$ if

$$
|f(x) - f(y)| \leq A|x - y|^\beta
$$

for some $A > 0$ and all $x, y \in U$. In this case, we can define the Hölder seminorm

$$
\|f\|_{H^\beta} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\beta}.
$$

The set of all Hölder continuous functions $C^{H\beta}_b(U)$ can then be completed to a Banach space with the norm $\|f\|_{H^\beta, \infty}$.

1. A first look at Banach and Hilbert spaces
1.7. Spaces of continuous and differentiable functions

\( C \) is called uniformly **Hölder continuous** with exponent \( \gamma \in (0, 1] \) if

\[
[f]_\gamma = \sup_{x \neq y \in U} \frac{|f(x) - f(y)|}{|x - y|^{\gamma}}
\]

is finite. Clearly, any Hölder continuous function is continuous and, in the special case \( \gamma = 1 \), we obtain the **Lipschitz continuous** functions.

It is easy to verify that this is a seminorm and that the corresponding space is complete.

**Theorem 1.40.** The space \( C^{k,\gamma}_b(U) \) of all functions whose partial derivatives up to order \( k \) are bounded and Hölder continuous with exponent \( \gamma \in (0, 1] \) form a Banach space with norm

\[
\|f\|_{\infty,k,\gamma} = \|f\|_{\infty,k} + \sum_{|\alpha| = k} [\partial_{\alpha}f]_\gamma.
\]

Note that by the mean value theorem all derivatives up to order lower than \( k \) are automatically Lipschitz continuous.

While the above spaces are able to cover a wide variety of situations, there are still cases where the above definitions are not suitable. However such cases cannot be covered with norm and we will postpone this to Section 4.7.

Note that in all the above spaces we could replace complex-valued by \( \mathbb{C}^n \)-valued functions.

**Problem 1.41.** Suppose \( X \) is a vector space and \( \|\cdot\|_j, 1 \leq j \leq n, \) is a finite family of seminorms. Show that \( \|x\| = \sum_{j=1}^{n} \|x\|_j \) is a seminorm. It is a norm if and only if \( \|x\|_j = 0 \) for all \( j \) implies \( x = 0 \).

**Problem 1.42.** Show that \( C_b(U) \) is a Banach space when equipped with the sup norm. Show that \( \overline{C_c(U)} = C_0(U) \). (Hint: The function \( m_\varepsilon(z) = \text{sign}(z) \max(0, |z| - \varepsilon) \in C(\mathbb{C}) \) might be useful.)

**Problem 1.43.** Suppose \( U \) is bounded. Show \( C^{k,\gamma_2}_b(U) \subseteq C^{k,\gamma_1}_b(U) \subseteq C^k_b(U) \) for \( 0 < \gamma_1 < \gamma_2 \leq 1 \).

**Problem 1.44.** Show that the product of two bounded Hölder continuous functions is again Hölder continuous:

\[
[fg]_\gamma \leq \|f\|_{\infty} [g]_\gamma + [f]_\gamma \|g\|_{\infty}.
\]

**Problem 1.45.** Show that if \( d \) is a pseudometric, then so is \( \frac{d}{1+d} \). (Hint: Note that \( f(x) = x/(1 + x) \) is concave.)
Chapter 2

Hilbert spaces

2.1. Orthonormal bases

In this section we will investigate orthonormal series and you will notice hardly any difference between the finite and infinite dimensional cases. Throughout this chapter $\mathcal{H}$ will be a Hilbert space.

As our first task, let us generalize the projection into the direction of one vector:

A set of vectors $\{u_j\}$ is called an orthonormal set if $\langle u_j, u_k \rangle = 0$ for $j \neq k$ and $\langle u_j, u_j \rangle = 1$. Note that every orthonormal set is linearly independent (show this).

Lemma 2.1. Suppose $\{u_j\}_{j=1}^n$ is an orthonormal set. Then every $f \in \mathcal{H}$ can be written as

$$f = f_\parallel + f_\perp, \quad f_\parallel = \sum_{j=1}^n \langle u_j, f \rangle u_j, \quad (2.1)$$

where $f_\parallel$ and $f_\perp$ are orthogonal. Moreover, $\langle u_j, f_\perp \rangle = 0$ for all $1 \leq j \leq n$. In particular,

$$\|f\|^2 = \sum_{j=1}^n |\langle u_j, f \rangle|^2 + \|f_\perp\|^2. \quad (2.2)$$

Moreover, every $\hat{f}$ in the span of $\{u_j\}_{j=1}^n$ satisfies

$$\|f - \hat{f}\| \geq \|f_\perp\| \quad (2.3)$$

with equality holding if and only if $\hat{f} = f_\parallel$. In other words, $f_\parallel$ is uniquely characterized as the vector in the span of $\{u_j\}_{j=1}^n$ closest to $f$.  

49
Proof. A straightforward calculation shows $\langle u_j, f - f\rangle = 0$ and hence $f\|$ and $f_\perp = f - f\|$ are orthogonal. The formula for the norm follows by applying (1.44) iteratively.

Now, fix a vector
\[
\hat{f} = \sum_{j=1}^{n} \alpha_j u_j
\]
in the span of $\{u_j\}_{j=1}^{n}$. Then one computes
\[
\|f - \hat{f}\|^2 = \|f\|^2 + \|f_\perp - \hat{f}\|^2 = \|f_\perp\|^2 + \|f - \hat{f}\|^2 \\
= \|f_\perp\|^2 + \sum_{j=1}^{n} |\alpha_j - \langle u_j, f \rangle|^2
\]
from which the last claim follows. \qed

From (2.2) we obtain Bessel’s inequality
\[
\sum_{j=1}^{n} |\langle u_j, f \rangle|^2 \leq \|f\|^2
\]
with equality holding if and only if $f$ lies in the span of $\{u_j\}_{j=1}^{n}$.

Of course, since we cannot assume $H$ to be a finite dimensional vector space, we need to generalize Lemma 2.1 to arbitrary orthonormal sets $\{u_j\}_{j \in J}$. We start by assuming that $J$ is countable. Then Bessel’s inequality (2.4) shows that
\[
\sum_{j \in J} |\langle u_j, f \rangle|^2
\]
converges absolutely. Moreover, for any finite subset $K \subset J$ we have
\[
\| \sum_{j \in K} \langle u_j, f \rangle u_j \|^2 = \sum_{j \in K} |\langle u_j, f \rangle|^2
\]
by the Pythagorean theorem and thus $\sum_{j \in J} \langle u_j, f \rangle u_j$ is a Cauchy sequence if and only if $\sum_{j \in J} |\langle u_j, f \rangle|^2$ is. Now let $J$ be arbitrary. Again, Bessel’s inequality shows that for any given $\varepsilon > 0$ there are at most finitely many $j$ for which $|\langle u_j, f \rangle| \geq \varepsilon$ (namely at most $\|f\|/\varepsilon$). Hence there are at most countably many $j$ for which $|\langle u_j, f \rangle| > 0$. Thus it follows that
\[
\sum_{j \in J} |\langle u_j, f \rangle|^2
\]
is well defined (as a countable sum over the nonzero terms) and (by completeness) so is
\[
\sum_{j \in J} \langle u_j, f \rangle u_j.
\]
Furthermore, it is also independent of the order of summation.
In particular, by continuity of the scalar product we see that Lemma 2.1 can be generalized to arbitrary orthonormal sets.

**Theorem 2.2.** Suppose \( \{u_j\}_{j \in J} \) is an orthonormal set in a Hilbert space \( \mathcal{H} \). Then every \( f \in \mathcal{H} \) can be written as

\[
f = f_\parallel + f_\perp, \quad f_\parallel = \sum_{j \in J} \langle u_j, f \rangle u_j,
\]

where \( f_\parallel \) and \( f_\perp \) are orthogonal. Moreover, \( \langle u_j, f_\perp \rangle = 0 \) for all \( j \in J \). In particular,

\[
\|f\|^2 = \sum_{j \in J} |\langle u_j, f \rangle|^2 + \|f_\perp\|^2. \tag{2.10}
\]

Furthermore, every \( \hat{f} \in \text{span}\{u_j\}_{j \in J} \) satisfies

\[
\|f - \hat{f}\| \geq \|f_\parallel\| \tag{2.11}
\]

with equality holding if and only if \( \hat{f} = f_\parallel \). In other words, \( f_\parallel \) is uniquely characterized as the vector in \( \text{span}\{u_j\}_{j \in J} \) closest to \( f \).

**Proof.** The first part follows as in Lemma 2.1 using continuity of the scalar product. The same is true for the last part except for the fact that every \( f \in \text{span}\{u_j\}_{j \in J} \) can be written as \( f = \sum_{j \in J} \alpha_j u_j \) (i.e., \( f = f_\parallel \)). To see this, let \( f_n \in \text{span}\{u_j\}_{j \in J} \) converge to \( f \). Then \( \|f - f_n\|^2 = \|f_\parallel - f_n\|^2 + \|f_\perp\|^2 \to 0 \) implies \( f_n \to f_\parallel \) and \( f_\perp = 0 \). □

Note that from Bessel’s inequality (which of course still holds), it follows that the map \( f \to f_\parallel \) is continuous.

Of course we are particularly interested in the case where every \( f \in \mathcal{H} \) can be written as \( \sum_{j \in J} \langle u_j, f \rangle u_j \). In this case we will call the orthonormal set \( \{u_j\}_{j \in J} \) an **orthonormal basis** (ONB).

If \( \mathcal{H} \) is separable it is easy to construct an orthonormal basis. In fact, if \( \mathcal{H} \) is separable, then there exists a countable total set \( \{f_j\}_{j=1}^N \). Here \( N \in \mathbb{N} \) if \( \mathcal{H} \) is finite dimensional and \( N = \infty \) otherwise. After throwing away some vectors, we can assume that \( f_{n+1} \) cannot be expressed as a linear combination of the vectors \( f_1, \ldots, f_n \). Now we can construct an orthonormal set as follows: We begin by normalizing \( f_1 \):

\[
u_1 = \frac{f_1}{\|f_1\|}. \tag{2.12}
\]

Next we take \( f_2 \) and remove the component parallel to \( u_1 \) and normalize again:

\[
u_2 = \frac{f_2 - \langle u_1, f_2 \rangle u_1}{\|f_2 - \langle u_1, f_2 \rangle u_1\|}. \tag{2.13}
\]
Proceeding like this, we define recursively

\[ u_n = \frac{f_n - \sum_{j=1}^{n-1} \langle u_j, f_n \rangle u_j}{\|f_n - \sum_{j=1}^{n-1} \langle u_j, f_n \rangle u_j\|}. \]  

(2.14)

This procedure is known as **Gram–Schmidt orthogonalization**. Hence we obtain an orthonormal set \( \{u_j\}_{j=1}^N \) such that \( \text{span}\{u_j\}_{j=1}^n = \text{span}\{f_j\}_{j=1}^n \) for any finite \( n \) and thus also for \( n = N \) (if \( N = \infty \)). Since \( \{f_j\}_{j=1}^N \) is total, so is \( \{u_j\}_{j=1}^N \). Now suppose there is some \( f = f_\parallel + f_\perp \in \mathcal{H} \) for which \( f_\perp \neq 0 \). Since \( \{u_j\}_{j=1}^N \) is total, we can find a \( \hat{f} \) in its span such that \( \|f - \hat{f}\| < \|f_\perp\| \), contradicting (2.11). Hence we infer that \( \{u_j\}_{j=1}^N \) is an orthonormal basis.

**Theorem 2.3.** Every separable Hilbert space has a countable orthonormal basis.

**Example.** In \( \mathcal{L}^2_{\text{cont}}(-1,1) \), we can orthogonalize the polynomial \( f_n(x) = x^n \) (which are total by the Weierstraß approximation theorem — Theorem 1.26). The resulting polynomials are up to a normalization equal to the Legendre polynomials

\[ P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{3x^2 - 1}{2}, \quad \ldots \]  

(2.15)

(which are normalized such that \( P_n(1) = 1 \)).

**Example.** The set of functions

\[ u_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}, \quad n \in \mathbb{Z}, \]  

(2.16)

forms an orthonormal basis for \( \mathcal{H} = \mathcal{L}^2_{\text{cont}}(0,2\pi) \). The corresponding orthogonal expansion is just the ordinary Fourier series. (That these functions are total will follow from the Stone–Weierstraß theorem — Problem 6.16.)

The following equivalent properties also characterize a basis.

**Theorem 2.4.** For an orthonormal set \( \{u_j\}_{j \in J} \) in a Hilbert space \( \mathcal{H} \), the following conditions are equivalent:

(i) \( \{u_j\}_{j \in J} \) is a maximal orthogonal set.

(ii) For every vector \( f \in \mathcal{H} \) we have

\[ f = \sum_{j \in J} \langle u_j, f \rangle u_j. \]  

(2.17)

(iii) For every vector \( f \in \mathcal{H} \) we have Parseval’s relation

\[ \|f\|^2 = \sum_{j \in J} |\langle u_j, f \rangle|^2. \]  

(2.18)

(iv) \( \langle u_j, f \rangle = 0 \) for all \( j \in J \) implies \( f = 0 \).
2.1. Orthonormal bases

Proof. We will use the notation from Theorem 2.2.
(i) ⇒ (ii): If \( f_\perp \neq 0 \), then we can normalize \( f_\perp \) to obtain a unit vector \( \tilde{f}_\perp \) which is orthogonal to all vectors \( u_j \). But then \( \{u_j\}_{j \in J} \cup \{f_\perp\} \) would be a larger orthonormal set, contradicting the maximality of \( \{u_j\}_{j \in J} \).
(ii) ⇒ (iii): This follows since (ii) implies \( f_\perp = 0 \).
(iii) ⇒ (iv): If \( \langle f, u_j \rangle = 0 \) for all \( j \in J \), we conclude \( \|f\|^2 = 0 \) and hence \( f = 0 \).
(iv) ⇒ (i): If \( \{u_j\}_{j \in J} \) were not maximal, there would be a unit vector \( g \) such that \( \{u_j\}_{j \in J} \cup \{g\} \) is a larger orthonormal set. But \( \langle u_j, g \rangle = 0 \) for all \( j \in J \) implies \( g = 0 \) by (iv), a contradiction. \( \square \)

By continuity of the norm it suffices to check (iii), and hence also (ii), for \( f \) in a dense set. In fact, by the inverse triangle inequality for \( \ell^2(\mathbb{N}) \) and the Bessel inequality we have

\[
\left| \sum_{j \in J} |\langle u_j, f \rangle|^2 - \sum_{j \in J} |\langle u_j, g \rangle|^2 \right| \leq \sqrt{\sum_{j \in J} |\langle u_j, f - g \rangle|^2} \sqrt{\sum_{j \in J} |\langle u_j, f + g \rangle|^2} \\
\leq \|f - g\| \|f + g\|
\]

(2.19)

implying \( \sum_{j \in J} |\langle u_j, f_n \rangle|^2 \to \sum_{j \in J} |\langle u_j, f \rangle|^2 \) if \( f_n \to f \).

It is not surprising that if there is one countable basis, then it follows that every other basis is countable as well.

Theorem 2.5. In a Hilbert space \( \mathcal{H} \) every orthonormal basis has the same cardinality.

Proof. Let \( \{u_j\}_{j \in J} \) and \( \{v_k\}_{k \in K} \) be two orthonormal bases. We first look at the case where one of them, say the first, is finite dimensional: \( J = \{1, \ldots, n\} \). Suppose the other basis has at least \( n \) elements \( \{1, \ldots, n\} \subseteq K \). Then \( v_k = \sum_{j=1}^n U_{k,j} u_j \), where \( U_{k,j} = \langle u_j, v_k \rangle \). By \( \delta_{j,k} = \langle v_j, v_k \rangle = \sum_{l=1}^n U_{j,l}^* U_{k,l} \) we see \( u_j = \sum_{k=1}^n U_{k,j}^* v_k \) showing that \( K \) cannot have more than \( n \) elements.

Now let us turn to the case where both \( J \) and \( K \) are infinite. Set \( K_j = \{k \in K | \langle v_k, u_j \rangle \neq 0\} \). Since these are the expansion coefficients of \( u_j \) with respect to \( \{v_k\}_{k \in K} \), this set is countable (and nonempty). Hence the set \( \hat{K} = \bigcup_{j \in J} K_j \) satisfies \( |\hat{K}| \leq |J \times \mathbb{N}| = |J| \) (Theorem A.9) But \( k \in K \setminus \hat{K} \) implies \( v_k = 0 \) and hence \( \hat{K} = K \). So \( |K| \leq |J| \) and reversing the roles of \( J \) and \( K \) shows \( |K| = |J| \). \( \square \)

The cardinality of an orthonormal basis is also called the Hilbert space dimension of \( \mathcal{H} \).

It even turns out that, up to unitary equivalence, there is only one separable infinite dimensional Hilbert space:
A bijective linear operator $U \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is called \textbf{unitary} if $U$ preserves scalar products:

$$\langle Ug, Uf \rangle_2 = \langle g, f \rangle_1, \quad g, f \in \mathcal{H}_1. \quad (2.20)$$

By the polarization identity, (1.51) this is the case if and only if $U$ preserves norms: $\|Uf\|_2 = \|f\|_1$ for all $f \in \mathcal{H}_1$ (note the a norm preserving linear operator is automatically injective). The two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ are called \textbf{unitarily equivalent} in this case.

Let $\mathcal{H}$ be an infinite dimensional Hilbert space and let $\{u_j\}_{j \in \mathbb{N}}$ be any orthogonal basis. Then the map $U : \mathcal{H} \to \ell^2(\mathbb{N}), f \mapsto (\langle u_j, f \rangle)_{j \in \mathbb{N}}$ is unitary (by Theorem 2.4 (ii) it is onto and by (iii) it is norm preserving). In particular,

\begin{theorem}
Any separable infinite dimensional Hilbert space is unitarily equivalent to $\ell^2(\mathbb{N})$.
\end{theorem}

Of course the same argument shows that every finite dimensional Hilbert space of dimension $n$ is unitarily equivalent to $\mathbb{C}^n$ with the usual scalar product.

Finally we briefly turn to the case where $\mathcal{H}$ is not separable.

\begin{theorem}
Every Hilbert space has an orthonormal basis.
\end{theorem}

\begin{proof}
To prove this we need to resort to Zorn’s lemma (see Appendix A): The collection of all orthonormal sets in $\mathcal{H}$ can be partially ordered by inclusion. Moreover, every linearly ordered chain has an upper bound (the union of all sets in the chain). Hence a fundamental result from axiomatic set theory, Zorn’s lemma, implies the existence of a maximal element, that is, an orthonormal set which is not a proper subset of every other orthonormal set. \hfill \Box
\end{proof}

Hence, if $\{u_j\}_{j \in J}$ is an orthogonal basis, we can show that $\mathcal{H}$ is unitarily equivalent to $\ell^2(J)$ and, by prescribing $J$, we can find a Hilbert space of any given dimension. Here $\ell^2(J)$ is the set of all complex valued functions $(a_j)_{j \in J}$ where at most countably many values are nonzero and $\sum_{j \in J} |a_j|^2 < \infty$.

\begin{problem}
Given some vectors $f_1, \ldots, f_n$ we define their \textbf{Gram determinant} as

$$\Gamma(f_1, \ldots, f_n) = \det (\langle f_j, f_k \rangle)_{1 \leq j, k \leq n}.$$ 

Show that the Gram determinant is nonzero if and only if the vectors are linearly independent. Moreover, show that in this case

$$\text{dist}(g, \text{span}\{f_1, \ldots, f_n\})^2 = \frac{\Gamma(f_1, \ldots, f_n, g)}{\Gamma(f_1, \ldots, f_n)}.$$ 

(Hint: How does $\Gamma$ change when you apply the Gram-Schmidt procedure?)
\end{problem}
Problem 2.2. Let \( \{u_j\} \) be some orthonormal basis. Show that a bounded linear operator \( A \) is uniquely determined by its matrix elements \( A_{jk} = \langle u_j, Au_k \rangle \) with respect to this basis.

### 2.2. The projection theorem and the Riesz lemma

Let \( M \subseteq \mathfrak{H} \) be a subset. Then \( M^\perp = \{ f | \langle g, f \rangle = 0, \forall g \in M \} \) is called the **orthogonal complement** of \( M \). By continuity of the scalar product it follows that \( M^\perp \) is a closed linear subspace and by linearity that \( \langle \text{span}(M) \rangle^\perp = M^\perp \). For example, we have \( \mathfrak{H}^\perp = \{ 0 \} \) since any vector in \( \mathfrak{H}^\perp \) must be in particular orthogonal to all vectors in some orthonormal basis.

**Theorem 2.8** (Projection theorem). Let \( M \) be a closed linear subspace of a Hilbert space \( \mathfrak{H} \). Then every \( f \in \mathfrak{H} \) can be uniquely written as \( f = f_{\parallel} + f_{\perp} \) with \( f_{\parallel} \in M \) and \( f_{\perp} \in M^\perp \). One writes

\[
M \oplus M^\perp = \mathfrak{H}
\]  
(2.21) in this situation.

**Proof.** Since \( M \) is closed, it is a Hilbert space and has an orthonormal basis \( \{u_j\}_{j \in J} \). Hence the existence part follows from Theorem 2.2. To see uniqueness, suppose there is another decomposition \( f = \tilde{f}_{\parallel} + \tilde{f}_{\perp} \). Then \( f_{\parallel} - \tilde{f}_{\parallel} = \tilde{f}_{\perp} - f_{\perp} \in M \cap M^\perp = \{0\} \).

**Corollary 2.9.** Every orthogonal set \( \{u_j\}_{j \in J} \) can be extended to an orthogonal basis.

**Proof.** Just add an orthogonal basis for \( (\{u_j\}_{j \in J})^\perp \). \( \square \)

Moreover, Theorem 2.8 implies that to every \( f \in \mathfrak{H} \) we can assign a unique vector \( f_{\parallel} \) which is the vector in \( M \) closest to \( f \). The rest, \( f - f_{\parallel} \), lies in \( M^\perp \). The operator \( P_M f = f_{\parallel} \) is called the **orthogonal projection** corresponding to \( M \). Note that we have

\[
P_M^2 = P_M \quad \text{and} \quad \langle P_M g, f \rangle = \langle g, P_M f \rangle
\]  
(2.22) since \( \langle P_M g, f \rangle = \langle g_{\parallel}, f_{\parallel} \rangle = \langle g, P_M f \rangle \). Clearly we have \( P_M f = f - P_M f = f_{\perp} \). Furthermore, (2.22) uniquely characterizes orthogonal projections (Problem 2.5).

Moreover, if \( M \) is a closed subspace, we have \( P_{M^\perp} = I - P_M = I - (I - P_M) = P_M \); that is, \( M^\perp = M \). If \( M \) is an arbitrary subset, we have at least

\[
M^\perp = \overline{\text{span}(M)}.
\]  
(2.23) Note that by \( \mathfrak{H}^\perp = \{0\} \) we see that \( M^\perp = \{0\} \) if and only if \( M \) is total.
Finally we turn to linear functionals, that is, to operators $\ell : \mathcal{H} \to \mathbb{C}$. By the Cauchy–Schwarz inequality we know that $\ell_g : f \mapsto \langle g, f \rangle$ is a bounded linear functional (with norm $\|g\|$). In turns out that, in a Hilbert space, every bounded linear functional can be written in this way.

**Theorem 2.10 (Riesz lemma).** Suppose $\ell$ is a bounded linear functional on a Hilbert space $\mathcal{H}$. Then there is a unique vector $g \in \mathcal{H}$ such that $\ell(f) = \langle g, f \rangle$ for all $f \in \mathcal{H}$.

In other words, a Hilbert space is equivalent to its own dual space $\mathcal{H}^* \cong \mathcal{H}$ via the map $f \mapsto \langle f, \cdot \rangle$ which is a conjugate linear isometric bijection between $\mathcal{H}$ and $\mathcal{H}^*$.

**Proof.** If $\ell \equiv 0$, we can choose $g = 0$. Otherwise $\text{Ker}(\ell) = \{f | \ell(f) = 0\}$ is a proper subspace and we can find a unit vector $\tilde{g} \in \text{Ker}(\ell) \perp$. For every $f \in \mathcal{H}$ we have $\ell(f)\tilde{g} - \ell(\tilde{g})f \in \text{Ker}(\ell)$ and hence

$$0 = \langle \tilde{g}, \ell(f)\tilde{g} - \ell(\tilde{g})f \rangle = \ell(f) - \ell(\tilde{g})\langle \tilde{g}, f \rangle.$$

In other words, we can choose $g = \ell(\tilde{g})^*\tilde{g}$. To see uniqueness, let $g_1$, $g_2$ be two such vectors. Then $\langle g_1 - g_2, f \rangle = \langle g_1, f \rangle - \langle g_2, f \rangle = \ell(f) - \ell(f) = 0$ for every $f \in \mathcal{H}$, which shows $g_1 - g_2 \in \mathcal{H} \perp = \{0\}$. \qed

In particular, this shows that $\mathcal{H}^*$ is again a Hilbert space whose scalar product (in terms of the above identification) is given by $\langle (f, \cdot), (g, \cdot) \rangle_{\mathcal{H}^*} = \langle f, g \rangle^*$.

We can even get a unitary map between $\mathcal{H}$ and $\mathcal{H}^*$ but such a map is not unique. To this end note that every Hilbert space has a conjugation $C$ which generalizes taking the complex conjugate of every coordinate. In fact, choosing an orthonormal basis (and different choices will produce different maps in general) we can set

$$Cf = \sum_{j \in J} \langle u_j, f \rangle^* u_j = \sum_{j \in J} \langle f, u_j \rangle u_j.$$  

Then $C$ is complex linear, isometric $\|Cf\| = \|f\|$, and idempotent $C^2 = I$. Note also $\langle Cf, Cg \rangle = \langle f, g \rangle^*$. As promised, the map $f \to \langle Cf, \cdot \rangle$ is a unitary map from $\mathcal{H}$ to $\mathcal{H}^*$.

**Problem 2.3.** Suppose $U : \mathcal{H} \to \mathcal{H}$ is unitary and $M \subseteq \mathcal{H}$. Show that $UM \perp = (UM) \perp$.

**Problem 2.4.** Show that an orthogonal projection $P_M \neq 0$ has norm one.

**Problem 2.5.** Suppose $P \in \mathcal{L}(\mathcal{H})$ satisfies

$$P^2 = P \quad \text{and} \quad \langle Pf, g \rangle = \langle f, Pg \rangle$$

and set $M = \text{Ran}(P)$. Show
• \( Pf = f \) for \( f \in M \) and \( M \) is closed,
• \( g \in M^\perp \) implies \( Pg \in M^\perp \) and thus \( Pg = 0 \),
and conclude \( P = P_M \). In particular
\[ \mathcal{H} = \ker(P) \oplus \text{ran}(P), \quad \ker(P) = (I - P)\mathcal{H}, \text{ ran}(P) = P\mathcal{H}. \]

2.3. Operators defined via forms

One of the key results about linear maps is that they are uniquely determined once we know the images of some basis vectors. In fact, the matrix elements with respect to some basis uniquely determine a linear map. Clearly this raises the question how this results extends to the infinite dimensional setting. As a first result we show that the Riesz lemma, Theorem 2.10, implies that a bounded operator \( A \) is uniquely determined by its associated sesquilinear form \( \langle f, Ag \rangle \). In fact, there is a one to one correspondence between bounded operators and bounded sesquilinear forms:

**Lemma 2.11.** Suppose \( s \) is a bounded sesquilinear form; that is,
\[ |s(f, g)| \leq C \|f\| \|g\|. \tag{2.24} \]
Then there is a unique bounded operator \( A \) such that
\[ s(f, g) = \langle f, Ag \rangle. \tag{2.25} \]
Moreover, the norm of \( A \) is given by
\[ \|A\| = \sup_{\|f\| = \|g\| = 1} |\langle f, Ag \rangle| \leq C. \tag{2.26} \]

**Proof.** For every \( g \in \mathcal{H} \) we have an associated bounded linear functional \( \ell_g(f) = s(f, g)^* \). By Theorem 2.10 there is a corresponding \( h \in \mathcal{H} \) (depending on \( g \)) such that \( \ell_g(f) = \langle h, f \rangle \), that is \( s(f, g) = \langle f, h \rangle \) and we can define \( A \) via \( Ag = h \). It is not hard to check that \( A \) is linear and from
\[ \|Af\|^2 = \langle Af, Af \rangle = s(Af, f) \leq C\|Af\|\|f\| \]
we infer \( \|Af\| \leq C\|f\| \), which shows that \( A \) is bounded with \( \|A\| \leq C \). Equation (2.26) is left as an exercise (Problem 2.7).

Note that if \( \{u_j\}_{j \in J} \) is some orthogonal basis, then the matrix elements \( A_{j,k} = \langle u_j, Au_k \rangle \) for all \( j, k \in J \) uniquely determine \( \langle f, Ag \rangle \) for arbitrary \( f, g \in \mathcal{H} \) (just expand \( f, g \) with respect to this basis) and thus \( A \) by our theorem.

**Example.** Consider \( \ell^2(\mathbb{N}) \) and let \( A \) be some compact operator. Let \( A_{jk} = \langle \delta^j, K\delta^k \rangle \) its matrix elements such that
\[ (Aa)_j = \sum_{k=1}^{\infty} A_{j,k}a_k. \]
Here the sum converges in $\ell^2(\mathbb{N})$ and hence in particular for every fixed $j$. Moreover, choosing $a^n_k = \alpha_n A_{jk}$ for $k \leq n$ and $a^n_k = 0$ for $k > n$ with $\alpha_n = (\sum_{j=1}^{n} |A_{jk}|^2)^{1/2}$ we see $a^n = |(Aa^n)_j| \leq \|A\| \|a^n\| = \|A\|$ we see that $\sum_{j=1}^{\infty} |A_{jk}|^2 \leq \|A\|$ and the sum is even absolutely convergent.

Moreover, by the polarization identity (Problem 1.26), $A$ is already uniquely determined by its quadratic form $q_A(f) = \langle f, Af \rangle$.

As a first application we introduce the adjoint operator via Lemma 2.11 as the operator associated with the sesquilinear form $s(f, g) = \langle Af, g \rangle$.

**Theorem 2.12.** For every bounded operator $A \in \mathfrak{L}(\mathfrak{H})$ there is a unique bounded operator $A^*$ defined via

$$\langle f, A^* g \rangle = \langle Af, g \rangle. \quad (2.27)$$

A bounded operator satisfying $A^* = A$ is called self-adjoint. Note that $q_{A^*}(f) = \langle Af, f \rangle = q_A(f)^*$ and hence a bounded operator is self-adjoint if and only if its quadratic form is real-valued.

**Example.** If $\mathfrak{H} = \mathbb{C}^n$ and $A = (a_{jk})_{1 \leq j, k \leq n}$, then $A^* = (a_{kj}^*)_{1 \leq j, k \leq n}$.

**Example.** Let $a \in \ell^\infty(\mathbb{N})$ and consider the multiplication operator

$$(Ab)_j = a_j b_j.$$ 

Then

$$\langle Ab, c \rangle = \sum_{j=1}^{\infty} (a_j b_j)^* c_j = \sum_{j=1}^{\infty} b_j^* (a_j^* c_j) = \langle b, A^* c \rangle$$

with $(A^* c)_j = a_j^* c_j$, that is, $A^*$ is the multiplication operator with $a^*$. 

**Example.** Let $\mathfrak{H} = \ell^2(\mathbb{N})$ and consider the shift operators defined via

$$(S^\pm a)_j = a_{j\pm 1}$$

with the convention that $a_0 = 0$. That is, $S^-$ shifts a sequence to the right and fills up the left most place by zero and $S^+$ shifts a sequence to the left dropping discarding the left most place:

$S^- (a_1, a_2, a_3, \cdots) = (0, a_1, a_2, \cdots), \quad S^+ (a_1, a_2, a_3, \cdots) = (a_2, a_3, a_4, \cdots).$

Then

$$\langle S^- a, b \rangle = \sum_{j=2}^{\infty} a_{j-1}^* b_j = \sum_{j=1}^{\infty} a_j^* b_{j+1} = \langle a, S^+ b \rangle,$$

which shows that $(S^-)^* = S^+$. Using symmetry of the scalar product we also get $\langle b, S^- a \rangle = \langle S^+ b, a \rangle$, that is, $(S^+)^* = S^-$. 

Note that $S^+$ is a left inverse of $S^-$, $S^+ S^- = \mathbb{1}$ but not a right inverse as $S^- S^+ \neq \mathbb{1}$. This is different from the finite dimensional case, where a left inverse is also a right inverse and vice versa.
Example. Suppose $U \in \mathcal{L}(\mathcal{H})$ is unitary. Then $U^* = U^{-1}$. This follows from Lemma 2.11 since $\langle f, g \rangle = \langle Uf, Ug \rangle = \langle f, U^*Ug \rangle$ implies $U^*U = I$. Since $U$ is bijective we can multiply this last equation from the right with $U^{-1}$ to obtain the claim.

A few simple properties of taking adjoints are listed below.

Lemma 2.13. Let $A, B \in \mathcal{L}(\mathcal{H})$ and $\alpha \in \mathbb{C}$. Then

(i) $(A + B)^* = A^* + B^*$, \hspace{1cm} (ii) $(\alpha A)^* = \alpha^* A^*$,

(iii) $(AB)^* = B^*A^*$,

(iv) $\|A^*\| = \|A\|$ and $\|A\|^2 = \|A^*A\| = \|AA^*\|$.

Proof. (i) is obvious. (ii) follows from $\langle f, A^*g \rangle = \langle A^*f, g \rangle = \langle f, Ag \rangle$. (iii) follows from $\langle f, (AB)g \rangle = \langle A^*f, Bg \rangle = \langle B^*A^*f, g \rangle$. (iv) follows using (2.26) from

$$
\|A^*\| = \sup_{\|f\| = \|g\| = 1} |\langle f, A^*g \rangle| = \sup_{\|f\| = \|g\| = 1} |\langle Af, g \rangle| = \sup_{\|f\| = \|g\| = 1} |\langle g, Af \rangle| = \|A\|
$$

and

$$
\|A^*A\| = \sup_{\|f\| = \|g\| = 1} |\langle f, A^*Ag \rangle| = \sup_{\|f\| = \|g\| = 1} |\langle Af, Ag \rangle| = \sup_{\|f\| = 1} \|Af\|^2 = \|A\|^2,
$$

where we have used that $|\langle Af, Ag \rangle|$ attains its maximum when $Af$ and $Ag$ are parallel (compare Theorem 1.28).

Note that $\|A\| = \|A^*\|$ implies that taking adjoints is a continuous operation. For later use also note that (Problem 2.9)

$$
\ker(A^*) = \mathrm{ran}(A)^\perp.
$$

(2.28)

A sesquilinear form is called nonnegative if $s(f, f) \geq 0$ and we will call $A$ nonnegative, $A \geq 0$, if its associated sesquilinear form is. We will write $A \geq B$ if $A - B \geq 0$. Observe that nonnegative operators are self-adjoint (as their quadratic forms are real-valued).

Example. For any operator $A$ the operators $A^*A$ and $AA^*$ are both nonnegative. In fact $\langle f, A^*Af \rangle = \langle Af, Af \rangle = \|Af\|^2 \geq 0$ and similarly $\langle f, AA^*f \rangle = \|A^*f\|^2 \geq 0$.

Lemma 2.14. Suppose $A \geq \varepsilon I$ for some $\varepsilon > 0$. Then $A$ is a bijection with bounded inverse, $\|A^{-1}\| \leq \frac{1}{\varepsilon}$. 

\[ \]
Proof. By definition \( \varepsilon \| f \|^2 \leq \langle f, Af \rangle \leq \| f \| \| Af \| \) and thus \( \varepsilon \| f \| \leq \| Af \| \). In particular, \( Af = 0 \) implies \( f = 0 \) and thus for every \( g \in \text{Ran}(A) \) there is a unique \( f = A^{-1}g \). Moreover, by \( \| A^{-1}g \| = \| f \| \leq \varepsilon^{-1} \| Af \| = \varepsilon^{-1} \| g \| \) the operator \( A^{-1} \) is bounded. So if \( g_n \in \text{Ran}(A) \) converges to some \( g \in \mathcal{F} \), then \( f_n = A^{-1}g_n \) converges to some \( f \). Taking limits in \( g_n = Af_n \) shows that \( g = Af \) is in the range of \( A \), that is, the range of \( A \) is closed. To show that \( \text{Ran}(A) = \mathcal{F} \) we pick \( h \in \text{Ran}(A) \perp \). Then \( 0 = \langle h, Ah \rangle \geq \varepsilon \| h \|^2 \) shows \( h = 0 \) and thus \( \text{Ran}(A) \perp = \{0\} \). \( \square \)

Combining the last two results we obtain the famous Lax–Milgram theorem which plays an important role in theory of elliptic partial differential equations.

**Theorem 2.15** (Lax–Milgram). Let \( s \) be a sesquilinear form which is
- **bounded**, \( |s(f, g)| \leq C \| f \| \| g \| \), and
- **coercive**, \( s(f, f) \geq \varepsilon \| f \|^2 \) for some \( \varepsilon > 0 \).

Then for every \( g \in \mathcal{F} \) there is a unique \( f \in \mathcal{F} \) such that

\[
\langle h, f \rangle = \langle h, g \rangle, \quad \forall h \in \mathcal{F}.
\]

(2.29)

**Proof.** Let \( A \) be the operator associated with \( s \). Then \( A \geq \varepsilon \) and \( f = A^{-1}g \). \( \square \)

**Example.** Consider \( \mathcal{F} = \ell^2(\mathbb{N}) \) and introduce the operator

\[
(Aa)_j = -a_{j+1} + 2a_j - a_{j-1}
\]

which is a discrete version of a second derivative (discrete one-dimensional Laplace operator). Here we use the convention \( a_0 = 0 \), that is, \( (Aa)_1 = -a_2 + 2a_1 \). In terms of the shift operators \( S^\pm \) we can write

\[
A = -S^+ + 2 - S^- = (S^+ - 1)(S^- - 1)
\]

and using \( (S^\pm)^* = S^\mp \) we obtain

\[
s_A(a, b) = \langle (S^- - 1)a, (S^- - 1)b \rangle = \sum_{j=1}^{\infty} (a_{j-1} - a_j)^* (b_{j-1} - b_j).
\]

In particular, this shows \( A \geq 0 \). Moreover, we have \( |s_A(a, b)| \leq 4 \| a \|_2 \| b \|_2 \) or equivalently \( \| A \| \leq 4 \).

Next, let

\[
(Qa)_j = q_j a_j
\]

for some sequence \( q \in \ell^\infty(\mathbb{N}) \). Then

\[
s_Q(a, b) = \sum_{j=1}^{\infty} q_j a_j^* b_j
\]
and \(|s_Q(a, b)| \leq \|q\|_{\infty} \|a\|_2 \|b\|_2\) or equivalently \(\|Q\| \leq \|q\|_{\infty}\). If in addition \(q_j \geq \varepsilon > 0\), then \(s_{A+Q}(a, b) = s_A(a, b) + s_Q(a, b)\) satisfies the assumptions of the Lax–Milgram theorem and

\[(A + Q)a = b\]

has a unique solution \(a = (A + Q)^{-1}b\) for every given \(b \in \ell^2(\mathbb{Z})\). Moreover, since \((A + Q)^{-1}\) is bounded, this solution depends continuously on \(b\).

\[\Diamond\]

**Problem 2.6.** Let \(\mathcal{H}\) be a Hilbert space and let \(u, v \in \mathcal{H}\). Show that the operator

\[Af = \langle u, f \rangle v\]

is bounded and compute its norm. Compute the adjoint of \(A\).

**Problem 2.7.** Prove \((2.26)\). (Hint: Use \(\|f\| = \sup_{\|f\| = 1} |\langle f, g \rangle| \quad \text{— compare Theorem 1.28.}\))

**Problem 2.8.** Suppose \(A\) has a bounded inverse \(A^{-1}\). Show \((A^{-1})^* = (A^*)^{-1}\).

**Problem 2.9.** Show \((2.28)\).

### 2.4. Orthogonal sums and tensor products

Given two Hilbert spaces \(\mathcal{H}_1\) and \(\mathcal{H}_2\), we define their **orthogonal sum** \(\mathcal{H}_1 \oplus \mathcal{H}_2\) to be the set of all pairs \((f_1, f_2) \in \mathcal{H}_1 \times \mathcal{H}_2\) together with the scalar product

\[\langle (g_1, g_2), (f_1, f_2) \rangle = \langle g_1, f_1 \rangle_{\mathcal{H}_1} + \langle g_2, f_2 \rangle_{\mathcal{H}_2}.\]  \hspace{1cm} (2.30)

It is left as an exercise to verify that \(\mathcal{H}_1 \oplus \mathcal{H}_2\) is again a Hilbert space. Moreover, \(\mathcal{H}_1\) can be identified with \(\{(f_1, 0) | f_1 \in \mathcal{H}_1\}\), and we can regard \(\mathcal{H}_1\) as a subspace of \(\mathcal{H}_1 \oplus \mathcal{H}_2\), and similarly for \(\mathcal{H}_2\). With this convention we have \(\mathcal{H}_1^\perp = \mathcal{H}_2\). It is also customary to write \(f_1 \oplus f_2\) instead of \((f_1, f_2)\). In the same way we can define the orthogonal sum \(\bigoplus_{j=1}^n \mathcal{H}_j\) of any finite number of Hilbert spaces.

**Example.** For example we have \(\bigoplus_{j=1}^n \mathbb{C} = \mathbb{C}^n\) and hence we will write \(\bigoplus_{j=1}^n \mathcal{H}_j = \mathcal{H}_n\).

More generally, let \(\mathcal{H}_j, j \in \mathbb{N}\), be a countable collection of Hilbert spaces and define

\[\bigoplus_{j=1}^\infty \mathcal{H}_j = \left\{ \bigoplus_{j=1}^\infty f_j | f_j \in \mathcal{H}_j, \sum_{j=1}^\infty \|f_j\|^2_{\mathcal{H}_j} < \infty \right\},\]  \hspace{1cm} (2.31)

which becomes a Hilbert space with the scalar product

\[\langle \bigoplus_{j=1}^\infty g_j, \bigoplus_{j=1}^\infty f_j \rangle = \sum_{j=1}^\infty \langle g_j, f_j \rangle_{\mathcal{H}_j},\]  \hspace{1cm} (2.32)
Example. $\bigoplus_{j=1}^{\infty} \mathbb{C} = \ell^2(\mathbb{N})$. 

Similarly, if $\mathcal{H}$ and $\tilde{\mathcal{H}}$ are two Hilbert spaces, we define their tensor product as follows: The elements should be products $f \otimes \tilde{f}$ of elements $f \in \mathcal{H}$ and $\tilde{f} \in \tilde{\mathcal{H}}$. Hence we start with the set of all finite linear combinations of elements of $\mathcal{H} \times \tilde{\mathcal{H}}$

$$\mathcal{F}(\mathcal{H}, \tilde{\mathcal{H}}) = \{ \sum_{j=1}^{n} \alpha_j (f_j, \tilde{f}_j) | (f_j, \tilde{f}_j) \in \mathcal{H} \times \tilde{\mathcal{H}}, \alpha_j \in \mathbb{C} \}. \quad (2.33)$$

Since we want $(f_1 + f_2) \otimes \tilde{f} = f_1 \otimes \tilde{f} + f_2 \otimes \tilde{f}$, $f \otimes (\tilde{f}_1 + \tilde{f}_2) = f \otimes \tilde{f}_1 + f \otimes \tilde{f}_2$, and $(\alpha f) \otimes \tilde{f} = f \otimes (\alpha \tilde{f}) = \alpha (f \otimes \tilde{f})$ we consider $\mathcal{F}(\mathcal{H}, \tilde{\mathcal{H}}) / \mathcal{N}(\mathcal{H}, \tilde{\mathcal{H}})$, where

$$\mathcal{N}(\mathcal{H}, \tilde{\mathcal{H}}) = \text{span}\{ \sum_{j,k=1}^{n} \alpha_j \beta_k (f_j, \tilde{f}_k) - (\sum_{j=1}^{n} \alpha_j f_j, \sum_{k=1}^{n} \beta_k \tilde{f}_k) \} \quad (2.34)$$

and write $f \otimes \tilde{f}$ for the equivalence class of $(f, \tilde{f})$. By construction, every element in this quotient space is a linear combination of elements of the type $f \otimes \tilde{f}$.

Next, we want to define a scalar product such that

$$\langle f \otimes \tilde{f}, g \otimes \tilde{g} \rangle = \langle f, g \rangle_{\mathcal{H}} \langle \tilde{f}, \tilde{g} \rangle_{\tilde{\mathcal{H}}} \quad (2.35)$$

holds. To this end we set

$$s(\sum_{j=1}^{n} \alpha_j (f_j, \tilde{f}_j), \sum_{k=1}^{n} \beta_k (g_k, \tilde{g}_k)) = \sum_{j,k=1}^{n} \alpha_j \beta_k \langle f_j, g_k \rangle_{\mathcal{H}} \langle \tilde{f}_j, \tilde{g}_k \rangle_{\tilde{\mathcal{H}}}, \quad (2.36)$$

which is a symmetric sesquilinear form on $\mathcal{F}(\mathcal{H}, \tilde{\mathcal{H}})$. Moreover, one verifies that $s(f, g) = 0$ for arbitrary $f \in \mathcal{F}(\mathcal{H}, \tilde{\mathcal{H}})$ and $g \in \mathcal{N}(\mathcal{H}, \tilde{\mathcal{H}})$ and thus

$$\langle \sum_{j=1}^{n} \alpha_j f_j \otimes \tilde{f}_j, \sum_{k=1}^{n} \beta_k g_k \otimes \tilde{g}_k \rangle = \sum_{j,k=1}^{n} \alpha_j \beta_k \langle f_j, g_k \rangle_{\mathcal{H}} \langle \tilde{f}_j, \tilde{g}_k \rangle_{\tilde{\mathcal{H}}} \quad (2.37)$$

is a symmetric sesquilinear form on $\mathcal{F}(\mathcal{H}, \tilde{\mathcal{H}}) / \mathcal{N}(\mathcal{H}, \tilde{\mathcal{H}})$. To show that this is in fact a scalar product, we need to ensure positivity. Let $f = \sum_i \alpha_i f_i \otimes \tilde{f}_i \neq 0$ and pick orthonormal bases $u_j, \tilde{u}_k$ for $\text{span}\{ f_i \}, \text{span}\{ \tilde{f}_i \}$, respectively. Then

$$f = \sum_{j,k} \alpha_{jk} u_j \otimes \tilde{u}_k, \quad \alpha_{jk} = \sum_i \alpha_i \langle u_j, f_i \rangle_{\mathcal{H}} \langle \tilde{u}_k, \tilde{f}_i \rangle_{\tilde{\mathcal{H}}} \quad (2.38)$$

and we compute

$$\langle f, f \rangle = \sum_{j,k} |\alpha_{jk}|^2 > 0. \quad (2.39)$$

The completion of $\mathcal{F}(\mathcal{H}, \tilde{\mathcal{H}}) / \mathcal{N}(\mathcal{H}, \tilde{\mathcal{H}})$ with respect to the induced norm is called the **tensor product** $\mathcal{H} \otimes \tilde{\mathcal{H}}$ of $\mathcal{H}$ and $\tilde{\mathcal{H}}$. 
Lemma 2.16. If $u_j, \tilde{u}_k$ are orthonormal bases for $\mathcal{H}, \tilde{\mathcal{H}}$, respectively, then $u_j \otimes \tilde{u}_k$ is an orthonormal basis for $\mathcal{H} \otimes \tilde{\mathcal{H}}$.

**Proof.** That $u_j \otimes \tilde{u}_k$ is an orthonormal set is immediate from (2.35). Moreover, since span$\{u_j\}$, span$\{\tilde{u}_k\}$ are dense in $\mathcal{H}, \tilde{\mathcal{H}}$, respectively, it is easy to see that $u_j \otimes \tilde{u}_k$ is dense in $\mathcal{F}(\mathcal{H}, \tilde{\mathcal{H}})/\mathcal{N}(\mathcal{H}, \tilde{\mathcal{H}})$. But the latter is dense in $\mathcal{H} \otimes \tilde{\mathcal{H}}$. □

Note that this in particular implies dim$(\mathcal{H} \otimes \tilde{\mathcal{H}}) = \text{dim}(\mathcal{H}) \cdot \text{dim}(\tilde{\mathcal{H}})$.

**Example.** We have $\mathcal{H} \otimes \mathbb{C}^n = \mathcal{H}^n$. ○

**Example.** We have $\ell^2(\mathbb{N}) \otimes \ell^2(\mathbb{N}) = \ell^2(\mathbb{N} \times \mathbb{N})$ by virtue of the identification $(a_{jk}) \mapsto \sum_{jk} a_{jk} \delta^j \otimes \delta^k$ where $\delta^j$ is the standard basis for $\ell^2(\mathbb{N})$. In fact, this follows from the previous lemma as in the proof of Theorem 2.6. ○

It is straightforward to extend the tensor product to any finite number of Hilbert spaces. We even note

$$
(\bigoplus_{j=1}^{\infty} \mathcal{H}_j) \otimes \mathcal{H} = \bigoplus_{j=1}^{\infty} (\mathcal{H}_j \otimes \mathcal{H}),
$$

(2.40)

where equality has to be understood in the sense that both spaces are unitarily equivalent by virtue of the identification

$$(\sum_{j=1}^{\infty} f_j) \otimes f = \sum_{j=1}^{\infty} f_j \otimes f.$$  

(2.41)

**Problem 2.10.** Show that $f \otimes \tilde{f} = 0$ if and only if $f = 0$ or $\tilde{f} = 0$.

**Problem 2.11.** We have $f \otimes \tilde{f} = g \otimes \tilde{g} \neq 0$ if and only if there is some $\alpha \in \mathbb{C} \setminus \{0\}$ such that $f = \alpha g$ and $\tilde{f} = \alpha^{-1} \tilde{g}$.

**Problem 2.12.** Show (2.40).

2.5. Applications to Fourier series

Given an integrable function $f$ we can define its **Fourier series**

$$S(f)(x) = \frac{a_0}{2} + \sum_{k \in \mathbb{N}} a_k \cos(kx) + b_k \sin(kx)$$

(2.42)

where the corresponding Fourier coefficients are given by

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(kx)f(x)dx, \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(kx)f(x)dx.$$  

(2.43)

At this point (2.42) is just a formal expression and it was (and to some extend still is) a fundamental question in mathematics to understand in what sense the above series converges. For example, does it converge at a
fixed point (e.g. at every point of continuity) or does it converge uniformly. We will give some first answers in the present section and then come back later to this when we have further tools at our disposal.

For our purpose the complex form

\[ S(f)(x) = \sum_{k \in \mathbb{Z}} \hat{f}_k e^{ikx}, \quad \hat{f}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ikx} f(x)dx \]  

(2.44)

will be more convenient, where the connection is given via \( \hat{f}_{\pm k} = \frac{a_k \mp b_k}{2} \). In this case the \( n \)’th partial sum can be written as

\[ S_n(f) = \sum_{k=-n}^{n} \hat{f}(k)e^{ikx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} D_n(x-y)f(x)dy, \]

where

\[ D_n(x) = \sum_{k=-n}^{n} e^{ikx} = \frac{\sin((n+1/2)x)}{\sin(x/2)} \]

is known as the Dirichlet kernel (to obtain the second form observe that the left-hand side is a geometric series). Note that \( D_n(-x) = -D(x) \) and that \( |D_n(x)| \) has a global maximum \( D_n(0) = 2n + 1 \) at \( x = 0 \).

Since

\[ \int_{-\pi}^{\pi} e^{-ikx}e^{ilx}dx = 2\pi\delta_{k,l} \]  

(2.45)
the functions \( e_k(x) = (2\pi)^{-1/2}e^{ikx} \) are orthonormal in \( L^2(-\pi, \pi) \) and hence the Fourier series is just the expansion with respect to this orthogonal set. Hence we obtain

**Theorem 2.17.** For every square integrable function \( f \in L^2(-\pi, \pi) \), the Fourier coefficients \( \hat{f}_k \) are square summable and the Fourier series converges to \( f \) in the sense of \( L^2 \).

**Proof.** To show this theorem it suffices to show that the functions \( e_k \) form a basis. This will be deduced as a special of Theorem 3.11 below (see the examples after this theorem). \( \square \)

This gives a satisfactory answer in the Hilbert space \( L^2(-\pi, \pi) \) but does not answer the question about pointwise or uniform convergence. The latter will be the case if the Fourier coefficients are summable. First of all we note that for integrable functions the Fourier coefficients will at least tend to zero.

**Lemma 2.18** (Riemann–Lebesgue lemma). Suppose \( f \) is integrable, then the Fourier coefficients converge to zero.

**Proof.** By our previous theorem this holds for continuous functions. But the map \( f \to \hat{f} \) is bounded from \( C[-\pi, \pi] \subset L^1(-\pi, \pi) \) to \( c_0(\mathbb{Z}) \) (the sequences vanishing as \( |k| \to \infty \)) since \( \|\hat{f}_k\| \leq (2\pi)^{-1}\|f\|_1 \) and there is a unique extension to all of \( L^1(-\pi, \pi) \). \( \square \)

It turns out that this result is best possible in general and we cannot say more without additional assumptions on \( f \). For example, if \( f \) is periodic and differentiable, then integration by parts shows

\[
\hat{f}_k = \frac{1}{2\pi i k} \int_{-\pi}^{\pi} e^{-ikx} f'(x) dx. \tag{2.46}
\]

Then, since both \( k^{-1} \) and the Fourier coefficients of \( f' \) are square summable, we conclude that \( \hat{f}_k \) are summable and hence the Fourier series converges uniformly. So we have a simple sufficient criterion for summability of the Fourier coefficients but can it be improved? Of course continuity of \( f \) is a necessary condition but this alone will not even be enough for pointwise convergence as we will see in the example on page 87. Moreover, continuity will not tell us more about the decay of the Fourier coefficients than what we already know in the integrable case from the Riemann–Lebesgue lemma (see the example on page 87).

A few improvements are easy: First of all, piecewise continuously differentiable would be sufficient for this argument. Or, slightly more general, an absolutely continuous function whose derivative is square integrable would
also do (cf. Lemma 9.26). However, even for an absolutely continuous function the Fourier coefficients might not be summable: For an absolutely continuous function $f$ we have a derivative which is integrable (Theorem 9.25) and hence the above formula combined with the Riemann–Lebesgue lemma implies $\hat{f}_k = o(\frac{1}{k})$. But on the other hand we can choose a summable sequence $c_k$ which does not obey this asymptotic requirement, say $c_k = \frac{1}{k}$ for $k = l^2$ and $c_k = 0$ else. Then

$$f(x) = \sum_{k \in \mathbb{Z}} c_k e^{ikx} = \sum_{l \in \mathbb{N}} \frac{1}{l^2} e^{il^2x}$$

is a function with summable Fourier coefficients $\hat{f}_k = c_k$ (by uniform convergence we can interchange summation and integration) but which is not absolutely continuous. There are further criteria for summability of the Fourier coefficients but no simple necessary and sufficient one.

Finally, we look at pointwise convergence.

**Theorem 2.19.** Suppose

$$\frac{f(x) - f(x_0)}{x - x_0}$$

is integrable (e.g. $f$ is Hölder continuous), then

$$\lim_{m,n \to \infty} \sum_{k=-m}^{n} \hat{f}(k)e^{ikx_0} = f(x_0).$$

**Proof.** Without loss of generality we can assume $x_0 = 0$ (by shifting $x \to x - x_0$ modulo $2\pi$ implying $\hat{f}_k \to e^{-ikx_0}\hat{f}_k$) and $f(x_0) = 0$ (by linearity since the claim is trivial for constant functions). Then by assumption

$$g(x) = \frac{f(x)}{e^{ix} - 1}$$

is integrable and $f(x) = (e^{ix} - 1)g(x)$ implies

$$\hat{f}_k = \hat{g}_{k-1} - \hat{g}_k$$

and hence

$$\sum_{k=m}^{n} \hat{f}_k = \hat{g}_{m-1} - \hat{g}_n.$$

Now the claim follows from the Riemann–Lebesgue lemma.

If one looks at symmetric partial sums $S_n(f)$ we can do even better.

**Corollary 2.20** (Dirichlet–Dini criterion). Suppose there is some $\alpha$ such that

$$\frac{f(x_0 + x) + f(x_0 - x) - 2\alpha}{x}$$

is integrable. Then $S_n(f)(0) \to \alpha$. 

Proof. Without loss of generality we can assume $x_0 = 0$. Now observe (since $D_n(-x) = D_n(x)$)

$$S_n(f)(0) = \alpha + S_n(g)(0), \quad g(x) = \frac{f(x) + f(-x) - 2\alpha}{2}$$

and apply the previous result. \qed
Chapter 3

Compact operators

3.1. Compact operators

A linear operator \( A : X \to Y \) defined between normed spaces \( X, Y \) is called \textbf{compact} if every sequence \( Af_n \) has a convergent subsequence whenever \( f_n \) is bounded. Equivalently (cf. Corollary 1.15), \( A \) is compact if it maps bounded sets to relatively compact ones. The set of all compact operators is denoted by \( \mathcal{C}(X,Y) \). If \( X = Y \) we will just write \( \mathcal{C}(X) = \mathcal{C}(X,X) \) as usual.

Example. Every linear map between finite dimensional spaces is compact by the Bolzano–Weierstraß theorem. Slightly more general, an operator is compact if its range is finite dimensional.

The following elementary properties of compact operators are left as an exercise (Problem 3.1):

\textbf{Theorem 3.1.} Let \( X, Y, \) and \( Z \) be normed spaces. Every compact linear operator is bounded, \( \mathcal{C}(X,Y) \subseteq \mathcal{L}(X,Y) \). Linear combinations of compact operators are compact, that is, \( \mathcal{C}(X,Y) \) is a subspace of \( \mathcal{L}(X,Y) \). Moreover, the product of a bounded and a compact operator is again compact, that is, \( A \in \mathcal{L}(X,Y), B \in \mathcal{C}(Y,Z) \) or \( A \in \mathcal{C}(X,Y), B \in \mathcal{L}(Y,Z) \) implies \( BA \in \mathcal{C}(X,Z) \).

In particular, the set of compact operators \( \mathcal{C}(X) \) is an ideal of the set of bounded operators. Moreover, if \( X \) is a Banach space this ideal is even closed:

\textbf{Theorem 3.2.} Suppose \( X \) is a normed and \( Y \) a Banach space. Let \( A_n \in \mathcal{C}(X,Y) \) be a convergent sequence of compact operators. Then the limit \( A \) is again compact.
Proof. Let \( f_j^{(0)} \) be a bounded sequence. Choose a subsequence \( f_j^{(1)} \) such that \( A_1 f_j^{(1)} \) converges. From \( f_j^{(1)} \) choose another subsequence \( f_j^{(2)} \) such that \( A_2 f_j^{(2)} \) converges and so on. Since \( f_j^{(n)} \) might disappear as \( n \to \infty \), we consider the diagonal sequence \( f_j = f_j^{(j)} \). By construction, \( f_j \) is a subsequence of \( f_j^{(n)} \) for \( j \geq n \) and hence \( A_n f_j \) is Cauchy for every fixed \( n \). Now

\[
\| A f_j - A f_k \| = \| (A - A_n)(f_j - f_k) + A_n(f_j - f_k) \| \\
\leq \| A - A_n \| \| f_j - f_k \| + \| A_n f_j - A_n f_k \|
\]

shows that \( Af_j \) is Cauchy since the first term can be made arbitrary small by choosing \( n \) large and the second by the Cauchy property of \( A_n f_j \).

Example. Let \( X = \ell^p(\mathbb{N}) \) and consider the operator

\[
(Qa)_j = q_j b_j
\]

for some sequence \( q = (q_j)_{j=1}^{\infty} \in c_0(\mathbb{N}) \) converging to zero. Let \( Q_n \) be associated with \( q_n^j = q_j \) for \( j \leq n \) and \( q_n^j = 0 \) for \( j > n \). Then the range of \( Q^n \) is finite dimensional and hence \( Q_n \) is compact. Moreover, by \( \|Q_n - Q\| = \sup_{j>n} |a_j| \) we see \( Q_n \to Q \) and thus \( Q \) is also compact by the previous theorem.

If \( A : X \to Y \) is a bounded operator there is a unique extension \( \overline{A} : \overline{X} \to \overline{Y} \) to the completion by Theorem 1.35. Moreover, if \( A \in \mathcal{C}(X,Y) \), then \( A \in \mathcal{C}(X,Y) \) is immediate. That we also have \( \overline{A} \in \mathcal{C}(\overline{X},\overline{Y}) \) will follow from the next theorem. In particular, it suffices to verify compactness on a dense set.

**Lemma 3.3.** Let \( X, Y \) be normed spaces and \( A \in \mathcal{C}(X,Y) \). Let \( \overline{X}, \overline{Y} \) be the completion of \( X, Y \), respectively. Then \( \overline{A} \in \mathcal{C}(\overline{X},\overline{Y}) \), where \( \overline{A} \) is the unique extension of \( A \).

**Proof.** Let \( f_n \in X \) be a given bounded sequence. We need to show that \( \overline{A} f_n \) has a convergent subsequence. Pick \( f_n^{(j)} \in X \) such that \( \|f_n^{(j)} - f_n\| \leq \frac{1}{j} \) and by compactness of \( A \) we can assume that \( Af_n^{(j)} \to g \). But then \( \|\overline{A} f_n - g\| \leq \|A\| \|f_n - f_n^{(j)}\| + \|Af_n^{(j)} - g\| \) shows that \( \overline{A} f_n \to g \).

One of the most important examples of compact operators are integral operators. To prove this we will need a good criterion when a sequence of continuous functions has a convergent subsequence. For the formulation of this result recall that a family of functions \( F \subset C(K) \), where \( K \) is some metric space, is called (uniformly) **equicontinuous** if for every \( \varepsilon > 0 \) there is a \( \delta > 0 \) such that

\[
|f_n(x) - f_n(y)| \leq \varepsilon \quad \text{if} \quad |x - y| < \delta, \quad \forall f \in F.
\]

(3.1)
3.1. Compact operators

Theorem 3.4 (Arzelà–Ascoli). Let $K$ be a compact metric space and let $F \subseteq C(K)$ be a family of continuous functions. Then every sequence from $F$ has a uniformly convergent subsequence if and only if $F$ is equicontinuous and the set $\{f(x) \mid f \in F\}$ is bounded for one $x$. In this case $F$ is even bounded: $\|f\|_{\infty} \leq C$ for all $f \in F$.

Proof. First of all note that if $F$ is equicontinuous and the set $\{f(x) \mid f \in F\}$ is bounded for one $x$ then $F$ is bounded. To see this fix $\varepsilon$ and cover $K$ by $N$ balls of radius less than $\delta$. Then $|f(y)| \leq N\varepsilon + |f(x)|$ for every $y \in [a, b]$ and every $f \in F$.

Now let $f_n$ be a sequence from $F$. Let $\{x_j\}_{j=1}^{\infty}$ be a dense subset of $K$ (cf. Corollary 1.16). Since $f_n(x_1)$ is bounded, we can choose a subsequence $f_n^{(1)}(x)$ such that $f_n^{(1)}(x_1)$ converges (Bolzano–Weierstraß). Similarly we can extract a subsequence $f_n^{(2)}(x)$ from $f_n^{(1)}(x)$ which converges at $x_2$ (and hence also at $x_1$ since it is a subsequence of $f_n^{(1)}(x)$). By induction we get a sequence $f_n^{(j)}(x)$ converging at $x_1, \ldots, x_j$. The diagonal sequence $\tilde{f}_n(x) = f_n^{(n)}(x)$ will hence converge for all $x = x_j$ (why?). We will show that it converges uniformly for all $x$:

Fix $\varepsilon > 0$ and chose $\delta$ such that $|f_n(x) - f_n(y)| \leq \frac{\varepsilon}{3}$ for $|x - y| < \delta$. The balls $B_\delta(x_j)$ cover $K$ and by compactness even finitely many, say $1 \leq j \leq p$, suffice. Furthermore, choose $N_{\varepsilon}$ such that $|\tilde{f}_m(x_j) - \tilde{f}_n(x_j)| \leq \frac{\varepsilon}{3}$ for $n, m \geq N_{\varepsilon}$ and $1 \leq j \leq p$.

Now pick $x$ and note that $x \in B_\delta(x_j)$ for some $j$. Thus

$$|\tilde{f}_m(x) - \tilde{f}_n(x)| \leq |\tilde{f}_m(x) - \tilde{f}_m(x_j)| + |\tilde{f}_m(x_j) - \tilde{f}_n(x_j)| + |\tilde{f}_n(x_j) - \tilde{f}_n(x)| \leq \varepsilon$$

for $n, m \geq N_{\varepsilon}$, which shows that $\tilde{f}_n$ is Cauchy with respect to the maximum norm. By completeness of $C(K)$ it has a limit.

To see the converse first note that if $\{f(x) \mid f \in F\}$ were unbounded for some $x$, then there would be a sequence of functions $\tilde{f}_n$ such that $|\tilde{f}_n(x)| \to \infty$. A contradiction. Similarly, if $F$ were not equicontinuous, there must be an $\varepsilon_0 > 0$ such that for every $n \in \mathbb{N}$ there is a function $f_n \in F$ and points $x_n, y_n$ with $|x_n - y_n| < \frac{1}{n}$ and $|f_n(x_n) - f_n(y_n)| \geq \varepsilon_0$. By passing to a subsequence we can assume $x_n \to x$ and hence also $y_n \to x$. Moreover, passing to yet another subsequence we can assume that $f_n \to f$ uniformly. But then $0 = |f(x) - f(x)| = \lim_{n \to \infty} |f_n(x_n) - f_n(y_n)| \geq \varepsilon_0$, a contradiction. □

Now we can show:
Lemma 3.5. Let $X = C([a, b])$ or $X = L^2_{\text{cont}}(a, b)$. The integral operator $K : X \to X$ defined by
\[
(Kf)(x) = \int_a^b K(x, y)f(y)\,dy, \quad (3.2)
\]
where $K(x, y) \in C([a, b] \times [a, b])$, is compact.

Proof. First of all note that $K(\ldots)$ is continuous on $[a, b] \times [a, b]$ and hence uniformly continuous. In particular, for every $\varepsilon > 0$ we can find a $\delta > 0$ such that $|K(y, t) - K(x, t)| \leq \varepsilon$ whenever $|y - x| \leq \delta$. Moreover, $\|K\|_{\infty} = \sup_{x, y \in [a, b]} |K(x, y)| < \infty$.

We begin with the case $X = L^2_{\text{cont}}(a, b)$. Let $g(x) = Kf(x)$. Then
\[
|g(x)| \leq \int_a^b |K(x, t)| |f(t)|\,dt \leq \|K\|_{\infty} \int_a^b |f(t)|\,dt \leq \|K\|_{\infty} \|1\| \|f\|,
\]
where we have used Cauchy–Schwarz in the last step (note that $\|1\| = \sqrt{b - a}$). Similarly,
\[
|g(x) - g(y)| \leq \int_a^b |K(y, t) - K(x, t)| |f(t)|\,dt
\leq \varepsilon \int_a^b |f(t)|\,dt \leq \varepsilon \|1\| \|f\|,
\]
whenever $|y - x| \leq \delta$. Hence, if $f_n(x)$ is a bounded sequence in $L^2_{\text{cont}}(a, b)$, then $g_n(x) = Kf_n(x)$ is bounded and equicontinuous and hence has a uniformly convergent subsequence by the Arzelà–Ascoli theorem (Theorem 3.4 below). But a uniformly convergent sequence is also convergent in the norm induced by the scalar product. Therefore $K$ is compact.

The case $X = C([a, b])$ follows by the same argument upon observing
\[
\int_a^b |f(t)|\,dt \leq (b - a)\|f\|_{\infty}.
\]

Compact operators are very similar to (finite) matrices as we will see in the next section.

Problem 3.1. Show Theorem 3.1.

Problem 3.2. Show that adjoint of the integral operator $K$ from Lemma 3.5 is the integral operator with kernel $K(y, x)^*$:
\[
(K^*f)(x) = \int_a^b K(y, x)^* f(y)\,dy.
\]
(Hint: Fubini.)

Problem 3.3. Show that the mapping $\frac{d}{dx} : C^1[a, b] \to C[a, b]$ is compact. (Hint: Arzelà–Ascoli.)
3.2. The spectral theorem for compact symmetric operators

Let $H$ be an inner product space. A linear operator $A$ is called symmetric if its domain is dense and if
\[ \langle g, Af \rangle = \langle Ag, f \rangle \quad f, g \in \mathcal{D}(A). \tag{3.3} \]
If $A$ is bounded (with $\mathcal{D}(A) = H$), then $A$ is symmetric precisely if $A = A^*$, that is, if $A$ is self-adjoint. However, for unbounded operators there is a subtle but important difference between symmetry and self-adjointness.

A number $z \in \mathbb{C}$ is called eigenvalue of $A$ if there is a nonzero vector $u \in \mathcal{D}(A)$ such that
\[ Au = zu. \tag{3.4} \]
The vector $u$ is called a corresponding eigenvector in this case. The set of all eigenvectors corresponding to $z$ is called the eigenspace
\[ \text{Ker}(A - z) \tag{3.5} \]
corresponding to $z$. Here we have used the shorthand notation $A - z$ for $A - z\mathbb{I}$. An eigenvalue is called simple if there is only one linearly independent eigenvector.

**Theorem 3.6.** Let $A$ be symmetric. Then all eigenvalues are real and eigenvectors corresponding to different eigenvalues are orthogonal.

**Proof.** Suppose $\lambda$ is an eigenvalue with corresponding normalized eigenvector $u$. Then $\lambda = \langle u, Au \rangle = \langle Au, u \rangle = \lambda^*$, which shows that $\lambda$ is real. Furthermore, if $Au_j = \lambda_j u_j$, $j = 1, 2$, we have
\[ (\lambda_1 - \lambda_2)\langle u_1, u_2 \rangle = \langle Au_1, u_2 \rangle - \langle u_1, Au_2 \rangle = 0 \]
finishing the proof. \qed

Note that while eigenvectors corresponding to the same eigenvalue $\lambda$ will in general not automatically be orthogonal, we can of course replace each set of eigenvectors corresponding to $\lambda$ by an set of orthonormal eigenvectors having the same linear span (e.g. using Gram–Schmidt orthogonalization).

**Example.** Let $\mathcal{H} = \ell^2(\mathbb{N})$ and consider the Jacobi operator $J = \frac{1}{2}(S^+ + S^-)$ associated with the sequences $a_j = \frac{1}{2}$, $b_j = 0$:
\[ (Je)_j = \frac{1}{2}(c_{j+1} + c_{j-1}) \]
with the convention $c_0 = 0$. Recall that $J^* = J$. If we look for an eigenvalue $Ju = zu$, we need to solve the corresponding recursion $u_{j+1} = 2zu_j - u_{j-1}$ starting from $u_0 = 0$ (our convention) and $u_1 = 1$ (normalization) which gives
\[ u_j(z) = \frac{k^j - k^{-j}}{k - k^{-1}}, \quad k = z - \sqrt{z^2 - 1}. \]
Note that \( k^{-1} = z + \sqrt{z^2 - 1} \) and in the case \( k = z = \pm 1 \) the above expression has to be understood as its limit \( u_j(\pm 1) = (\pm 1)^{j+1}j \). In fact, \( T_j(z) = u_{j-1}(z) \) are polynomials of degree \( j \) known as Chebyshev polynomials.

Now for \( z \in \mathbb{R} \setminus [-1, 1] \) we have \( |k| < 1 \) and \( u_j \) explodes exponentially. For \( z \in (-1, 1) \) we have \( |k| = 1 \) and hence we can write \( k = e^{i\kappa} \) with \( \kappa \in \mathbb{R} \). Thus \( u_j = \frac{\sin(j\kappa)}{\sin(\kappa)} \) is oscillating. So for no value of \( z \in \mathbb{R} \) our potential eigenvector \( u_j \) is square summable and thus \( J \) has no eigenvalues.

The previous example shows that in the infinite dimensional case symmetry is not enough to guarantee existence of even a single eigenvalue. We will see that compactness provides a suitable extra condition to obtain an orthonormal basis of eigenfunctions. The crucial step is to prove existence of one eigenvalue, the rest then follows as in the finite dimensional case.

**Theorem 3.7.** Let \( \mathcal{H} \) be an inner product space. A symmetric compact operator \( A \) has an eigenvalue \( \alpha_1 \) which satisfies \( |\alpha_1| = \|A\| \).

**Proof.** We set \( \alpha = \|A\| \) and assume \( \alpha \neq 0 \) (i.e, \( A \neq 0 \)) without loss of generality. Since
\[
\|A\|^2 = \sup_{f : \|f\| = 1} \|Af\|^2 = \sup_{f : \|f\| = 1} \langle Af, Af \rangle = \sup_{f : \|f\| = 1} \langle f, A^2 f \rangle
\]
there exists a normalized sequence \( u_n \) such that
\[
\lim_{n \to \infty} \langle u_n, A^2 u_n \rangle = \alpha^2.
\]
Since \( A \) is compact, it is no restriction to assume that \( A^2 u_n \) converges, say \( \lim_{n \to \infty} A^2 u_n = \alpha^2 u \). Now
\[
\| (A^2 - \alpha^2) u_n \|^2 = \| A^2 u_n \|^2 - 2\alpha^2 \langle u_n, A^2 u_n \rangle + \alpha^4 
\leq 2\alpha^2 (\alpha^2 - \langle u_n, A^2 u_n \rangle)
\]
(where we have used \( \|A^2 u_n\| \leq \|A\| \|A u_n\| \leq \|A\|^2 \|u_n\| = \alpha^2 \)) implies \( \lim_{n \to \infty} (A^2 u_n - \alpha^2 u_n) = 0 \) and hence \( \lim_{n \to \infty} u_n = u \). In addition, \( u \) is a normalized eigenvector of \( A^2 \) since \( (A^2 - \alpha^2) u = 0 \). Factorizing this last equation according to \( (A - \alpha)u = v \) and \( (A + \alpha)v = 0 \) shows that either \( v \neq 0 \) is an eigenvector corresponding to \( -\alpha \) or \( v = 0 \) and hence \( u \neq 0 \) is an eigenvector corresponding to \( \alpha \).

Note that for a bounded operator \( A \), there cannot be an eigenvalue with absolute value larger than \( \|A\| \), that is, the set of eigenvalues is bounded by \( \|A\| \) (Problem 3.4).

Now consider a symmetric compact operator \( A \) with eigenvalue \( \alpha_1 \) (as above) and corresponding normalized eigenvector \( u_1 \). Setting
\[
\mathcal{H}_1 = \{ u_1 \}^\perp = \{ f \in \mathcal{H} | \langle u_1, f \rangle = 0 \}
\]
we can restrict $A$ to $\mathcal{H}_1$ since $f \in \mathcal{H}_1$ implies
\[
\langle u_1, Af \rangle = \langle Au_1, f \rangle = \alpha_1 \langle u_1, f \rangle = 0
\]
and hence $Af \in \mathcal{H}_1$. Denoting this restriction by $A_1$, it is not hard to see that $A_1$ is again a symmetric compact operator. Hence we can apply Theorem 3.7 iteratively to obtain a sequence of eigenvalues $\alpha_j$ with corresponding normalized eigenvectors $u_j$. Moreover, by construction, $u_j$ is orthogonal to all $u_k$ with $k < j$ and hence the eigenvectors $\{u_j\}$ form an orthonormal set.

By construction we have $|\alpha_j| = \|A_j\| \leq \|A_{j-1}\| = |\alpha_{j-1}|$. This procedure will not stop unless $\mathcal{H}$ is finite dimensional. However, note that $\alpha_j = 0$ for $j \geq n$ might happen if $A_n = 0$.

**Theorem 3.8.** Suppose $\mathcal{H}$ is an infinite dimensional Hilbert space and $A : \mathcal{H} \to \mathcal{H}$ is a compact symmetric operator. Then there exists a sequence of real eigenvalues $\alpha_j$ converging to 0. The corresponding normalized eigenvectors $u_j$ form an orthonormal set and every $f \in \mathcal{H}$ can be written as
\[
f = \sum_{j=1}^{\infty} \langle u_j, f \rangle u_j + h,
\]
where $h$ is in the kernel of $A$, that is, $Ah = 0$.

In particular, if 0 is not an eigenvalue, then the eigenvectors form an orthonormal basis (in addition, $\mathcal{H}$ need not be complete in this case).

**Proof.** Existence of the eigenvalues $\alpha_j$ and the corresponding eigenvectors $u_j$ has already been established. Since the sequence $|\alpha_j|$ is decreasing it has a limit $\varepsilon \geq 0$ and we have $|\alpha_j| \geq \varepsilon$. If this limit is nonzero, then $v_j = \alpha_j^{-1} u_j$ is a bounded sequence ($\|v_j\| \leq \frac{1}{\varepsilon}$) for which $Av_j$ has no convergent subsequence since $\|Av_j - Av_k\|^2 = \|u_j - u_k\|^2 = 2$, a contradiction.

Next, setting
\[
f_n = \sum_{j=1}^{n} \langle u_j, f \rangle u_j,
\]
we have
\[
\|A(f - f_n)\| \leq |\alpha_n| \|f - f_n\| \leq |\alpha_n| \|f\|
\]
since $f - f_n \in \mathcal{H}_n$ and $\|A_n\| = |\alpha_n|$. Letting $n \to \infty$ shows $A(f_\infty - f) = 0$ proving (3.8). \( \Box \)

By applying $A$ to (3.8) we obtain the following canonical form of compact symmetric operators.

**Corollary 3.9.** Every compact symmetric operator $A$ can be written as
\[
Af = \sum_{j=1}^{\infty} \alpha_j \langle u_j, f \rangle u_j,
\]
where \( \alpha_j \) are the nonzero eigenvalues with corresponding eigenvectors \( u_j \) from the previous theorem.

Remark: There are two cases where our procedure might fail to construct an orthonormal basis of eigenvectors. One case is where there is an infinite number of nonzero eigenvalues. In this case \( \alpha_n \) never reaches 0 and all eigenvectors corresponding to 0 are missed. In the other case, 0 is reached, but there might not be a countable basis and hence again some of the eigenvectors corresponding to 0 are missed. In any case by adding vectors from the kernel (which are automatically eigenvectors), one can always extend the eigenvectors \( u_j \) to an orthonormal basis of eigenvectors.

**Corollary 3.10.** Every compact symmetric operator has an associated orthonormal basis of eigenvectors.

**Example.** Let \( a, b \in c_0(\mathbb{N}) \) be real-valued sequences and consider the operator

\[
(Jc)_j = a_j c_{j+1} + b_j c_j + a_{j-1} c_{j-1}.
\]

If \( A, B \) denote the multiplication operators by the sequences \( a, b \), respectively, then we already know that \( A \) and \( B \) are compact. Moreover, using the shift operators \( S^\pm \) we can write

\[
J = AS^+ + B + S^- A.
\]

which shows that \( J \) is self-adjoint since \( A^* = A \), \( B^* = B \), and \( (S^\pm)^* = S^\mp \). Hence we can conclude that \( J \) has a countable number of eigenvalues converging to zero and a corresponding orthonormal basis of eigenvectors.

This is all we need and it remains to apply these results to Sturm–Liouville operators.

**Problem 3.4.** Show that if \( A \) is bounded, then every eigenvalue \( \alpha \) satisfies \(|\alpha| \leq \|A\|\).

**Problem 3.5.** Find the eigenvalues and eigenfunctions of the integral operator

\[
(Kf)(x) = \int_0^1 u(x)v(y)f(y)dy
\]

in \( \mathcal{L}_2^{cont}(0,1) \), where \( u(x) \) and \( v(x) \) are some given continuous functions.

**Problem 3.6.** Find the eigenvalues and eigenfunctions of the integral operator

\[
(Kf)(x) = 2 \int_0^1 (2xy - x - y + 1)f(y)dy
\]

in \( \mathcal{L}_2^{cont}(0,1) \).
3.3. Applications to Sturm–Liouville operators

Now, after all this hard work, we can show that our Sturm–Liouville operator

\[ L = -\frac{d^2}{dx^2} + q(x), \]  

where \( q \) is continuous and real, defined on

\[ \mathcal{D}(L) = \{ f \in C^2[0,1] | f(0) = f(1) = 0 \} \subset L^2_{\text{cont}}(0,1), \]

has an orthonormal basis of eigenfunctions.

The corresponding eigenvalue equation \( Lu = zu \) explicitly reads

\[ -u''(x) + q(x)u(x) = zu(x). \]

It is a second order homogeneous linear ordinary differential equations and hence has two linearly independent solutions. In particular, specifying two initial conditions, e.g. \( u(0) = 0, u'(0) = 1 \) determines the solution uniquely. Hence, if we require \( u(0) = 0 \), the solution is determined up to a multiple and consequently the additional requirement \( u(1) = 0 \) cannot be satisfied by a nontrivial solution in general. However, there might be some \( z \in \mathbb{C} \) for which the solution corresponding to the initial conditions \( u(0) = 0, u'(0) = 1 \) happens to satisfy \( u(1) = 0 \) and these are precisely the eigenvalues we are looking for.

Note that the fact that \( L^2_{\text{cont}}(0,1) \) is not complete causes no problems since we can always replace it by its completion \( \mathcal{H} = L^2(0,1) \). A thorough investigation of this completion will be given later, at this point this is not essential.

We first verify that \( L \) is symmetric:

\[ \langle f, Lg \rangle = \int_0^1 f(x)^*(-g''(x) + q(x)g(x))dx \]
\[ = \int_0^1 f'(x)^*g'(x)dx + \int_0^1 f(x)^*q(x)g(x)dx \]
\[ = \int_0^1 -f''(x)^*g(x)dx + \int_0^1 f(x)^*q(x)g(x)dx \]
\[ = \langle Lf, g \rangle. \]

Here we have used integration by part twice (the boundary terms vanish due to our boundary conditions \( f(0) = f(1) = 0 \) and \( g(0) = g(1) = 0 \)).

Of course we want to apply Theorem 3.8 and for this we would need to show that \( L \) is compact. But this task is bound to fail, since \( L \) is not even bounded (see the example in Section 1.5)!

So here comes the trick: If \( L \) is unbounded its inverse \( L^{-1} \) might still be bounded. Moreover, \( L^{-1} \) might even be compact and this is the case
here! Since \( L \) might not be injective (0 might be an eigenvalue), we consider \( R_L(z) = (L - z)^{-1} \), \( z \in \mathbb{C} \), which is also known as the \textbf{resolvent} of \( L \).

In order to compute the resolvent, we need to solve the inhomogeneous equation \((L - z)f = g\). This can be done using the variation of constants formula from ordinary differential equations which determines the solution up to an arbitrary solution of the homogeneous equation. This homogeneous equation has to be chosen such that \( f \in \mathcal{D}(L) \), that is, such that \( f(0) = f(1) = 0 \).

Define

\[
    f(x) = \frac{u_+(z, x)}{W(z)} \left( \int_0^x u_-(z, t) g(t) dt \right) + \frac{u_-(z, x)}{W(z)} \left( \int_x^1 u_+(z, t) g(t) dt \right),
\]

where \( u_\pm(z, x) \) are the solutions of the homogeneous differential equation

\[
    -u_\pm''(z, x) + (q(x) - z) u_\pm(z, x) = 0
\]

satisfying the initial conditions \( u_-(z, 0) = 0, u'_-(z, 0) = 1 \) respectively \( u_+(z, 1) = 0, u'_+(z, 1) = 1 \) and

\[
    W(z) = W(u_+(z), u_-(z)) = u'_-(z, x) u_+(z, x) - u_-(z, x) u'_+(z, x)
\]

is the Wronski determinant, which is independent of \( x \) (check this!).

Then clearly \( f(0) = 0 \) since \( u_-(z, 0) = 0 \) and similarly \( f(1) = 0 \) since \( u_+(z, 1) = 0 \). Furthermore, \( f \) is differentiable and a straightforward computation verifies

\[
    f'(x) = \frac{u_+(z, x)'}{W(z)} \left( \int_0^x u_-(z, t) g(t) dt \right) + \frac{u_-(z, x)'}{W(z)} \left( \int_x^1 u_+(z, t) g(t) dt \right).
\]

Thus we can differentiate once more giving

\[
    f''(x) = \frac{u_+(z, x)''}{W(z)} \left( \int_0^x u_-(z, t) g(t) dt \right) + \frac{u_-(z, x)''}{W(z)} \left( \int_x^1 u_+(z, t) g(t) dt \right) - g(x)
\]

\[
    = (q(x) - z) f(x) - g(x).
\]

In summary, \( f \) is in the domain of \( L \) and satisfies \((L - z)f = g\).

Note that \( z \) is an eigenvalue if and only if \( W(z) = 0 \). In fact, in this case \( u_+(z, x) \) and \( u_-(z, x) \) are linearly dependent and hence \( u_-(z, 1) = c u_+(z, 1) = 0 \) which shows that \( u_-(z, x) \) satisfies both boundary conditions and is thus an eigenfunction.
Introducing the Green function
\[ G(z, x, t) = \frac{1}{W(u_+(z), u_-(z))} \begin{cases} u_+(z, x)u_-(z, t), & x \geq t \\ u_+(z, t)u_-(z, x), & x \leq t \end{cases} \] (3.18)
we see that \((L - z)^{-1}\) is given by
\[ (L - z)^{-1}g(x) = \int_0^1 G(z, x, t)g(t)dt. \] (3.19)
Moreover, from \(G(z, x, t) = G(z, t, x)\) it follows that \((L - z)^{-1}\) is symmetric for \(z \in \mathbb{R}\) (Problem 3.7) and from Lemma 3.5 it follows that it is compact. Hence Theorem 3.8 applies to \((L - z)^{-1}\) and we obtain:

**Theorem 3.11.** The Sturm–Liouville operator \(L\) has a countable number of eigenvalues \(E_n\). All eigenvalues are discrete and simple. The corresponding normalized eigenfunctions \(u_n\) form an orthonormal basis for \(L^2_{\text{cont}}(0, 1)\).

**Proof.** Pick a value \(\lambda \in \mathbb{R}\) such that \(R_L(\lambda)\) exists. By Lemma 3.5 \(R_L(\lambda)\) is compact and by Lemma 3.3 this remains true if we replace \(L^2_{\text{cont}}(0, 1)\) by its completion. By Theorem 3.8 there are eigenvalues \(\alpha_n\) of \(R_L(\lambda)\) with corresponding eigenfunctions \(u_n\). Moreover, \(R_L(\lambda)u_n = \alpha_n u_n\) is equivalent to \(Lu_n = (\lambda + \frac{1}{\alpha_n})u_n\), which shows that \(E_n = \lambda + \frac{1}{\alpha_n}\) are eigenvalues of \(L\) with corresponding eigenfunctions \(u_n\). Now everything follows from Theorem 3.8 except that the eigenvalues are simple. To show this, observe that if \(u_n\) and \(v_n\) are two different eigenfunctions corresponding to \(E_n\), then \(u_n(0) = v_n(0) = 0\) implies \(W(u_n, v_n) = 0\) and hence \(u_n\) and \(v_n\) are linearly dependent. \(\Box\)

**Example.** Let us look at the Sturm–Liouville problem with \(q = 0\). Then the underlying differential equation is
\[-u''(x) = z u(x)\]
whose solution is given by \(u(x) = c_1 \sin(\sqrt{z}x) + c_2 \cos(\sqrt{z}x)\). The solutions satisfying the boundary condition at the left endpoint is \(u_-(z, x) = \sin(\sqrt{z}x)\) and it will be an eigenfunction if and only if \(u_-(z, 1) = \sin(\sqrt{z}) = 0\). Hence the corresponding eigenvalues and normalized eigenfunctions are
\[ E_n = \pi^2 n^2, \quad u_n(x) = \sqrt{2} \sin(n\pi x), \quad n \in \mathbb{N}. \]
Moreover, every function \(f \in \mathcal{S}_0\) can be expanded into a —bf Fourier sine series
\[ f(x) = \sum_{n=1}^{\infty} f_n u_n(x), \quad f_n = \int_0^1 u_n(x)f(x)dx, \]
which is convergent with respect to our scalar product. \(\diamond\)
Example. We could also look at the same equation as in the previous problem but with different boundary conditions
\[ u'(0) = u'(1) = 0. \]

Then
\[ E_n = \frac{\pi^2 n^2}{2}, \quad u_n(x) = \begin{cases} 1, & n = 0, \\ \sqrt{2} \cos(n\pi x), & n \in \mathbb{N}. \end{cases} \]

Moreover, every function \( f \in C_0 \) can be expanded into a Fourier cosine series
\[ f(x) = \sum_{n=1}^{\infty} f_n u_n(x), \quad f_n = \int_0^1 u_n(x) f(x) dx, \]
which is convergent with respect to our scalar product. ◊

Example. Combining the last two examples we see that every symmetric function on \([-1,1]\) can be expanded into a Fourier cosine series and every anti-symmetric function into a Fourier sine series. Moreover, since every function \( f(x) \) can be written as the sum of a symmetric function \( f(x)+f(-x) \) and an anti-symmetric function \( \frac{f(x)-f(-x)}{2} \), it can be expanded into a Fourier series. ◊

Problem 3.7. Show that for our Sturm–Liouville operator \( u_\pm(z,x)^* = u_\pm(z^*,x) \). Conclude \( R_L(z)^* = R_L(z^*) \). (Hint: Problem 3.2.)

Problem 3.8. Show that the resolvent \( R_A(z) = (A-z)^{-1} \) (provided it exists and is densely defined) of a symmetric operator \( A \) is again symmetric for \( z \in \mathbb{R} \). (Hint: \( g \in \mathcal{D}(R_A(z)) \) if and only if \( g = (A-z)f \) for some \( f \in \mathcal{D}(A) \)).

Problem 3.9. Consider the Sturm–Liouville problem on a compact interval \([a,b]\) with domain
\[ \mathcal{D}(L) = \{ f \in C^2[a,b] | f'(a) - \alpha f(a) = f'(b) - \beta f(b) = 0 \} \]
for some real constants \( \alpha, \beta \in \mathbb{R} \). Show that Theorem 3.11 continues to hold.

3.4. Estimating eigenvalues

In general, there is no way of computing eigenvalues and their corresponding eigenfunctions explicitly. Hence it is important to be able to determine the eigenvalues at least approximately.

Let \( A \) be a self-adjoint operator which has a lowest eigenvalue \( \alpha_1 \) (e.g., \( A \) is compact or \( A \) is a Sturm–Liouville operator). Suppose we have a vector
3.4. Estimating eigenvalues

$f$ which is an approximation for the eigenvector $u_1$ of this lowest eigenvalue $\alpha_1$. Moreover, suppose we can write

$$A = \sum_{j=1}^{\infty} \alpha_j \langle u_j, \cdot \rangle u_j, \quad \mathcal{D}(A) = \{ f \in H \mid \sum_{j=1}^{\infty} |\alpha_j \langle u_j, f \rangle|^2 < \infty \}, \quad (3.20)$$

where $\{u_j\}_{j \in \mathbb{N}}$ is an orthonormal basis of eigenvectors. Since $\alpha_1$ is supposed to be the lowest eigenvalue we have $\alpha_j \geq \alpha_1$ for all $j \in \mathbb{N}$.

Writing $f = \sum_j \gamma_j u_j$, $\gamma_j = \langle u_j, u \rangle$, one computes

$$\langle f, Af \rangle = \langle f, \sum_{j=1}^{\infty} \alpha_j \gamma_j u_j \rangle = \sum_{j=1}^{\infty} \alpha_j |\gamma_j|^2, \quad f \in \mathcal{D}(A), \quad (3.21)$$

and we clearly have

$$\alpha_1 \leq \frac{\langle f, Af \rangle}{\|f\|^2}, \quad f \in \mathcal{D}(A), \quad (3.22)$$

with equality for $f = u_1$. In particular, any $u$ will provide an upper bound and if we add some free parameters to $u$, one can optimize them and obtain quite good upper bounds for the first eigenvalue. For example we could take some orthogonal basis, take a finite number of coefficients and optimize them. This is known as the Rayleigh–Ritz method.

Example. Consider the Sturm–Liouville operator $L$ with potential $q(x) = x$ and Dirichlet boundary conditions $f(0) = f(1) = 0$ in the interval $[0, 1]$. First of all note that integration by parts shows

$$\langle f, Lf \rangle = \int_0^1 (|f'(x)|^2 + q(x)|f(x)|^2) \, dx.$$  

which gives us the lower bound

$$\langle f, Lf \rangle \geq \min_{0 \leq x \leq 1} q(x) = 0.$$  

While the corresponding differential equation can in principle be solved in terms of Airy functions, there is no closed form for the eigenvalues.

First of all we can improve the above bound upon observing $0 \leq q(x) \leq 1$ which implies

$$\langle f, L_0 f \rangle \leq \langle f, Lf \rangle \leq \langle f, (L_0 + 1)f \rangle, \quad f \in \mathcal{D}(L) = \mathcal{D}(L_0),$$

where $L_0$ is the Sturm–Liouville operator corresponding to $q(x) = 0$. Since the lowest eigenvalue of $L_0$ is $\pi^2$ we obtain

$$\pi^2 \leq E_1 \leq \pi^2 + 1$$

for the lowest eigenvalue $E_1$ of $L$. 
Moreover, using the lowest eigenfunction \( f_1(x) = \sqrt{2} \sin(\pi x) \) of \( L_0 \) one obtains the improved upper bound
\[
E_1 \leq \langle f_1, Af_1 \rangle = \pi^2 + \frac{1}{2} \approx 10.3696.
\]
Taking the second eigenfunction \( f_2(x) = \sqrt{2} \sin(2\pi x) \) of \( L_0 \) we can make the ansatz \( f(x) = (1 + \gamma^2)^{-1/2} (f_1(x) + \gamma f_2(x)) \) which gives
\[
\langle f, Af \rangle = \pi^2 + \frac{1}{2} + \frac{\gamma}{1 + \gamma^2} (3\pi^2 \gamma - \frac{32}{9\pi^2}).
\]
The right-hand side has a unique minimum at \( \gamma = \frac{32}{27\pi^4 + \sqrt{1024 + 729\pi^8}} \) giving the bound
\[
E_1 \leq \frac{5}{2} \pi^2 + \frac{1}{2} - \frac{\sqrt{1024 + 729\pi^8}}{18\pi^2} \approx 10.3685
\]
which coincides with the exact eigenvalue up to five digits. \( \diamond \)

But is there also something one can say about the next eigenvalues? Suppose we know the first eigenfunction \( u_1 \). Then we can restrict \( A \) to the orthogonal complement of \( u_1 \) and proceed as before: \( E_2 \) will be the minimum of \( \langle f, Af \rangle \) over all \( f \) restricted to this subspace. If we restrict to the orthogonal complement of an approximating eigenfunction \( f_1 \), there will still be a component in the direction of \( u_1 \) left and hence the infimum of the expectations will be lower than \( E_2 \). Thus the optimal choice \( f_1 = u_1 \) will give the maximal value \( E_2 \).

**Theorem 3.12 (Max-min).** Let \( A \) be a self-adjoint operator and let \( \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_N \) be eigenvalues of \( A \) with corresponding orthonormal eigenvectors \( u_1, u_2, \ldots, u_N \). Suppose
\[
A = \sum_{j=1}^{N} \alpha_j \langle u_j, \cdot \rangle u_j + \tilde{A} \tag{3.23}
\]
with \( \tilde{A} \geq \alpha_N \). Then
\[
\alpha_j = \sup_{f_1, \ldots, f_{j-1}} \inf_{f \in U(f_1, \ldots, f_{j-1})} \langle f, Af \rangle, \quad 1 \leq j \leq N, \tag{3.24}
\]
where
\[
U(f_1, \ldots, f_j) = \{ f \in \mathcal{D}(A) \| f \| = 1, \ f \in \text{span}\{f_1, \ldots, f_j \}^\perp \}. \tag{3.25}
\]

**Proof.** We have
\[
\inf_{f \in U(f_1, \ldots, f_{j-1})} \langle f, Af \rangle \leq \alpha_j.
\]
3.4. Estimating eigenvalues

In fact, set \( f = \sum_{k=1}^{j} \gamma_k u_k \) and choose \( \gamma_k \) such that \( f \in U(f_1, \ldots, f_{j-1}) \). Then

\[
\langle f, Af \rangle = \sum_{k=1}^{j} |\gamma_k|^2 \alpha_k \leq \alpha_j
\]

and the claim follows.

Conversely, let \( \gamma_k = \langle u_k, f \rangle \) and write \( f = \sum_{k=1}^{j} \gamma_k u_k + \tilde{f} \). Then

\[
\inf_{f \in U(u_1, \ldots, u_{j-1})} \langle f, Af \rangle = \inf_{f \in U(u_1, \ldots, u_{j-1})} \left( \sum_{k=j}^{N} |\gamma_k|^2 \alpha_k + \langle \tilde{f}, A \tilde{f} \rangle \right) = \alpha_j. \quad \Box
\]

Of course if we are interested in the largest eigenvalues all we have to do is consider \(-A\).

Note that this immediately gives an estimate for eigenvalues if we have an corresponding estimate for the operators. To this end we will write

\[
A \leq B \iff \langle f, Af \rangle \leq \langle f, Bf \rangle, \quad f \in \mathcal{D}(A) \cap \mathcal{D}(B). \quad (3.26)
\]

**Corollary 3.13.** Suppose \( A \) and \( B \) are self-adjoint operators with corresponding eigenvalues \( \alpha_j \) and \( \beta_j \) as in the previous theorem. If \( A \leq B \) and \( \mathcal{D}(B) \subseteq \mathcal{D}(A) \) then \( \alpha_j \leq \beta_j \).

**Proof.** By assumption we have \( \langle f, Af \rangle \leq \langle f, Bf \rangle \) for \( f \in \mathcal{D}(B) \) implying

\[
\inf_{f \in U_A(f_1, \ldots, f_{j-1})} \langle f, Af \rangle \leq \inf_{f \in U_B(f_1, \ldots, f_{j-1})} \langle f, Af \rangle \leq \inf_{f \in U_B(f_1, \ldots, f_{j-1})} \langle f, Bf \rangle,
\]

where we have indicated the dependence of \( U \) on the operator via a subscript. Taking the sup on both sides the claim follows. \quad \Box

**Example.** Let \( L \) be again our Sturm–Liouville operator and \( L_0 \) the corresponding operator with \( q(x) = 0 \). Set \( q_+ = \min_{0 \leq x \leq 1} q(x) \) and \( q_- = \max_{0 \leq x \leq 1} q(x) \). Then \( L_0 + q_- \leq L \leq L_0 + q_+ \) implies

\[
\pi^2 n^2 + q_- \leq E_n \leq \pi^2 n^2 + q_+.
\]

In particular, we have proven the famous **Weyl asymptotic**

\[
E_n = \pi^2 n^2 + O(1)
\]

for the eigenvalues. \quad \diamond

There is also an alternative version which can be proven similar (Problem 3.10):

**Theorem 3.14 (Min-max).** Let \( A \) be as in the previous theorem. Then

\[
\alpha_j = \inf_{V_j \subseteq \mathcal{D}(A), \dim(V_j) = j} \sup_{f \in V_j, \|f\| = 1} \langle f, Af \rangle, \quad (3.27)
\]

where the inf is taken over subspaces with the indicated properties.

Problem 3.11. Suppose $A, A_n$ are self-adjoint, bounded and $A_n \to A$. Then $\alpha_k(A_n) \to \alpha_k(A)$. (Hint: $\|B\| \leq \varepsilon$ is equivalent to $-\varepsilon \leq B \leq \varepsilon$.)
Chapter 4

The main theorems about Banach spaces

4.1. The Baire theorem and its consequences

Recall that the interior of a set is the largest open subset (that is, the union of all open subsets). A set is called nowhere dense if its closure has empty interior. The key to several important theorems about Banach spaces is the observation that a Banach space cannot be the countable union of nowhere dense sets.

**Theorem 4.1** (Baire category theorem). Let $X$ be a (nonempty) complete metric space. Then $X$ cannot be the countable union of nowhere dense sets.

**Proof.** Suppose $X = \bigcup_{n=1}^{\infty} X_n$. We can assume that the sets $X_n$ are closed and none of them contains a ball; that is, $X \setminus X_n$ is open and nonempty for every $n$. We will construct a Cauchy sequence $x_n$ which stays away from all $X_n$.

Since $X \setminus X_1$ is open and nonempty, there is a closed ball $B_{r_1}(x_1) \subseteq X \setminus X_1$. Reducing $r_1$ a little, we can even assume $\overline{B_{r_1}(x_1)} \subseteq X \setminus X_1$. Moreover, since $X_2$ cannot contain $B_{r_1}(x_1)$, there is some $x_2 \in B_{r_2}(x_1)$ that is not in $X_2$. Since $B_{r_1}(x_1) \cap (X \setminus X_2)$ is open, there is a closed ball $\overline{B_{r_2}(x_2)} \subseteq \overline{B_{r_1}(x_1)} \cap (X \setminus X_2)$. Proceeding recursively, we obtain a sequence (here we use the axiom of choice) of balls such that

$$\overline{B_{r_n}(x_n)} \subseteq \overline{B_{r_{n-1}}(x_{n-1})} \cap (X \setminus X_n).$$

Now observe that in every step we can choose $r_n$ as small as we please; hence without loss of generality $r_n \to 0$. Since by construction $x_n \in B_{r_N}(x_N)$ for
n ≥ N, we conclude that xₙ is Cauchy and converges to some point x ∈ X. But x ∈ Bᵣₙ(xₙ) ⊆ X \ Xₙ for every n, contradicting our assumption that the Xₙ cover X.

Remark: The set of rational numbers Q can be written as a countable union of its elements. This shows that the completeness assumption is crucial.

(Sets which can be written as the countable union of nowhere dense sets are said to be of first category. All other sets are second category. Hence we have the name category theorem.)

In other words, if Xₙ ⊆ X is a sequence of closed subsets which cover X, at least one Xₙ contains a ball of radius ε > 0.

Since a closed set is nowhere dense if and only if its complement is open and dense (cf. Problem 1.5), there is a reformulation which is also worthwhile noting:

Corollary 4.2. Let X be a complete metric space. Then any countable intersection of open dense sets is again dense.

Proof. Let Oₙ be open dense sets whose intersection is not dense. Then this intersection must be missing some closed ball Bᵣ. This ball will lie in ∪ₙ Xₙ, where Xₙ = X \ Oₙ are closed and nowhere dense. Now note that Xₙ = Xₙ ∪ B are closed nowhere dense sets in Bᵣ. But Bᵣ is a complete metric space, a contradiction.

Now we come to the first important consequence, the uniform boundedness principle.

Theorem 4.3 (Banach–Steinhaus). Let X be a Banach space and Y some normed vector space. Let \{Aₐ\} ⊆ L(X,Y) be a family of bounded operators. Suppose \|Aₐx\| ≤ C(x) is bounded for fixed x ∈ X. Then \{Aₐ\} is uniformly bounded, \|Aₐ\| ≤ C.

Proof. Let

\[ Xₙ = \{x| \|Aₐx\| ≤ n \text{ for all } α\} = \bigcap_α \{x| \|Aₐx\| ≤ n\}. \]

Then ∪ₙ Xₙ = X by assumption. Moreover, by continuity of Aₐα and the norm, each Xₙ is an intersection of closed sets and hence closed. By Baire’s theorem at least one contains a ball of positive radius: Bᵣ(x₀) ⊂ Xₙ. Now observe

\[ \|Aₐy\| ≤ \|Aₐ(y + x₀)\| + \|Aₐx₀\| ≤ n + C(x₀) \]
for \( \|y\| \leq \varepsilon \). Setting \( y = \varepsilon \frac{x}{\|x\|} \), we obtain
\[
\|A_\alpha x\| \leq \frac{n + C(x_0)}{\varepsilon} \|x\|
\]
for every \( x \).

Note that if the assumptions of the previous theorem fail for one \( x \), they must in fact fail for \( x \) in a dense set. In fact, if
\[
\sup_\alpha \|A_\alpha x_0\| = \infty
\]
for one \( x_0 \in X \) then for any other \( x \in X \) there are two possibilities. Either
\[
\sup_\alpha \|A_\alpha x\| = \infty
\]
or
\[
\sup_\alpha \|A_\alpha (x + \varepsilon x_0)\| = \infty
\]
for every \( \varepsilon \neq 0 \).

**Example.** Consider the Fourier series \((2.44)\) of a continuous periodic function \( f \in C_{\text{per}}[-\pi, \pi] = \{f \in C[-\pi, \pi]|f(-\pi) = f(\pi)\} \). (Note that this is a closed subspace of \( C[-\pi, \pi] \) and hence a Banach space — it is the kernel of the linear functional \( \ell(f) = f(-\pi) - f(\pi) \).) We want to show that for every fixed \( x \in [-\pi, \pi] \) there is a dense set of function in \( C_{\text{per}}[-\pi, \pi] \) for which the Fourier series will diverge at \( x \).

To show this, suppose this were true. Without loss of generality we fix \( x = 0 \) as our point. Then \( n \)’th partial sum gives rise to the linear functional
\[
\ell_n(f) = S_n(f) = (2\pi)^{-1} \int_{-\pi}^{\pi} D_n(x) f(x) dx
\]
which converges for fixed \( f \) and hence in particular must be bounded. \( |\ell_n(f)| \leq C(f) \). Hence by the uniform boundedness principle we have \( \|\ell_n\| \leq C \). On the other hand, by the example on page 42 (adapted to our present periodic setting) we have
\[
\|\ell_n\| = \|D_n\|_1.
\]
Now we estimate
\[
\|D_n\|_1 = \frac{1}{\pi} \int_0^{\pi} |D_n(x)| dx \geq \frac{1}{\pi} \int_0^{\pi} \frac{|\sin((n + 1/2)x)|}{x/2} dx
\]
\[
= 2 \pi \int_0^{(n+1/2)\pi} |\sin(y)| \frac{dy}{y} \geq 2 \pi \sum_{k=1}^{n} \int_{(k-1)\pi}^{k\pi} |\sin(y)| \frac{dy}{k\pi}
\]
\[
= \frac{4}{\pi^2} \sum_{k=1}^{n} \frac{1}{k}
\]
and we get a contradiction since the harmonic series diverges.

**Example.** Recall that the Fourier coefficients of an absolutely continuous function satisfy the estimate
\[
|\hat{f}_k| \leq \begin{cases} 
\|f\|_\infty, & k = 0, \\
\|f'_\infty\|_\infty / |k|, & k \neq 0.
\end{cases}
\]
and this raises the question if a similar estimate can be true for continuous functions. More precisely, can we find a sequence \( c_k > 0 \) such that
\[
|\hat{f}_k| \leq C f c_k,
\]
where \( C_f \) is some constant depending on \( f \). If this were true the linear functionals
\[
\ell_k(f) = \frac{\hat{f}_k}{c_k}, \quad k \in \mathbb{Z},
\]
satisfy the assumptions of the uniform boundedness principle implying \( \|\ell_k\| \leq C \). In other words, we must have an estimate of the type
\[
|\hat{f}_k| \leq C \|f\|_{\infty} c_k
\]
which implies \( 1 \leq C c_k \) upon choosing \( f(x) = e^{ikx} \). Hence our assumption cannot hold for any sequence \( c_k \) converging to zero. Hence there is no universal decay rate for the Fourier coefficients of continuous functions beyond the fact that they must converge to zero by the Riemann–Lebesgue lemma. ⋄

The next application is

**Theorem 4.4 (Open mapping).** Let \( A \in \mathcal{L}(X,Y) \) be a bounded linear operator from one Banach space onto another. Then \( A \) is open (i.e., maps open sets to open sets).

**Proof.** Denote by \( B^X_r(x) \subseteq X \) the open ball with radius \( r \) centered at \( x \) and let \( B^X_r = B^X_r(0) \). Similarly for \( B^Y_r(y) \). By scaling and translating balls (using linearity of \( A \)), it suffices to prove \( B^Y_\varepsilon \subseteq A(B^X_1) \) for some \( \varepsilon > 0 \). Since \( A \) is surjective we have
\[
Y = \bigcup_{n=1}^{\infty} A(B^X_n)
\]
and the Baire theorem implies that for some \( n, \overline{A(B^X_n)} \) contains a ball \( B^Y_\varepsilon(y) \). Without restriction \( n = 1 \) (just scale the balls). Since \( -\overline{A(B^X_1)} = \overline{A(-B^X_1)} = \overline{A(B^X_1)} \) we see \( B^Y_\varepsilon(-y) \subseteq \overline{A(B^X_1)} \) and by convexity of \( \overline{A(B^X_1)} \) we also have \( B^Y_\varepsilon \subseteq \overline{A(B^X_1)} \).

So we have \( B^Y_\varepsilon \subseteq \overline{A(B^X_1)} \), but we would need \( B^Y_\varepsilon \subseteq A(B^X_1) \). To complete the proof we will show \( \overline{A(B^X_1)} \subseteq A(B^X_{1/2}) \) which implies \( B^Y_{\varepsilon/2} \subseteq A(B^X_{1/2}) \).

For every \( y \in \overline{A(B^X_1)} \) we can choose some sequence \( y_n \in A(B^X_1) \) with \( y_n \to y \). Moreover, there even is some \( x_n \in B^X_1 \) with \( y_n = A(x_n) \). However, \( x_n \) might not converge, so we need to argue more carefully and ensure convergence along the way: start with \( x_1 \in B^X_1 \) such that \( y - Ax_1 \in B^Y_{\varepsilon/2} \). Scaling the relation \( B^Y_\varepsilon \subseteq \overline{A(B^X_1)} \) we have \( B^Y_{\varepsilon/2} \subseteq \overline{A(B^X_{1/2})} \) and hence we can
choose \( x_2 \in B_{1/2}^X \) such that \((y - Ax_1) - Ax_2 \in B_{\varepsilon/4}^Y \subseteq A(B_{1/4}^X)\). Proceeding like this we obtain a sequence of points \( x_n \in B_{2^{-n}}^X \) such that
\[
y - \sum_{k=1}^{n} Ax_k \in B_{\varepsilon/2^{-n}}^Y.
\]
By \( \|x_k\| < 2^{1-k} \) the limit \( x = \sum_{k=1}^{\infty} x_k \) exists and satisfies \( \|x\| < 2 \). Hence \( y = Ax \in A(B_2^X) \) as desired. \( \square \)

Remark: The requirement that \( A \) is onto is crucial (just look at the one-dimensional case \( X = \mathbb{C} \)). Moreover, the converse is also true: If \( A \) is open, then the image of the unit ball contains again some ball \( B_Y^\varepsilon \subseteq A(B_X^1) \). Hence by scaling \( B_Y^\varepsilon \subseteq A(B_X^r) \) and letting \( r \to \infty \) we see that \( A \) is onto: \( Y = A(X) \).

As an immediate consequence we get the inverse mapping theorem:

**Theorem 4.5 (Inverse mapping).** Let \( A \in \mathcal{L}(X,Y) \) be a bounded linear bijection between Banach spaces. Then \( A^{-1} \) is continuous.

**Example.** Consider the operator \((Aa)^n_{j=1} = (\frac{1}{j}a_j)^n_{j=1} \) in \( \ell^2(\mathbb{N}) \). Then its inverse \((A^{-1}a)^n_{j=1} = (j a_j)^n_{j=1} \) is unbounded (show this!). This is in agreement with our theorem since its range is dense (why?) but not all of \( \ell^2(\mathbb{N}) \): For example, \((b_j = \frac{1}{j})^\infty_{j=1} \notin \text{Ran}(A)\) since \( b = Aa \) gives the contradiction
\[
\infty = \sum_{j=1}^\infty 1 = \sum_{j=1}^\infty |jb_j|^2 = \sum_{j=1}^\infty |a_j|^2 < \infty.
\]
This should also be compared with Corollary 4.8 below. \( \diamond \)

Another important consequence is the closed graph theorem. The **graph** of an operator \( A \) is just
\[
\Gamma(A) = \{(x, Ax) \mid x \in \mathcal{D}(A)\}. \tag{4.1}
\]
If \( A \) is linear, the graph is a subspace of the Banach space \( X \oplus Y \) (provided \( X \) and \( Y \) are Banach spaces), which is just the cartesian product together with the norm
\[
\|(x, y)\|_{X \oplus Y} = \|x\|_X + \|y\|_Y. \tag{4.2}
\]
Note that \( (x_n, y_n) \to (x, y) \) if and only if \( x_n \to x \) and \( y_n \to y \). We say that \( A \) has a closed graph if \( \Gamma(A) \) is a closed set in \( X \oplus Y \).

**Theorem 4.6 (Closed graph).** Let \( A : X \to Y \) be a linear map from a Banach space \( X \) to another Banach space \( Y \). Then \( A \) is bounded if and only if its graph is closed.
**Proof.** If $\Gamma(A)$ is closed, then it is again a Banach space. Now the projection $\pi_1(x, Ax) = x$ onto the first component is a continuous bijection onto $X$. So by the inverse mapping theorem its inverse $\pi_1^{-1}$ is again continuous. Moreover, the projection $\pi_2(x, Ax) = Ax$ onto the second component is also continuous and consequently so is $A = \pi_2 \circ \pi_1^{-1}$. The converse is easy. □

Remark: The crucial fact here is that $A$ is defined on all of $X$!

Operators whose graphs are closed are called **closed operators**. Being closed is the next option you have once an operator turns out to be unbounded. If $A$ is closed, then $x_n \to x$ does not guarantee you that $Ax_n$ converges (like continuity would), but it at least guarantees that if $Ax_n$ converges, it converges to the right thing, namely $Ax$:

- $A$ bounded (with $\mathcal{D}(A) = X$): $x_n \to x$ implies $Ax_n \to Ax$.
- $A$ closed (with $\mathcal{D}(A) \subseteq X$): $x_n \to x$, $x_n \in \mathcal{D}(A)$, and $Ax_n \to y$ implies $x \in \mathcal{D}(A)$ and $y = Ax$.

If an operator is not closed, you can try to take the closure of its graph, to obtain a closed operator. If $A$ is bounded this always works (which is just the content of Theorem 1.35). However, in general, the closure of the graph might not be the graph of an operator as we might pick up points $(x, y_2) - (x, y_1) = (0, y_2 - y_1) \in \Gamma(A)$ in this case and thus $\Gamma(A)$ is the graph of some operator if and only if

$$\Gamma(A) \cap \{(0, y) | y \in Y\} = \{(0, 0)\}. \tag{4.3}$$

If this is the case, $A$ is called **closable** and the operator $\overline{A}$ associated with $\Gamma(A)$ is called the **closure** of $A$.

In particular, $A$ is closable if and only if $x_n \to 0$ and $Ax_n \to y$ implies $y = 0$. In this case

$$\mathcal{D}(\overline{A}) = \{x \in X | \exists x_n \in \mathcal{D}(A), y \in Y : x_n \to x \text{ and } Ax_n \to y\},$$

$$\overline{A}x = y. \tag{4.4}$$

For yet another way of defining the closure see Problem 4.4.

**Example.** Consider the operator $A$ in $\ell^p(\mathbb{N})$ defined by $Aa_j = ja_j$ on $\mathcal{D}(A) = \{a \in \ell^p(\mathbb{N}) | a_j \neq 0 \text{ for finitely many } j\}$.

(i). $A$ is closable. In fact, if $a^n \to 0$ and $Aa^n \to b$ then we have $a^n_j \to 0$ and thus $ja^n_j \to 0 = b_j$ for any $j \in \mathbb{N}$.

(ii). The closure of $A$ is given by

$$\mathcal{D}(\overline{A}) = \begin{cases} \{a \in \ell^p(\mathbb{N}) | (ja_j)_{j=1}^\infty \in \ell^p(\mathbb{N})\}, & 1 \leq p < \infty, \\ \{a \in c_0(\mathbb{N}) | (ja_j)_{j=1}^\infty \in c_0(\mathbb{N})\}, & p = \infty, \end{cases}$$
and $\overline{A}a_j = ja_j$. In fact, if $a^n \to a$ and $Aa^n \to b$ then we have $a^n_j \to a_j$ and $ja^n_j \to b_j$ for any $j \in \mathbb{N}$ and thus $b_j = ja_j$ for any $j \in \mathbb{N}$. In particular, $(ja_j)_j=1^\infty = (b_j)_j=1^\infty \in \ell^p(\mathbb{N})$ ($c_0(\mathbb{N})$ if $p = \infty$). Conversely, suppose $(ja_j)_j=1^\infty \in \ell^p(\mathbb{N})$ ($c_0(\mathbb{N})$ if $p = \infty$) and consider

$$a^n_j = \begin{cases} a_j, & j \leq n, \\ 0, & j > n. \end{cases}$$

Then $a^n \to a$ and $Aa^n \to (ja_j)_j=1^\infty$.

(iii). Note that the inverse of $\overline{A}$ is the bounded operator $\overline{A}^{-1}a_j = j^{-1}a_j$ defined on all of $\ell^p(\mathbb{N})$. Thus $\overline{A}^{-1}$ is closed. However, since its range $\text{Ran}(\overline{A}^{-1}) = \mathcal{D}(\overline{A})$ is dense but not all of $\ell^p(\mathbb{N})$, $\overline{A}^{-1}$ does not map closed sets to closed sets in general. In particular, the concept of a closed operator should not be confused with the concept of a closed map in topology!

(iv). Extending the basis vectors $\{\delta_n\}_{n \in \mathbb{N}}$ to a Hamel basis (Problem 1.14) and setting $Aa = 0$ for every other element from this Hamel basis we obtain a (still unbounded) operator which is everywhere defined. However, this extension cannot be closed!

Example. Here is a simple example of a nonclosable operator: Let $X = \ell^2(\mathbb{N})$ and consider $Ba = (\sum_{j=1}^\infty a_j)\delta_1$ defined on $\ell^1(\mathbb{N}) \subset \ell^2(\mathbb{N})$. Let $a_j^n = \frac{1}{n}$ for $1 \leq j \leq n$ and $a_j^n = 0$ for $j > n$. Then $\|a^n\|_2 = \frac{1}{\sqrt{n}}$ implying $a^n \to 0$ but $Ba^n = \delta_1 \not\to 0$.

Lemma 4.7. Suppose $A$ is closable and $\overline{A}$ is injective. Then $\overline{A}^{-1} = \overline{A}^{-1}$. 

Proof. If we set

$$\Gamma^{-1} = \{(y,x)|(x,y) \in \Gamma\}$$

then $\Gamma(A^{-1}) = \Gamma^{-1}(A)$ and

$$\Gamma(A^{-1}) = \overline{\Gamma(A)}^{-1} = \overline{\Gamma(A)}^{-1} = \overline{\Gamma(A)}^{-1} = \overline{\Gamma(A)}^{-1}.$$ 

Note that $A$ injective does not imply $\overline{A}$ injective in general.

Example. Let $P_M$ be the projection in $\ell^2(\mathbb{N})$ on $M = \{b\}^\perp$, where $b = (2^{-j})_{j=1}^\infty$. Explicitly we have $P_Ma = a - \langle b,a \rangle b$. Then $P_M$ restricted to the space of sequences with finitely many nonzero terms is injective, but its closure is not.

As a consequence of the closed graph theorem we obtain:

Corollary 4.8. Suppose $A : \mathcal{D}(A) \to \text{Ran}(A)$ is closed and injective. Then $A^{-1}$ defined on $\mathcal{D}(A^{-1}) = \text{Ran}(A)$ is closed. Moreover, in this case $\text{Ran}(A)$ is closed if and only if $A^{-1}$ is bounded.
The closed graph theorem tells us that closed linear operators can be defined on all of $X$ if and only if they are bounded. So if we have an unbounded operator we cannot have both! That is, if we want our operator to be at least closed, we have to live with domains. This is the reason why in quantum mechanics most operators are defined on domains. In fact, there is another important property which does not allow unbounded operators to be defined on the entire space:

**Theorem 4.9** (Hellinger–Toeplitz). Let $A : \mathcal{H} \to \mathcal{H}$ be a linear operator on some Hilbert space $\mathcal{H}$. If $A$ is symmetric, that is $\langle g, Af \rangle = \langle Ag, f \rangle, \ f, g \in \mathcal{H}$, then $A$ is bounded.

**Proof.** It suffices to prove that $A$ is closed. In fact, $f_n \to f$ and $Af_n \to g$ implies
\[
\langle h, g \rangle = \lim_{n \to \infty} \langle h, Af_n \rangle = \lim_{n \to \infty} \langle Ah, f_n \rangle = \langle Ah, f \rangle = \langle h, Af \rangle
\]
for every $h \in \mathcal{H}$. Hence $Af = g$. □

**Problem 4.1.** Show that a compact symmetric operator in an infinite-dimensional Hilbert space cannot be surjective.

**Problem 4.2.** Show that if $A$ is closed and $B$ bounded, then $A + B$ is closed. Show that this in general fails if $B$ is not bounded. (Here $A + B$ is defined on $\mathcal{D}(A + B) = \mathcal{D}(A) \cap \mathcal{D}(B)$.)

**Problem 4.3.** Show that the differential operator $A = \frac{d}{dx}$ defined on $\mathcal{D}(A) = C^1[0,1] \subset C[0,1]$ (sup norm) is a closed operator. (Compare the example in Section 1.5.)

**Problem 4.4.** Consider a linear operator $A : \mathcal{D}(A) \subseteq X \to Y$, where $X$ and $Y$ are Banach spaces. Define the graph norm associated with $A$ by
\[
\|x\|_A = \|x\|_X + \|Ax\|_Y. \tag{4.5}
\]
Show that $A : \mathcal{D}(A) \to Y$ is bounded if we equip $\mathcal{D}(A)$ with the graph norm. Show that the completion $X_A$ of $(\mathcal{D}(A), \|\cdot\|_A)$ can be regarded as a subset of $X$ if and only if $A$ is closable. Show that in this case the completion can be identified with $\mathcal{D}(A)$ and that the closure of $A$ in $X$ coincides with the extension from Theorem 1.35 of $A$ in $X_A$.

**Problem 4.5.** Let $X = \ell^2(\mathbb{N})$ and $(Aa)_j = j a_j$ with $\mathcal{D}(A) = \{a \in \ell^2(\mathbb{N})| (ja_j)_{j \in \mathbb{N}} \in \ell^2(\mathbb{N})\}$ and $Ba = (\sum_{j \in \mathbb{N}} a_j)\delta_1$. Then we have seen that $A$ is closed while $B$ is not closable. Show that $A + B, \mathcal{D}(A + B) = \mathcal{D}(A) \cap \mathcal{D}(B) = \mathcal{D}(A)$ is closed.
4.2. The Hahn–Banach theorem and its consequences

Let $X$ be a Banach space. Recall that we have called the set of all bounded linear functionals the dual space $X^*$ (which is again a Banach space by Theorem 1.36).

**Example.** Consider the Banach space $\ell^p(\mathbb{N})$, $1 \leq p < \infty$. Taking the Kronecker deltas $\delta^n$ as a Schauder basis the $n$'th term $x_n$ of a sequence $x \in \ell^p(\mathbb{N})$ can also be considered as the $n$'th coordinate of $x$ with respect to this basis. Moreover, the map $l_n(x) = x_n$ is a bounded linear functional, that is, $l_n \in \ell^p(\mathbb{N})^*$, since $|l_n(x)| = |x_n| \leq \|x\|_p$. It is a special case of following more general example (in fact, we have $l_n = l_{\delta^n}$). Since the coordinates of a vector carry all the information this explains why understanding linear functionals if of a key importance.

**Example.** Consider the Banach space $\ell^p(\mathbb{N})$, $1 \leq p < \infty$. We have already seen that by Hölder’s inequality (1.29), every $y \in \ell^q(\mathbb{N})$ gives rise to a bounded linear functional

$$l_y(x) = \sum_{n \in \mathbb{N}} y_n x_n$$

(4.6)

whose norm is $\|l_y\| = \|y\|_q$ (Problem 4.8). But can every element of $\ell^p(\mathbb{N})^*$ be written in this form?

Suppose $p = 1$ and choose $l \in \ell^1(\mathbb{N})^*$. Now define

$$y_n = l(\delta^n),$$

(4.7)

where $\delta^n_n = 1$ and $\delta^n_m = 0$, $n \neq m$. Then

$$|y_n| = |l(\delta^n)| \leq \|l\| \|\delta^n\|_1 = \|l\|$$

(4.8)

shows $\|y\|_\infty \leq \|l\|$, that is, $y \in \ell^\infty(\mathbb{N})$. By construction $l(x) = l_y(x)$ for every $x \in \text{span}\{\delta^n\}$. By continuity of $l$ it even holds for $x \in \overline{\text{span}\{\delta^n\}} = \ell^1(\mathbb{N})$. Hence the map $y \mapsto l_y$ is an isomorphism, that is, $\ell^1(\mathbb{N})^* \cong \ell^\infty(\mathbb{N})$. A similar argument shows $\ell^p(\mathbb{N})^* \cong \ell^q(\mathbb{N})$, $1 \leq p < \infty$ (Problem 4.9). One usually identifies $\ell^p(\mathbb{N})^*$ with $\ell^q(\mathbb{N})$ using this canonical isomorphism and simply writes $\ell^p(\mathbb{N})^* = \ell^q(\mathbb{N})$. In the case $p = \infty$ this is not true, as we will see soon.

It turns out that many questions are easier to handle after applying a linear functional $l \in X^*$. For example, suppose $x(t)$ is a function $\mathbb{R} \to X$ (or $\mathbb{C} \to X$), then $\ell(x(t))$ is a function $\mathbb{R} \to \mathbb{C}$ (respectively $\mathbb{C} \to \mathbb{C}$) for any $\ell \in X^*$. So to investigate $\ell(x(t))$ we have all tools from real/complex analysis at our disposal. But how do we translate this information back to $x(t)$? Suppose we have $\ell(x(t)) = \ell(y(t))$ for all $\ell \in X^*$. Can we conclude $x(t) = y(t)$? The answer is yes and will follow from the Hahn–Banach theorem.
4. The main theorems about Banach spaces

We first prove the real version from which the complex one then follows easily.

**Theorem 4.10** (Hahn–Banach, real version). Let $X$ be a real vector space and $\varphi : X \to \mathbb{R}$ a convex function (i.e., $\varphi(\lambda x + (1-\lambda)y) \leq \lambda \varphi(x) + (1-\lambda)\varphi(y)$ for $\lambda \in (0,1)$).

If $\ell$ is a linear functional defined on some subspace $Y \subset X$ which satisfies $\ell(y) \leq \varphi(y)$, $y \in Y$, then there is an extension $\tilde{\ell}$ to all of $X$ satisfying $\tilde{\ell}(x) \leq \varphi(x)$, $x \in X$.

**Proof.** Let us first try to extend $\ell$ in just one direction: Take $x \not\in Y$ and set $\tilde{Y} = \text{span}\{x, Y\}$. If there is an extension $\tilde{\ell}$ to $\tilde{Y}$ it must clearly satisfy

$$\tilde{\ell}(y + \alpha x) = \ell(y) + \alpha \tilde{\ell}(x).$$

So all we need to do is to choose $\tilde{\ell}(x)$ such that $\tilde{\ell}(y + \alpha x) \leq \varphi(y + \alpha x)$. But this is equivalent to

$$\sup_{\alpha > 0, y \in Y} \frac{\varphi(y - \alpha x) - \ell(y)}{-\alpha} \leq \frac{\varphi(y + \alpha x) - \ell(y)}{\alpha}$$

and is hence only possible if

$$\frac{\varphi(y_1 - \alpha_1 x) - \ell(y_1)}{-\alpha_1} \leq \frac{\varphi(y_2 + \alpha_2 x) - \ell(y_2)}{\alpha_2}$$

for every $\alpha_1, \alpha_2 > 0$ and $y_1, y_2 \in Y$. Rearranging this last equations we see that we need to show

$$\alpha_2 \ell(y_1) + \alpha_1 \ell(y_2) \leq \alpha_2 \varphi(y_1 - \alpha_1 x) + \alpha_1 \varphi(y_2 + \alpha_2 x).$$

Starting with the left-hand side we have

$$\alpha_2 \ell(y_1) + \alpha_1 \ell(y_2) = (\alpha_1 + \alpha_2) \ell(\lambda y_1 + (1 - \lambda)y_2)$$

$$\leq (\alpha_1 + \alpha_2) \varphi(\lambda y_1 + (1 - \lambda)y_2)$$

$$= (\alpha_1 + \alpha_2) \varphi(\lambda(y_1 - \alpha_1 x) + (1 - \lambda)(y_2 + \alpha_2 x))$$

$$\leq \alpha_2 \varphi(y_1 - \alpha_1 x) + \alpha_1 \varphi(y_2 + \alpha_2 x),$$

where $\lambda = \frac{\alpha_2}{\alpha_1 + \alpha_2}$. Hence one dimension works.

To finish the proof we appeal to Zorn’s lemma (see Appendix A): Let $E$ be the collection of all extensions $\tilde{\ell}$ satisfying $\tilde{\ell}(x) \leq \varphi(x)$. Then $E$ can be partially ordered by inclusion (with respect to the domain) and every linear chain has an upper bound (defined on the union of all domains). Hence there is a maximal element $\tilde{\ell}$ by Zorn’s lemma. This element is defined on $X$, since if it were not, we could extend it as before contradicting maximality. \hfill \Box

Note that linearity gives us a corresponding lower bound $-\varphi(-x) \leq \tilde{\ell}(x)$, $x \in X$, for free. In particular, if $\varphi(x) = \varphi(-x)$ then $|\tilde{\ell}(x)| \leq \varphi(x)$. 
4.2. The Hahn–Banach theorem and its consequences

**Theorem 4.11** (Hahn–Banach, complex version). Let $X$ be a complex vector space and $\varphi : X \to \mathbb{R}$ a convex function satisfying $\varphi(\alpha x) \leq \varphi(x)$ if $|\alpha| = 1$.

If $\ell$ is a linear functional defined on some subspace $Y \subset X$ which satisfies $|\ell(y)| \leq \varphi(y)$, $y \in Y$, then there is an extension $\overline{\ell}$ to all of $X$ satisfying $|\overline{\ell}(x)| \leq \varphi(x)$, $x \in X$.

**Proof.** Set $\ell_r = \text{Re}(\ell)$ and observe $\ell(x) = \ell_r(x) - i\ell_r(ix)$.

By our previous theorem, there is a real linear extension $\overline{\ell}_r$ satisfying $\overline{\ell}_r(x) \leq \varphi(x)$. Now set $\overline{\ell}(x) = \overline{\ell}_r(x) - i\overline{\ell}_r(ix)$. Then $\overline{\ell}(x)$ is real linear and by $\overline{\ell}(ix) = \overline{\ell}_r(iz) + i\overline{\ell}_r(x) = i\overline{\ell}(x)$ also complex linear. To show $|\overline{\ell}(x)| \leq \varphi(x)$ we abbreviate $\alpha = \frac{\overline{\ell}(x)^*}{\overline{\ell}(x)}$ and use

$$|\overline{\ell}(x)| = \alpha \overline{\ell}(x) = \overline{\ell}(\alpha x) = \overline{\ell}_r(\alpha x) \leq \varphi(\alpha x) \leq \varphi(x),$$

which finishes the proof. \(\square\)

Note that $\varphi(\alpha x) \leq \varphi(x)$, $|\alpha| = 1$ is in fact equivalent to $\varphi(\alpha x) = \varphi(x)$, $|\alpha| = 1$.

If $\ell$ is a linear functional defined on some subspace, the choice $\varphi(x) = \|\ell\|\|x\|$ implies:

**Corollary 4.12.** Let $X$ be a normed space and let $\ell$ be a bounded linear functional defined on some subspace $Y \subseteq X$. Then there is an extension $\overline{\ell} \in X^*$ preserving the norm.

Moreover, we can now easily prove our anticipated result

**Corollary 4.13.** Let $X$ be a normed space and $x \in X$ fixed. Suppose $\ell(x) = 0$ for all $\ell$ in some total subset $Y \subseteq X^*$. Then $x = 0$.

**Proof.** Clearly, if $\ell(x) = 0$ holds for all $\ell$ in some total subset, this holds for all $\ell \in X^*$. If $x \neq 0$ we can construct a bounded linear functional on span$\{x\}$ by setting $\ell(\alpha x) = \alpha$ and extending it to $X^*$ using the previous corollary. But this contradicts our assumption. \(\square\)

**Example.** Let us return to our example $\ell^\infty(\mathbb{N})$. Let $c(\mathbb{N}) \subset \ell^\infty(\mathbb{N})$ be the subspace of convergent sequences. Set

$$l(x) = \lim_{n \to \infty} x_n, \quad x \in c(\mathbb{N}),$$

(4.9)

then $l$ is bounded since

$$|l(x)| = \lim_{n \to \infty} |x_n| \leq \|x\|_\infty.$$ 

(4.10)
Hence we can extend it to $\ell^\infty(\mathbb{N})$ by Corollary 4.12. Then $l(x)$ cannot be written as $l(x) = l_y(x)$ for some $y \in \ell^1(\mathbb{N})$ (as in (4.6)) since $y_n = l(\delta_n) = 0$ shows $y = 0$ and hence $\ell_y = 0$. The problem is that $\operatorname{span}\{\delta_n\} = c_0(\mathbb{N}) \neq \ell^\infty(\mathbb{N})$, where $c_0(\mathbb{N})$ is the subspace of sequences converging to 0.

Moreover, there is also no other way to identify $\ell^\infty(\mathbb{N})^*$ with $\ell^1(\mathbb{N})$, since $\ell^1(\mathbb{N})$ is separable whereas $\ell^\infty(\mathbb{N})$ is not. This will follow from Lemma 4.17 (iii) below.

Another useful consequence is

Corollary 4.14. Let $Y \subseteq X$ be a subspace of a normed vector space and let $x_0 \in X \setminus \overline{Y}$. Then there exists an $\ell \in X^*$ such that (i) $\ell(y) = 0$, $y \in Y$, (ii) $\ell(x_0) = \operatorname{dist}(x_0,Y)$, and (iii) $\|\ell\| = 1$.

Proof. Replacing $Y$ by $\overline{Y}$ we see that it is no restriction to assume that $Y$ is closed. (Note that $x_0 \in X \setminus \overline{Y}$ if and only if $\operatorname{dist}(x_0,Y) > 0$.) Let $\tilde{Y} = \operatorname{span}\{x_0,Y\}$. Since every element of $\tilde{Y}$ can be uniquely written as $y + \alpha x_0$ we can define

$$\ell(y + \alpha x_0) = \alpha \operatorname{dist}(x_0,Y).$$

By construction $\ell$ is linear on $\tilde{Y}$ and satisfies (i) and (ii). Moreover, by $\operatorname{dist}(x_0,Y) \leq \|x_0 - \frac{y}{\alpha}\|$ for every $y \in Y$ we have

$$|\ell(y + \alpha x_0)| = |\alpha| \operatorname{dist}(x_0,Y) \leq \|y + \alpha x_0\|, \quad y \in Y.$$

Hence $\|\ell\| \leq 1$ and there is an extension to $X^*$ by Corollary 4.12. To see that the norm is in fact equal to one, take a sequence $y_n \in Y$ such that $\operatorname{dist}(x_0,Y) \geq (1 - \frac{1}{n})\|x_0 + y_n\|$. Then

$$|\ell(y_n + x_0)| = \operatorname{dist}(x_0,Y) \geq (1 - \frac{1}{n})\|y_n + x_0\|$$

establishing (iii). \qed

A straightforward consequence of the last corollary is also worthwhile noting:

Corollary 4.15. Let $Y \subseteq X$ be a subspace of a normed vector space. Then $x \in \overline{Y}$ if and only if $\ell(x) = 0$ for every $\ell \in X^*$ which vanishes on $Y$.

If we take the bidual (or double dual) $X^{**}$, then the Hahn–Banach theorem tells us, that $X$ can be identified with a subspace of $X^{**}$. In fact, consider the linear map $J : X \to X^{**}$ defined by $J(x)(\ell) = \ell(x)$ (i.e., $J(x)$ is evaluation at $x$). Then

Theorem 4.16. Let $X$ be a normed space. Then $J : X \to X^{**}$ is isometric (norm preserving).
4.2. The Hahn–Banach theorem and its consequences

Proof. Fix \( x_0 \in X \). By \( |J(x_0)(\ell)| = |\ell(x_0)| \leq \|\ell\|_*\|x_0\| \) we have at least \( \|J(x_0)\|_{**} \leq \|x_0\| \). Next, by Hahn–Banach there is a linear functional \( \ell_0 \) with norm \( \|\ell_0\|_* = 1 \) such that \( \ell_0(x_0) = \|x_0\| \). Hence \( |J(x_0)(\ell_0)| = |\ell_0(x_0)| = \|x_0\| \) shows \( \|J(x_0)\|_{**} = \|x_0\| \).

Example. This gives another quick way of showing that a normed space has a completion: Take \( X = J(X) \subseteq X^{**} \) and recall that a dual space is always complete (Theorem 1.36).

Thus \( J : X \to X^{**} \) is an isometric embedding. In many cases we even have \( J(X) = X^{**} \) and \( X \) is called reflexive in this case.

Example. The Banach spaces \( \ell^p(\mathbb{N}) \) with \( 1 < p < \infty \) are reflexive: Identify \( \ell^p(\mathbb{N})^* \) with \( \ell^q(\mathbb{N}) \) and choose \( z \in \ell^p(\mathbb{N})^* \). Then there is some \( x \in \ell^p(\mathbb{N}) \) such that
\[
z(y) = \sum_{j \in \mathbb{N}} y_j x_j, \quad y \in \ell^q(\mathbb{N}) \cong \ell^p(\mathbb{N})^*.
\]
But this implies \( z(y) = y(x) \), that is, \( z = J(x) \), and thus \( J \) is surjective. (Warning: It does not suffice to just argue \( \ell^p(\mathbb{N})^{**} \cong \ell^p(\mathbb{N})^* \cong \ell^p(\mathbb{N}) \).

However, \( \ell^1 \) is not reflexive since \( \ell^1(\mathbb{N})^* \cong \ell^\infty(\mathbb{N}) \) but \( \ell^\infty(\mathbb{N})^* \not\cong \ell^1(\mathbb{N}) \) as noted earlier.

Example. By the same argument, every Hilbert space is reflexive. In fact, by the Riesz lemma we can identify \( \mathcal{H}^* \) with \( \mathcal{H} \) via the (complex linear) map \( x \mapsto \langle x, \cdot \rangle \). Taking \( z \in \mathcal{H}^* \), we have, again by the Riesz lemma, that \( z(y) = \langle \langle x, \cdot \rangle, y, \cdot \rangle \rangle_{\mathcal{H}^*} = \langle x, y \rangle^* = \langle y, x \rangle = J(x)(y) \).

Lemma 4.17. Let \( X \) be a Banach space.

(i) If \( X \) is reflexive, so is every closed subspace.

(ii) \( X \) is reflexive if and only if \( X^* \) is.

(iii) If \( X^* \) is separable, so is \( X \).

Proof. (i) Let \( Y \) be a closed subspace. Denote by \( j : Y \hookrightarrow X \) the natural inclusion and define \( j_{**, Y} : Y^{**} \to X^{**} \) via \( (j_{**, Y}(y''))(\ell) = y''(\ell|_Y) \) for \( y'' \in Y^{**} \) and \( \ell \in X^* \). Note that \( j_{**, Y} \) is isometric by Corollary 4.12. Then
\[
\begin{array}{ccc}
X & J_X & X^{**} \\
\uparrow & & \uparrow \\
Y & J_Y & Y^{**}
\end{array}
\]
commutes. In fact, we have \( j_{**, Y}(J_Y(y))(\ell) = J_Y(y)(\ell|_Y) = \ell(y) = J_X(y)(\ell) \). Moreover, since \( J_X \) is surjective, for every \( y'' \in Y^{**} \) there is an \( x \in X \) such that \( j_{**, X}(y'') = J_X(x) \). Since \( j_{**, X}(y'') = y''(\ell|_Y) \) vanishes on all \( \ell \in X^* \) which vanish on \( Y \), so does \( \ell(x) = J_X(x)(\ell) = j_{**, X}(y'')(\ell) \) and thus \( x \in Y \).
by Corollary 4.15. That is, \( j_{**}(Y^{**}) = J_X(Y) \) and \( J_Y = j \circ J_X \circ j_{**}^{-1} \) is surjective.

(ii) Suppose \( X \) is reflexive. Then the two maps

\[
(J_X)_* : X^* \rightarrow X^{***} \quad (J_X)^* : X^{***} \rightarrow X^*
\]

are inverse of each other. Moreover, fix \( x'' \in X^{**} \) and let \( x = J_X^{-1}(x'') \).

Then \( J_X^*(x')(x'') = x''(x') = J(x)(x') = x'(J_X^{-1}(x'')) \), that is \( J_X^* = (J_X)_* \), respectively \( (J_X^*)^{-1} = (J_X)^* \), which shows \( X^* \) reflexive if \( X \) reflexive.

To see the converse, observe that \( X^* \) reflexive implies \( X^{**} \) reflexive and hence \( J_X(X) \cong X \) is reflexive by (i).

(iii) Let \( \{\ell_n\}_{n=1}^\infty \) be a dense set in \( X^* \). Then we can choose \( x_n \in X \) such that \( \|x_n\| = 1 \) and \( \ell_n(x_n) \geq \|\ell_n\|/2 \). We will show that \( \{x_n\}_{n=1}^\infty \) is total in \( X \). If it were not, we could find some \( x \in X \setminus \operatorname{span}\{x_n\}_{n=1}^\infty \) and hence there is a functional \( \ell \in X^* \) as in Corollary 4.14. Choose a subsequence \( \ell_{n_k} \rightarrow \ell \).

Then \( \|\ell - \ell_{n_k}\| \geq |(\ell - \ell_{n_k})(x_{n_k})| = |\ell_{n_k}(x_{n_k})| \geq \|\ell_{n_k}\|/2 \), which implies \( \ell_{n_k} \rightarrow 0 \) and contradicts \( \|\ell\| = 1 \).

If \( X \) is reflexive, then the converse of (iii) is also true (since \( X \cong X^{**} \) separable implies \( X^* \) separable), but in general this fails as the example \( \ell^1(\mathbb{N})^* = \ell^\infty(\mathbb{N}) \) shows.

**Problem 4.6.** Let \( X \) be some normed space. Show that

\[
\|x\| = \sup_{\ell \in V, \|\ell\| = 1} |\ell(x)|,
\]  

where \( V \subseteq X^* \) is some dense subspace. Show that equality is attained if \( V = X^* \).

**Problem 4.7.** Let \( X, Y \) be some normed spaces and \( A : \mathcal{D}(A) \subseteq X \rightarrow Y \).

Show

\[
\|A\| = \sup_{x \in X, \|x\| = 1; \ell \in V, \|\ell\| = 1} |\ell(Ax)|,
\]

where \( V \subseteq Y^* \) is a dense subspace.

**Problem 4.8.** Show that \( \|l_y\| = \|y\|_q \), where \( l_y \in \ell^p(\mathbb{N})^* \) as defined in (4.6).

(Hint: Choose \( x \in \ell^p \) such that \( x_n y_n = |y_n|^q \).)

**Problem 4.9.** Show that every \( l \in \ell^p(\mathbb{N})^* \), \( 1 \leq p < \infty \), can be written as

\[
l(x) = \sum_{n \in \mathbb{N}} y_n x_n
\]

with some \( y \in \ell^q(\mathbb{N}) \). (Hint: To see \( y \in \ell^q(\mathbb{N}) \) consider \( x^N \) defined such that \( x^N_n = |y_n|^q/y_n \) for \( n \leq N \) with \( y_n \neq 0 \) and \( x^N_n = 0 \) else. Now look at \( |l(x^N)| \leq \|l\|\|x^N\|_p \).)
Problem 4.10. Let \( c_0(N) \subset \ell^\infty(N) \) be the subspace of sequences which converge to 0, and \( c(N) \subset \ell^\infty(N) \) the subspace of convergent sequences.

(i) Show that \( c_0(N), c(N) \) are both Banach spaces and that \( c(N) = \operatorname{span}\{c_0(N), e\} \), where \( e = (1, 1, 1, \ldots) \in c(N) \).

(ii) Show that every \( l \in c_0(N)^* \) can be written as
\[
l(a) = \sum_{n \in \mathbb{N}} b_n a_n
\]
with some \( b \in \ell^1(N) \) which satisfies \( \|b\|_1 = \|l\| \).

(iii) Show that every \( l \in c(N)^* \) can be written as
\[
l(a) = \sum_{n \in \mathbb{N}} b_n a_n + b_0 \lim_{n \to \infty} a_n
\]
with some \( b \in \ell^1(N) \) which satisfies \( |b_0| + \|b\|_1 = \|l\| \).

Problem 4.11. Let \( u_n \in X \) be a Schauder basis and suppose the complex numbers \( c_n \) satisfy \( |c_n| \leq c \|u_n\| \). Is there a bounded linear functional \( \ell \in X^* \) with \( \ell(u_n) = c_n \)? (Hint: Consider e.g. \( X = \ell^2(\mathbb{Z}) \).)

4.3. The adjoint operator

Given two normed spaces \( X \) and \( Y \) and a bounded operator \( A \in \mathcal{L}(X,Y) \) we can defined its adjoint \( A' : Y^* \to X^* \) via
\[
A' y' = y' \circ A, \quad y' \in Y^*.
\]

It is immediate that \( A' \) is linear and boundedness follows from
\[
\|A'\| = \sup_{y' \in Y^*: \|y'\|=1} \|A'y'\| = \sup_{y' \in Y^*: \|y'\|=1} \left( \sup_{x \in X: \|x\|=1} |(A'y')(x)| \right)
= \sup_{y' \in Y^*: \|y'\|=1} \left( \sup_{x \in X: \|x\|=1} |y'(Ax)| \right)
= \sup_{x \in X: \|x\|=1} \|Ax\| = \|A\|,
\]
where we have used Problem 4.6 to obtain the fourth equality. In summary,

Theorem 4.18. Let \( A \in \mathcal{L}(X,Y) \), then \( A' \in \mathcal{L}(Y^*,X^*) \) with \( \|A\| = \|A'\| \).

Note that for \( A \in \mathcal{L}(X,Y) \) and \( B \in \mathcal{L}(Y,Z) \) we have
\[
(BA)' = A'B'
\]
which is immediate from the definition.

Example. Given a Hilbert space \( \mathcal{H} \) we have the conjugate linear map \( C : \mathcal{H} \to \mathcal{H}^*, \> f \mapsto \langle f, \cdot \rangle \). hence for given \( A \in \mathcal{L}(\mathcal{H}) \) we have \( A'Cf = \langle f, A \cdot \rangle = \langle A^*f, \cdot \rangle \) which shows \( A' = CA^*C^{-1} \). \( \diamond \)
Example. Let \( X = Y = \ell^p(N), 1 \leq p < \infty \) such that \( X^* = \ell^q(N) \), \( \frac{1}{p} + \frac{1}{q} = 1 \). Consider the left shift \( S \in \mathcal{L}(\ell^p(N)) \) given by
\[
Sx = (0, x_1, x_2, \ldots).
\]
Then for \( y' \in \ell^q(N) \)
\[
y'(Sx) = \sum_{j=1}^{\infty} y'_j(Sx)_j = \sum_{j=2}^{\infty} y'_j x_{j-1} = \sum_{j=1}^{\infty} y'_{j+1} x_j
\]
which shows \((S'y')_k = y_{k+1}\) upon choosing \( x = \delta^k \). Hence \( S' \) is the right shift. \( \diamond \)

Of course we can also consider the doubly adjoint operator \( A'' \). Then a simple computation
\[
A''(J_X(x))(y') = J_X(x)(A'y') = (A'y')(x) = y'(Ax) = J_Y(Ax)(y') \tag{4.14}
\]
shows that the following diagram commutes
\[
\begin{array}{ccc}
X & \xrightarrow{A} & Y \\
\downarrow J_X & & \downarrow J_Y \\
X^{**} & \xrightarrow{A''} & Y^{**}
\end{array}
\]
Hence, regarding \( X \) as a subspace \( J(X) \subseteq X^{**} \) then \( A'' \) is an extension of \( A \) to \( X^{**} \) but with values in \( Y^{**} \). In particular, note that \( B \in \mathcal{L}(Y^*, X^*) \) is the adjoint of some other operator \( B = A' \) if and only if \( B'(J_X(X)) = A''(J_X(X)) \subseteq J_Y(Y) \). This can be used to show that not every operator is an adjoint (Problem 4.12).

Theorem 4.19 (Schauder). Suppose \( X, Y \) are Banach spaces and \( A \in \mathcal{L}(X,Y) \). Then \( A \) is compact if and only if \( A' \) is.

Proof. If \( A \) is compact, then \( A(B_1^X(0)) \) is relatively compact and hence \( K = \overline{A(B_1^X(0))} \) is a compact metric space. Let \( y'_n \in Y^* \) be a bounded sequence and consider the family of functions \( f_n = y'_n|_K \in C(K) \). Then this family is bounded and equicontinuous since
\[
|f_n(y_1) - f_n(y_2)| \leq ||y'_n|| ||y_1 - y_2|| \leq C ||y_1 - y_2||.
\]
Hence the Arzelà–Ascoli theorem (Theorem 3.4) implies existence of a uniformly converging subsequence \( f_{n_j} \). For this subsequence we have
\[
||A'y'_{n_j} - A'y'_{n_k}|| \leq \sup_{x \in B_1^X(0)} |y'_{n_j}(Ax) - y'_{n_k}(Ax)| = ||f_{n_j} - f_{n_k}||_{\infty}
\]
since \( A(B_1^X(0)) \subseteq K \) is dense. Thus \( y'_{n_j} \) is the required subsequence and \( A' \) is compact.

To see the converse assume that \( A' \) is compact and hence also \( A'' \) by the first part. It remains compact when we restrict it to \( J_X(X) \) and since the
image of this restriction is contained in the closed set \( J_Y(Y) \), the operator
\[ A'' : J_X(X) \to J_Y(Y) \]
is compact. But this restriction is isometrically isomorphic to \( K \).

Finally we discuss the relation between solvability of \( Ax = y \) and the corresponding adjoint equation \( A'y' = x' \). To this end we need the analog of the orthogonal complement of a set. Given subsets \( M \subseteq X \) and \( N \subseteq X^* \) we define their annihilator as
\[
M^\perp = \{ \ell \in X^* | \ell(x) = 0 \ \forall x \in M \} = \{ \ell \in X^* | M \subseteq \ker(\ell) \},
\]
\[
N_\perp = \{ x \in X | \ell(x) = 0 \ \forall \ell \in N \} = \bigcap_{\ell \in N} \ker(\ell).
\]

The following properties are immediate from the definition (by linearity and continuity)

- \( M^\perp \) is a closed subspace of \( X^* \) and \( M^\perp = (\overline{\text{span}(M)})^\perp \).
- \( N_\perp \) is a closed subspace of \( X \) and \( N_\perp = (\overline{\text{span}(N)})_\perp \).

Note also that \( \overline{\text{span}(M)} = X^* \) if and only if \( M^\perp = \{0\} \) (cf. Corollary 4.13) and \( \overline{\text{span}(N)} = X \) if and only if \( N_\perp = \{0\} \).

**Lemma 4.20.** We have \( (M^\perp)_\perp = \overline{\text{span}(M)} \) and \( (N_\perp)^\perp \subseteq \overline{\text{span}(N)} \).

**Proof.** Be the preceding remarks we can assume \( M, N \) to be closed subspaces. The first part is Corollary 4.15 and for the second part one just has to spell out the definition:
\[
(N_\perp)^\perp = \{ \ell \in X^* | \bigcap_{\tilde{\ell} \in N} \ker(\tilde{\ell}) \subseteq \ker(\ell) \} \tag{4.15}
\]

Note that we have equality in the preceding lemma if \( N \) is finite dimensional (Problem 4.14).

Furthermore, we have the following analog of (2.28).

**Theorem 4.21.** Suppose \( X, Y \) are normed spaces and \( A \in \mathcal{L}(X,Y) \). Then
\( \text{Ran}(A')_\perp = \ker(A) \) and \( \text{Ran}(A)^\perp = \ker(A') \).

**Proof.** For the first claim observe: \( x \in \ker(A) \Leftrightarrow Ax = 0 \Leftrightarrow \ell(Ax) = 0, \forall \ell \in X^* \Leftrightarrow (A'\ell)(x) = 0, \forall \ell \in X^* \Leftrightarrow x \in \text{Ran}(A')^\perp \).

For the second claim observe: \( \ell \in \ker(A') \Leftrightarrow A'\ell = 0 \Leftrightarrow (A'\ell)(x) = 0, \forall x \in X \Leftrightarrow \ell(Ax) = 0, \forall x \in X \Leftrightarrow \ell \in \text{Ran}(A)^\perp \).
Theorem 4.22. Let $M$ be a closed subspace of a normed space. Then there are canonical isometries

$$(X/M)^* \cong M^\perp, \quad M^* \cong X^*/M^\perp.$$  \hspace{1cm} (4.16)

**Proof.** In the first case the isometry is given by $\ell \mapsto \ell \circ j$, where $j : X \to X/M$ is the quotient map. In the second case $x' + M^\perp \mapsto x'|_M$. The details are easy to check. \quad \square

**Problem 4.12.** Let $X = Y = c_0(\mathbb{N})$ and recall that $X^* = \ell^1(\mathbb{N})$ and $X^{**} = \ell^\infty(\mathbb{N})$. Consider the operator $A \in \mathcal{L}(\ell^1(\mathbb{N}))$ given by

$$Ax = \left( \sum_{n \in \mathbb{N}} x_n, 0, \ldots \right).$$

Show that $A'x' = (x'_1, x'_2, \ldots)$. Conclude that $A$ is not the adjoint of an operator from $\mathcal{L}(C_0(\mathbb{N}))$.

**Problem 4.13.** Let $X$ be a normed vector space and $Y \subset X$ some subspace. Show that if $Y \neq X$, then for every $\varepsilon \in (0, 1)$ there exists an $x_\varepsilon$ with $\|x_\varepsilon\| = 1$ and

$$\inf_{y \in Y} \|x_\varepsilon - y\| \geq 1 - \varepsilon.$$ \hspace{1cm} (4.17)

Note: In a Hilbert space the lemma holds with $\varepsilon = 0$ for any normalized $x$ in the orthogonal complement of $Y$ and hence $x_\varepsilon$ can be thought of a replacement of an orthogonal vector. (Hint: Choose an $y_\varepsilon \in Y$ which is close to $x$ and look at $x - y_\varepsilon$.)

**Problem 4.14.** Suppose $X$ is a vector space and $\ell, \ell_1, \ldots, \ell_n$ are linear functionals such that $\bigcap_{j=1}^n \ker(\ell_j) \subseteq \ker(\ell)$. Then $\ell = \sum_{j=0}^n \alpha_j \ell_j$ for some constants $\alpha_j \in \mathbb{C}$. (Hint: Choose a dual basis $x_k \in X$ such that $\ell_j(x_k) = \delta_{j,k}$ and look at $x - \sum_{j=1}^n \ell_j(x)x_j$.)

### 4.4. The geometric Hahn–Banach theorem

Finally we turn to a geometric version of the Hahn–Banach theorem. Let $X$ be a vector space. For every subset $U \subset X$ we define its **Minkowski functional** (or gauge)

$$p_U(x) = \inf \{ t > 0 | x \in tU \}.$$  \hspace{1cm} (4.18)

Here $tU = \{ tx | x \in U \}$. Note that $0 \in U$ implies $p_U(0) = 0$ and $p_U(x)$ will be finite for all $x$ when $U$ is **absorbing**, that is, for every $x \in X$ there is some $r$ such that $x \in \alpha U$ for every $|\alpha| \geq r$. Note that every absorbing set contains 0 and every neighborhood of 0 in a Banach space is absorbing.
4.4. The geometric Hahn–Banach theorem

**Figure 1.** Separation of convex sets via a hyperplane

**Example.** Let $X$ be a Banach space and $U = B_1(0)$, then $p_U(x) = \|x\|$. If $X = \mathbb{R}^2$ and $U = (-1, 1) \times \mathbb{R}$ then $p_U(x) = |x_1|$. If $X = \mathbb{R}^2$ and $U = (-1, 1) \times \{0\}$ then $p_U(x) = |x_1|$ if $x_2 = 0$ and $p_U(x) = \infty$ else.

**Lemma 4.23.** Let $X$ be a vector space and $U$ a convex subset containing $0$. Then

$$p_U(x + y) \leq p_U(x) + p_U(y), \quad p_U(\lambda x) = \lambda p_U(x), \quad \lambda \geq 0. \quad (4.19)$$

Moreover, $\{x \mid p_U(x) < 1\} \subseteq U \subseteq \{x \mid p_U(x) \leq 1\}$. If, in addition, $X$ is a topological vector space and $U$ is open, then $U = \{x \mid p_U(x) < 1\}$.

**Proof.** The homogeneity condition $p(\lambda x) = \lambda p(x)$ for $\lambda > 0$ is straightforward. To see the sublinearity Let $t, s > 0$ with $x \in tU$ and $y \in sU$, then

$$t \cdot \frac{x}{t} + s \cdot \frac{y}{s} = \frac{t}{t + s} x + \frac{s}{t + s} y$$

is in $U$ by convexity. Moreover, $p_U(x + y) \leq s + t$ and taking the infimum over all $t$ and $s$ we find $p_U(x + y) \leq p_U(x) + p_U(y)$.

Suppose $p_U(x) < 1$, then $t^{-1}x \in U$ for some $t < 1$ and thus $x \in U$ by convexity. Similarly, if $x \in U$ then $t^{-1}x \in U$ for $t \geq 1$ by convexity and thus $p_U(x) \leq 1$. Finally, let $U$ be open and $x \in U$, then $(1 + \varepsilon)x \in U$ for some $\varepsilon > 0$ and thus $p(x) \leq (1 + \varepsilon)^{-1}$. \qed

Note that (4.19) implies convexity

$$p_U(\lambda x + (1 - \lambda)y) \leq \lambda p_U(x) + (1 - \lambda)p_U(y), \quad \lambda \in [0, 1]. \quad (4.20)$$

**Theorem 4.24** (geometric Hahn–Banach, real version). Let $U, V$ be disjoint nonempty convex subsets of a real topological vector space $X$ space and let $U$ be open. Then there is a linear functional $\ell \in X^*$ and some $c \in \mathbb{R}$ such that

$$\ell(x) < c \leq \ell(y), \quad x \in U, \ y \in V. \quad (4.21)$$

If $V$ is also open, then the second inequality is also strict.
Proof. Choose \( x_0 \in U \) and \( y_0 \in V \), then
\[
W = (U - x_0) + (V - y_0) = \{(x - x_0) - (y - y_0) | x \in U, y \in V \}
\]
is open (since \( U \) is), convex (since \( U \) and \( V \) are) and contains 0. Moreover, since \( U \) and \( V \) are disjoint we have \( z_0 = y_0 - x_0 \notin W \). By the previous lemma, the associated Minkowski functional \( p_W \) is convex and by the Hahn–Banach theorem there is a linear functional satisfying
\[
\ell(tz_0) = t, \quad |\ell(x)| \leq p_W(x).
\]
Note that since \( z_0 \notin W \) we have \( p_W(z_0) \geq 1 \). Moreover, \( W = \{x|p_U(x) < 1\} \subseteq \{x|p(x) < 1\} \) which shows that \( \ell \) is continuous at 0 by scaling and by translations \( \ell \) is continuous everywhere.

Finally we again use \( p_W(z) < 1 \) for \( z \in W \) implying
\[
\ell(x) - \ell(y) + 1 = \ell(x - y + z_0) \leq p_W(x - y + z_0) < 1
\]
and hence \( \ell(x) < \ell(y) \) for \( x \in U \) and \( y \in V \). Therefore \( \ell(U) \) and \( \ell(V) \) are disjoint convex subsets of \( \mathbb{R} \). Finally, let us suppose that there is some \( x_1 \) for which \( \ell(x_1) = \sup \ell(U) \). Then, by continuity of the map \( t \mapsto x_1 + tz_0 \) there is some \( \varepsilon > 0 \) such that \( x_1 + \varepsilon z_0 \in U \). But this gives a contradiction \( \ell(x_1) + \varepsilon = \ell(x_1 + \varepsilon z_0) \leq \ell(x_1) \). Thus the claim holds with \( c = \sup \ell(U) \). If \( V \) is also open an analogous argument shows \( \inf \ell(V) < \ell(y) \) for all \( y \in V \). \( \square \)

Of course there is also a complex version.

**Theorem 4.25** (geometric Hahn–Banach, complex version). Let \( U, V \) be disjoint nonempty convex subsets of a topological vector space \( X \) and let \( U \) be open. Then there is a linear functional \( \ell \in X^* \) and some \( c \in \mathbb{R} \) such that
\[
\text{Re}(\ell(x)) < c \leq \text{Re}(\ell(y)), \quad x \in U, y \in V. \tag{4.22}
\]
If \( V \) is also open, then the second inequality is also strict.

**Proof.** Consider \( X \) as a real Banach space. Then there is a continuous real-linear functional \( \ell_r : X \to \mathbb{R} \) by the real version of the geometric Hahn–Banach theorem. Then \( \ell(x) = \ell_r(x) - i\ell_r(ix) \) is the functional we are looking for (check this). \( \square \)

**Example.** The assumption that one set is open is crucial as the following example shows. Let \( X = c_0(\mathbb{N}) \), \( U = \{a \in c_0(\mathbb{N}) \exists N : a_N > 0 \text{ and } a_n = 0, n > N \} \) and \( V = \{0\} \). Note that \( U \) is convex but not open and that \( U \cap V = \emptyset \). Suppose we could find a linear functional \( \ell \) as in the geometric Hahn–Banach theorem (of course we can choose \( \alpha = \ell(0) = 0 \) in this case). Then by Problem 4.10 there is some \( b_j \in \ell^\infty(\mathbb{N}) \) such that \( \ell(a) = \sum_{j=1}^\infty b_j a_j \). Moreover, we must have \( b_j = \ell(\delta^j) < 0 \). But then \( a = (b_2, -b_1, 0, \ldots) \in U \) and \( \ell(a) = 0 \neq 0 \). \( \diamond \)
4.5. Weak convergence

Note that two disjoint closed convex sets can be separated strictly if one of them is compact. However, this will require that every point has a neighborhood base of convex open sets. Such topological vector spaces are called locally convex spaces.

**Corollary 4.26.** Let $U$, $V$ be disjoint nonempty closed convex subsets of a locally convex space $X$ and let $U$ be compact. Then there is a linear functional $\ell \in X^*$ and some $c, d \in \mathbb{R}$ such that

$$\text{Re}(\ell(x)) \leq d < c \leq \text{Re}(\ell(y)), \quad x \in U, \ y \in V.$$  \hspace{1cm} (4.23)

**Proof.** Since $V$ is closed for every $x \in U$ there is a convex open neighborhood $N_x$ of 0 such that $x + N_x$ does not intersect $V$. By compactness of $U$ there are $x_1, \ldots, x_n$ such that the corresponding neighborhoods $x_j + \frac{1}{2} N_{x_j}$ cover $U$. Set $N = \bigcap_{j=1}^n N_{x_j}$ which as a convex open neighborhood of 0.

$$\tilde{U} = U + \frac{1}{2} N \subseteq \bigcup_{j=1}^n (x_j + \frac{1}{2} N_{x_j}) + \frac{1}{2} N \subseteq \bigcup_{j=1}^n (x_j + \frac{1}{2} N_{x_j} + \frac{1}{2} N_{x_j}) = \bigcup_{j=1}^n (x_j + N_{x_j})$$

is a convex open set which is disjoint from $V$. Hence by the previous theorem we can find some $\ell$ such that $\text{Re}(\ell(x)) < c \leq \text{Re}(\ell(y))$ for all $x \in \tilde{U}$ and $y \in V$. Moreover, since $\ell(U)$ is a compact interval $[e, d]$ the claim follows. \hfill $\square$

Note that if $U$ and $V$ are absolutely convex (i.e., $\alpha U + \beta U \subseteq U$ for $|\alpha| + |\beta| \leq 1$), then we can write the previous condition equivalently as

$$|\ell(x)| \leq d < c \leq |\ell(y)|, \quad x \in U, \ y \in V.$$  \hspace{1cm} (4.24)

since $x \in U$ implies $\theta x \in U$ for $\theta = \text{sign}(\ell(x))$ and thus $|\ell(x)| = \theta \ell(x) = \ell(\theta x) = \text{Re}(\ell(\theta x))$.

**Problem 4.15.** Show that Corollary 4.26 fails even in $\mathbb{R}^2$ unless one set is compact.

4.5. Weak convergence

In Section 4.2 we have seen that $\ell(x) = 0$ for all $\ell \in X^*$ implies $x = 0$. Now what about convergence? Does $\ell(x_n) \to \ell(x)$ for every $\ell \in X^*$ imply $x_n \to x$? Unfortunately the answer is no:

**Example.** Let $u_n$ be an infinite orthonormal set in some Hilbert space. Then $\langle g, u_n \rangle \to 0$ for every $g$ since these are just the expansion coefficients of $g$ which are in $\ell^2(\mathbb{N})$ by Bessel’s inequality. Since by the Riesz lemma (Theorem 2.10), every bounded linear functional is of this form, we have $\ell(u_n) \to 0$ for every bounded linear functional. (Clearly $u_n$ does not converge to 0, since $\|u_n\| = 1$.)

$\diamond$
4. The main theorems about Banach spaces

If \( \ell(x_n) \to \ell(x) \) for every \( \ell \in X^* \) we say that \( x_n \) **converges weakly** to \( x \) and write

\[
\text{w-lim}_{n \to \infty} x_n = x \quad \text{or} \quad x_n \rightharpoonup x.
\]  

(4.25)

Clearly, \( x_n \to x \) implies \( x_n \rightharpoonup x \) and hence this notion of convergence is indeed weaker. Moreover, the weak limit is unique, since \( \ell(x_n) \to \ell(x) \) and \( \ell(x_n) \to \ell(\tilde{x}) \) imply \( \ell(x - \tilde{x}) = 0 \). A sequence \( x_n \) is called a **weak Cauchy sequence** if \( \ell(x_n) \) is Cauchy (i.e. converges) for every \( \ell \in X^* \).

**Lemma 4.27.** Let \( X \) be a Banach space.

(i) \( x_n \rightharpoonup x \) implies \( \|x\| \leq \lim \inf \|x_n\| \).

(ii) Every weak Cauchy sequence \( x_n \) is bounded: \( \|x_n\| \leq C \).

(iii) If \( X \) is reflexive, then every weak Cauchy sequence converges weakly.

(iv) A sequence \( x_n \) is Cauchy if and only if \( \ell(x_n) \) is Cauchy, uniformly for \( \ell \in X^* \) with \( \|\ell\| = 1 \).

**Proof.**

(i) Choose \( \ell \in X^* \) such that \( \ell(x) = \|x\| \) (for the limit \( x \)) and \( \|\ell\| = 1 \). Then

\[
\|x\| = \ell(x) = \lim \inf \ell(x_n) \leq \lim \inf \|x_n\|.
\]

(ii) For every \( \ell \) we have that \( |J(x_n)(\ell)| = |\ell(x_n)| \leq C(\ell) \) is bounded. Hence by the uniform boundedness principle we have \( \|x_n\| = \|J(x_n)\| \leq C \).

(iii) If \( x_n \) is a weak Cauchy sequence, then \( \ell(x_n) \) converges and we can define \( j(\ell) = \lim \ell(x_n) \). By construction \( j \) is a linear functional on \( X^* \). Moreover, by (ii) we have \( \|j(\ell)\| \leq \sup \|\ell(x_n)\| \leq \|\ell\| \sup \|x_n\| \leq C\|\ell\| \) which shows \( j \in X^{**} \). Since \( X \) is reflexive, \( j = J(x) \) for some \( x \in X \) and by construction \( \ell(x_n) \to J(x)(\ell) = \ell(x) \), that is, \( x_n \rightharpoonup x \).

(iv) This follows from

\[
\|x_n - x_m\| = \sup_{\|\ell\| = 1} |\ell(x_n - x_m)|
\]

(cf. Problem 4.6). \( \square \)

**Remark:** One can equip \( X \) with the weakest topology for which all \( \ell \in X^* \) remain continuous. This topology is called the **weak topology** and it is given by taking all finite intersections of inverse images of open sets as a base. By construction, a sequence will converge in the weak topology if and only if it converges weakly. By Corollary 4.14 the weak topology is Hausdorff, but it will not be metrizable in general. In particular, sequences do not suffice to describe this topology. Nevertheless we will stick with sequences for now and come back to this more general point of view in the next section.

In a Hilbert space there is also a simple criterion for a weakly convergent sequence to converge in norm.
Lemma 4.28. Let \( H \) be a Hilbert space and let \( f_n \to f \). Then \( f_n \to f \) if and only if \( \limsup \|f_n\| \leq \|f\| \).

Proof. By (i) of the previous lemma we have \( \lim \|f_n\| = \|f\| \) and hence
\[
\|f - f_n\|^2 = \|f\|^2 - 2\text{Re}(\langle f, f_n \rangle) + \|f_n\|^2 \to 0.
\]
The converse is straightforward.

Now we come to the main reason why weakly convergent sequences are of interest: A typical approach for solving a given equation in a Banach space is as follows:

(i) Construct a (bounded) sequence \( x_n \) of approximating solutions (e.g. by solving the equation restricted to a finite dimensional subspace and increasing this subspace).

(ii) Use a compactness argument to extract a convergent subsequence.

(iii) Show that the limit solves the equation.

Our aim here is to provide some results for the step (ii). In a finite dimensional vector space the most important compactness criterion is boundedness (Heine–Borel theorem, Theorem 1.17). In infinite dimensions this breaks down:

Theorem 4.29. The closed unit ball in \( X \) is compact if and only if \( X \) is finite dimensional.

Proof. If \( X \) is finite dimensional, then \( X \) is isomorphic to \( \mathbb{C}^n \) and the closed unit ball is compact by the Heine–Borel theorem (Theorem 1.17).

Conversely, suppose \( S = \{ x \in X | \|x\| = 1 \} \) is compact. Then the open cover \( \{ X \setminus \text{Ker}(\ell) \}_{\ell \in X^*} \) has a finite subcover, \( S \subset \bigcup_{j=1}^n X \setminus \text{Ker}(\ell_j) = X \setminus \bigcap_{j=1}^n \text{Ker}(\ell_j) \). Hence \( \bigcap_{j=1}^n \text{Ker}(\ell_j) = \{0\} \) and the map \( X \to \mathbb{C}^n, \; x \mapsto (\ell_1(x), \ldots, \ell_n(x)) \) is injective, that is, \( \dim(X) \leq n \).

If we are willing to treat convergence for weak convergence, the situation looks much brighter!

Theorem 4.30. Let \( X \) be a reflexive Banach space. Then every bounded sequence has a weakly convergent subsequence.

Proof. Let \( x_n \) be some bounded sequence and consider \( Y = \overline{\text{span}\{x_n\}} \). Then \( Y \) is reflexive by Lemma 4.17 (i). Moreover, by construction \( Y \) is separable and so is \( Y^* \) by the remark after Lemma 4.17.

Let \( \ell_k \) be a dense set in \( Y^* \). Then by the usual diagonal sequence argument we can find a subsequence \( x_{nm} \) such that \( \ell_k(x_{nm}) \) converges for
The main theorems about Banach spaces

every \( k \). Denote this subsequence again by \( x_n \) for notational simplicity. Then,
\[
\| \ell(x_n) - \ell(x_m) \| \leq \| \ell(x_n) - \ell_k(x_n) \| + \| \ell_k(x_n) - \ell_k(x_m) \| \\
+ \| \ell_k(x_m) - \ell(x_m) \| \\
\leq 2C \| \ell - \ell_k \| + \| \ell_k(x_n) - \ell_k(x_m) \|
\]
shows that \( \ell(x_n) \) converges for every \( \ell \in \text{span}\{\ell_k\} = Y^* \). Thus there is a limit by Lemma 4.27 (iii). \( \square \)

Note that this theorem breaks down if \( X \) is not reflexive.

**Example.** Consider the sequence of vectors \( \delta^n \) (with \( \delta^n \equiv 1 \) and \( \delta^n_m = 0, \ n \neq m \)) in \( \ell^p(\mathbb{N}), \ 1 \leq p < \infty \). Then \( \delta^n \rightharpoonup 0 \) for \( 1 < p < \infty \). In fact, since every \( l \in \ell^p(\mathbb{N})^* \) is of the form \( l = ly \) for some \( y \in \ell^q(\mathbb{N}) \) we have \( ly(\delta^n) = yn \to 0 \).

If we consider the same sequence in \( \ell^1(\mathbb{N}) \) there is no weakly convergent subsequence. In fact, since \( ly(\delta^n) \to 0 \) for every sequence \( y \in \ell^\infty(\mathbb{N}) \) with finitely many nonzero entries, the only possible weak limit is zero. On the other hand choosing the constant sequence \( y = (1)_{j=1}^{\infty} \) we see \( ly(\delta^n) = 1 \nrightarrow 0 \), a contradiction. \( \diamond \)

**Example.** Let \( X = L^1(\mathbb{R}) \). Every bounded \( \varphi \) gives rise to a linear functional
\[
\ell_\varphi(f) = \int f(x) \varphi(x) \, dx
\]
in \( L^1(\mathbb{R})^* \). Take some nonnegative \( u_1 \) with compact support, \( \|u_1\|_1 = 1 \), and set \( u_k(x) = ku_1(kx) \). Then we have
\[
\int u_k(x) \varphi(x) \, dx \to \varphi(0)
\]
(see Problem 8.13) for every continuous \( \varphi \). Furthermore, if \( u_{k_j} \rightharpoonup u \) we conclude
\[
\int u(x) \varphi(x) \, dx = \varphi(0).
\]
In particular, choosing \( \varphi_k(x) = \max(0, 1-k|x|) \) we infer from the dominated convergence theorem
\[
1 = \int u(x) \varphi_k(x) \, dx \to \int u(x) \chi_{\{0\}}(x) \, dx = 0,
\]
a contradiction.

In fact, \( u_k \) converges to the Dirac measure centered at 0, which is not in \( L^1(\mathbb{R}) \). \( \diamond \)
4.5. Weak convergence

Note that the above theorem also shows that in an infinite dimensional reflexive Banach space weak convergence is always weaker than strong convergence since otherwise every bounded sequence had a weakly, and thus by assumption also norm, convergent subsequence contradicting Theorem 4.29. In a non-reflexive space this situation can however occur.

**Example.** In $\ell^1(\mathbb{N})$ every weakly convergent sequence is in fact (norm) convergent. If this were not the case we could find a sequence $a^n \rightarrow 0$ for which $\lim \inf_{n} ||a^n||_1 \geq \varepsilon > 0$. After passing to a subsequence we can assume $||a^n||_1 \geq \varepsilon/2$ and after rescaling the norm even $||a^n||_1 = 1$. Now weak convergence $a^n \rightarrow 0$ implies $a^n_j = l_\delta_j(a^n) \rightarrow 0$ for every fixed $j \in \mathbb{N}$. Hence the main contribution to the norm of $a^n$ must move towards $\infty$ and we can find a subsequence $n_j$ and a corresponding increasing sequence of integers $k_j$ such that $\sum_{k_j \leq k < k_{j+1}} |a^n_k| \geq \frac{2}{3}$. Now set $b_k = \text{sign}(a^n_k)$, $k_j \leq k < k_{j+1}$.

Then $|l_b(a^{n_j})| = \sum_{k_j \leq k < k_{j+1}} |a^n_k| + \sum_{1 \leq k < k_j; k_{j+1} \leq k} b_k a^n_k \geq \frac{2}{3} - \frac{1}{3} = \frac{1}{3}$, contradicting $a^{n_j} \rightarrow 0$.

Let me remark that similar concepts can be introduced for operators. This is of particular importance for the case of unbounded operators, where convergence in the operator norm makes no sense at all.

A sequence of operators $A_n$ is said to **converge strongly** to $A$,

$$\text{s-lim }_n A_n = A \quad :\iff \quad A_n x \rightarrow Ax \quad \forall x \in \mathcal{D}(A) \subseteq \mathcal{D}(A_n).$$

(4.26)

It is said to **converge weakly** to $A$,

$$\text{w-lim }_n A_n = A \quad :\iff \quad A_n x \rightharpoonup Ax \quad \forall x \in \mathcal{D}(A) \subseteq \mathcal{D}(A_n).$$

(4.27)

Clearly norm convergence implies strong convergence and strong convergence implies weak convergence.

**Example.** Consider the operator $S_n \in \mathcal{L}(\ell^2(\mathbb{N}))$ which shifts a sequence $n$ places to the left, that is,

$$S_n (x_1, x_2, \ldots) = (x_{n+1}, x_{n+2}, \ldots)$$

(4.28)

and the operator $S^*_n \in \mathcal{L}(\ell^2(\mathbb{N}))$ which shifts a sequence $n$ places to the right and fills up the first $n$ places with zeros, that is,

$$S^*_n (x_1, x_2, \ldots) = (0, \ldots, 0, x_1, x_2, \ldots)_{n \text{ places}}.$$  

(4.29)
Then $S_n$ converges to zero strongly but not in norm (since $\|S_n\| = 1$) and $S_n^*$ converges weakly to zero (since $\langle x, S_n^* y \rangle = \langle S_n x, y \rangle$) but not strongly (since $\|S_n^* x\| = \|x\|$).

Lemma 4.31. Suppose $A_n \in \mathcal{L}(X,Y)$ is a sequence of bounded operators.

(i) $s\lim_{n \to \infty} A_n = A$ implies $\|A\| \leq \liminf_{n \to \infty} \|A_n\|$.

(ii) Every strong Cauchy sequence $A_n$ is bounded: $\|A_n\| \leq C$.

(iii) If $A_n y \to Ay$ for $y$ in a total set and $\|A_n\| \leq C$, then $s\lim_{n \to \infty} A_n = A$.

The same result holds if strong convergence is replaced by weak convergence.

Proof. (i) follows from

$$\|Ax\| = \lim_{n \to \infty} \|A_n x\| \leq \liminf_{n \to \infty} \|A_n\|$$

for every $x$ with $\|x\| = 1$.

(ii) follows as in Lemma 4.27 (ii).

(iii) By taking linear combinations we can replace the total set by a dense one. Now just use

$$\|A_n x - Ax\| \leq \|A_n x - A_n y\| + \|A_n y - Ay\| + \|Ay - Ax\|$$

$$\leq 2C\|x - y\| + \|A_n y - Ay\|$$

and choose $y$ in the dense subspace such that $\|x - y\| \leq \frac{\varepsilon}{4C}$ and $n$ large such that $\|A_n y - Ay\| \leq \frac{\varepsilon}{2}$.

The case of weak convergence is left as an exercise.

Lemma 4.32. Suppose $A_n \in \mathcal{L}(Y,Z)$, $A_n, B_n \in \mathcal{L}(X,Y)$ are two sequences of bounded operators.

(i) $s\lim_{n \to \infty} A_n = A$ and $s\lim_{n \to \infty} B_n = B$ implies $s\lim_{n \to \infty} A_n B_n = AB$.

(ii) $w\lim_{n \to \infty} A_n = A$ and $s\lim_{n \to \infty} B_n = B$ implies $w\lim_{n \to \infty} A_n B_n = AB$.

(iii) $\lim_{n \to \infty} A_n = A$ and $w\lim_{n \to \infty} B_n = B$ implies $w\lim_{n \to \infty} A_n B_n = AB$.

Proof. For the first case just observe

$$\|(A_n B_n - AB)x\| \leq \|(A_n - A)Bx\| + \|A_n\|\|(B_n - B)x\| \to 0.$$
converges to 0 weakly (in fact even strongly) but 

\[ S_n S_n^*(x_1, x_2, \ldots) = (x_1, x_2, \ldots) \]

does not! Hence the order in the second claim is important. ⋄

For a sequence of linear functionals \( \ell_n \), strong convergence is also called weak-∗ convergence. That is, the weak-∗ limit of \( \ell_n \) is \( \ell \) if 

\[ \ell_n(x) \to \ell(x) \quad \forall x \in X. \] (4.30)

Note that this is the same as strong convergence on \( X^* = \mathcal{L}(X, \mathbb{C}) \) but not the same as weak convergence on \( X^* \), since \( \ell \) is the weak limit of \( \ell_n \) if 

\[ j(\ell_n) \to j(\ell) \quad \forall j \in X^{**}, \] (4.31)

whereas for the weak-∗ limit this is only required for \( j \in J(X) \subseteq X^{**} \) (recall \( J(x)(\ell) = \ell(x) \)).

With this notation it is also possible to slightly generalize Theorem 4.30 (Problem 4.19):

**Lemma 4.33.** Suppose \( X \) is a separable Banach space. Then every bounded sequence \( \ell_n \in X^* \) has a weak-∗ convergent subsequence.

**Example.** Let us return to the example after Theorem 4.30. Consider the Banach space of bounded continuous functions \( X = C(\mathbb{R}) \). Using \( \ell_f(\varphi) = \int \varphi f \, dx \) we can regard \( L^1(\mathbb{R}) \) as a subspace of \( X^* \). Then the Dirac measure centered at 0 is also in \( X^* \) and it is the weak-∗ limit of the sequence \( u_k \). ⋄

Finally, let me discuss a simple application of the above ideas to the **calculus of variations**. Many problems lead to finding the minimum of a given function. For example, many physical problems can be described by an energy functional and one seeks a solution which minimizes this energy. So we have a Banach space \( X \) (typically some function space) and a functional \( F : M \subseteq X \to \mathbb{R} \) (of course this functional will in general be nonlinear). If \( M \) is compact and \( F \) is continuous, then we can proceed as in the finite-dimensional case to show that there is a minimizer: Start with a sequence \( x_n \) such that \( F(x_n) \to \inf_M F \). By compactness we can assume that \( x_n \to x_0 \) after passing to a subsequence and by continuity \( F(x_n) \to F(x_0) = \inf_M F \). Now in the infinite dimensional case we will use weak convergence to get compactness and hence we will also need weak (sequential) continuity of \( F \). However, since there are more weak than strong convergent subsequences, weak (sequential) continuity is in fact a stronger property than just continuity! Hence it might be difficult to check in concrete applications. In this respect note that for our argument to work lower semicontinuity (cf. Problem 7.11) will already be sufficient:
Theorem 4.34. Let $X$ be a reflexive Banach space and let $F : M \subseteq X \to \mathbb{R}$. Suppose that either $M = X$ and $F$ is weakly coercive, that is $F(x) \to \infty$ whenever $\|x\| \to \infty$ or that $M$ is weakly sequentially closed and bounded. If in addition, $F$ is weakly sequentially lower semicontinuous, then there exists some $x_0 \in M$ with $F(x_0) = \inf_M F$.

Proof. As above we start with a sequence $x_n \in M$ such that $F(x_n) \to \inf_M F$. If $M = X$ then $F$ coercive implies that $x_n$ is bounded. Otherwise, it is bounded since we assumed $M$ to be bounded. Hence we can pass to a subsequence such that $x_n \rightharpoonup x_0$ with $x_0 \in M$ since $M$ is assumed sequentially closed. Now since $F$ is weakly sequentially lower semicontinuous we finally get $\inf_M F = \lim_{n \to \infty} F(x_n) = \liminf_{n \to \infty} F(x_n) \geq F(x_0)$. □

Of course in a metric space the definitions closedness in terms of sequences agree with the corresponding topological definition. However, for arbitrary topological spaces it will be weaker in general. So in general we have the following connection: weakly closed implies sequentially weakly closed (as the complement is open, a sequence from the set cannot converge to a point in the complement) which in turn implies sequentially closed (since every weakly convergent sequence is convergent) which in turn implies closed (since for every limit point there is a sequence converging to it). If convexity is added to the requirements it turns out that they all agree!

Lemma 4.35. Suppose $M \subseteq X$ is convex. Then every closed set is sequentially weakly closed.

Proof. Suppose $x$ is in the weak sequential closure of $M$, that is, there is a sequence $x_n \rightharpoonup x$. If $x \not\in M$, then by Corollary 4.26 we can find a linear functional $\ell$ which separates $\{x\}$ and $M$. But this contradicts $\ell(x) = d < c < \ell(x_n) \to \ell(x)$. □

Similarly, the same is true with lower semicontinuity.

Lemma 4.36. Suppose $M \subseteq X$ is a closed convex set and suppose $F : M \to \mathbb{R}$ is convex. Then $F$ is weakly sequentially lower semicontinuous if and only if it is (sequentially) lower semicontinuous.

Proof. Suppose $F$ is lower semicontinuous and thus sequentially lower semicontinuous. Let $x_n \rightharpoonup x_0$, then, by the previous lemma there is a sequence of convex combinations $y_n \to x_0$. Then $\liminf_{n \to \infty} F(x_n) \geq \liminf_{n \to \infty} F(y_n) \geq F(x_0)$ since the liminf of a convex combination of real numbers is greater or equal than the liminf of the original numbers. □

Corollary 4.37. Let $X$ be a reflexive Banach space and let $M$ be a closed convex subset. If $F : M \subseteq X \to \mathbb{R}$ is convex, lower semicontinuous, and,
if $M$ is unbounded, weakly coercive, then there exists some $x_0 \in M$ with $F(x_0) = \inf_M F$. If $F$ is strictly convex then $x_0$ is unique.

**Proof.** It remains to show uniqueness. Let $x_0$ and $x_1$ be two different minima. Then $F(\lambda x_0 + (1 - \lambda)x_1) < \lambda F(x_0) + (1 - \lambda)F(x_1) = \inf_M F$, a contradiction. \qed

**Example.** Let $X$ be a reflexive Banach space. Suppose $M \subseteq X$ is a closed convex set. Then for every $x \in X$ there is a point $x_0 \in M$ with minimal distance, $\|x - x_0\| = \text{dist}(x, M)$. Indeed, first of all choose a (sufficiently large) ball $B_r(x)$ such that $B_r(x) \cap M \neq \emptyset$. Then, since the points $y \in M$ with $\|x - y\| > r$ have larger distance, than the ones with $\|x - y\| \leq r$ we can replace $M$ by $M \cap B_r(x)$. Moreover, $F(z) = \text{dist}(x, z)$ is convex, continuous and hence the claim follows from Corollary 4.37. Note that the assumption that $X$ is reflexive is crucial (Problem 4.20).

**Example.** Let $H$ be a Hilbert space and consider the quadratic form $q_A(f) = \langle f, Af \rangle$ of a positive operator $A \geq 0$. Then since $q_A^{1/2}$ is a seminorm (Problem 1.28) and taking squares is convex, $q_A$ is convex. If $\text{Ker}(A) = 0$ it will also be weakly coercive. If we consider it on $M = B_1(0)$ we get existence of a minimum from Theorem 4.34. However this minimum is just $q_A(0) = 0$ which is not very interesting. In order to obtain a minimal eigenvalue we would need to take $M = S_1 = \{f \mid \|f\| = 1\}$ however, this set is not weakly closed (its weak closure is $B_1(0)$ as we will see in the next section). Moreover, considering an operator whose eigenvalues accumulate at 0 but where 0 itself is no eigenvalue (e.g. the multiplication operator with the sequence $\frac{1}{n}$ in $\ell^2(\mathbb{N})$) shows that the minimum is not attained on $M$ in this case.

Note that our problem with the trivial minimum at 0 would disappear if we would search for a maximum instead. However, our lemma above only guarantees us weak sequential lower semicontinuity but not weak sequential upper semicontinuity. In fact, note that not even the norm (the quadratic form of the identity) is weakly sequentially upper continuous (cf. Lemma 4.27 (i) versus Lemma 4.28). \diamond

**Problem 4.16.** Suppose $\ell_n \to \ell$ in $X^*$ and $x_n \to x$ in $X$. Then $\ell_n(x_n) \to \ell(x)$.

Similarly, suppose $\text{s-lim} \ell_n \to \ell$ and $x_n \to x$. Then $\ell_n(x_n) \to \ell(x)$.

**Problem 4.17.** Show that $x_n \to x$ implies $Ax_n \to Ax$ for $A \in \mathcal{L}(X)$.

**Problem 4.18.** Show that if $\{\ell_j\} \subseteq X^*$ is some total set, then $x_n \to x$ if and only if $x_n$ is bounded and $\ell_j(x_n) \to \ell_j(x)$ for all $j$. Show that this is wrong without the boundedness assumption (Hint: Take e.g. $X = \ell^2(\mathbb{N})$).

Problem 4.20. Consider $X = C[0,1]$ and $M = \{f | \int_0^1 f(x)dx = 1, f(0) = 0\}$. Show that $M$ is closed and convex. Show that $d(0,M) = 1$ but there is no minimizer. If we replace the boundary condition by $f(0) = 1$ there is a unique minimizer and for $f(0) = 2$ there are infinitely many minimizers.

4.6. Weak topologies

In the previous section we have defined weak convergence for sequences and this raises the question about a natural topology associated with this convergence. To this end we define the weak topology on $X$ as the weakest topology for which all $\ell \in X^*$ remain continuous and the weak-* topology on $X^*$ as the weakest topology for which all $j \in J(X) \subseteq X^{**}$ remain continuous. In particular, the weak-* topology is weaker than the weak topology on $X^*$ and both are equal if $X$ is reflexive.

Note that given a total set $\{x_n\}_{n \in \mathbb{N}} \subset X$ of (w.l.o.g.) normalized vectors

$$d(\ell, \tilde{\ell}) = \sum_{n=1}^{\infty} \frac{1}{2^n} |\ell(x_n) - \tilde{\ell}(x_n)|$$

(4.32)

defines a metric on the unit ball $B_1^*(0) \subset X^*$ which can be shown to generate the weak-* topology (cf. (iii) of Lemma 4.31). Hence Lemma 4.33 could also be stated as $B_1^*(0) \subset X^*$ being weak-* compact. This is in fact true without assuming $X$ to be separable and is known as Banach–Alaoglu theorem.

Theorem 4.38 (Banach–Alaoglu). Let $X$ be a Banach space. Then $B_1^*(0) \subset X^*$ is compact in the weak-* topology.

Proof. Abbreviate $B = \bar{B}_1^X(0)$, $B^* = \bar{B}_1^{X^*}(0)$, and $B_x = \bar{B}_{||x||}^C(0)$. Consider the (injective) map $\Phi : X^* \to \mathbb{C}^X$ given by $|\Phi(\ell)(x)| = \ell(x)$ and identify $X^*$ with $\Phi(X^*)$. Then the weak-* topology on $X^*$ coincides with the relative topology on $\Phi(X^*) \subseteq \mathbb{C}^X$ (recall that the product topology on $\mathbb{C}^X$ is the weakest topology which makes all point evaluations continuous). Moreover, $\Phi(\ell) \leq ||\ell||||x||$ implies $\Phi(B^*) \subset \prod_{x \in X} B_x$ where the last product is compact by Tychonoff’s theorem. Hence it suffices to show that $\Phi(B^*)$ is closed. To this end let $l \in \Phi(B^*)$. We need to show that $l$ is linear and bounded. Fix $x_1, x_2 \in X$, $\alpha \in \mathbb{C}$, and consider the open neighborhood

$$U(l) = \left\{ h \in \prod_{x \in B} \left| h(x_1 + x_2) - l(x_1 + \alpha x_2) \right| < \varepsilon, \quad \left| h(x_1) - l(x_1) \right| < \varepsilon, \quad \left| \alpha ||h(x_2) - l(x_2)\right| < \varepsilon \right\}$$

of $l$. Since $U(l) \cap \Phi(X^*)$ is nonempty we can choose an element $h$ from this intersection to show $|l(x_1 + \alpha x_2) - l(x_1) - \alpha l(x_2)| < 3\varepsilon$. Since $\varepsilon > 0$
4.6. Weak topologies

is arbitrary we conclude \( l(x_1 + \alpha x_2) = l(x_1) - \alpha l(x_2) \). Moreover, \( |l(x_1)| \leq |h(x_1)| + \varepsilon \leq \|x_1\| + \varepsilon \) shows \( \|l\| \leq 1 \) and thus \( l \in \Phi(B^*) \).

Since the weak topology is weaker than the norm topology every weakly closed set is also (norm) closed. Moreover, the weak closure of a set will in general be larger than the norm closure. However, for convex sets the converse is also true. Moreover, we have the following characterization in terms of closed (affine) half-spaces, that is, sets of the form \( \{ x \in X | \Re(\ell(x)) \leq \alpha \} \) for some \( \ell \in X^* \) and some \( \alpha \in \mathbb{R} \).

**Theorem 4.39** (Mazur). The weak as well as the norm closure of a convex set \( K \) is the intersection of all half-spaces containing \( K \). In particular, a convex set \( K \subseteq X \) is weakly closed if and only if it is closed.

**Proof.** Since the intersection of closed-half spaces is (weakly) closed, it suffices to show that for every \( x \) not in the (weak) closure there is a closed half-plane not containing \( x \). Moreover, if \( x \) is not in the weak closure it is also not in the norm closure (the norm closure is contained in the weak closure) and by Theorem 4.25 with \( U = B_{\text{dist}(x,K)}(x) \) and \( V = K \) there is a functional \( \ell \in X^* \) such that \( K \subseteq \Re(\ell)^{-1}([c,\infty)) \) and \( x \not\in \Re(\ell)^{-1}([c,\infty)) \).

**Example.** Suppose \( X \) is infinite dimensional. The weak closure \( \overline{S}^w \) of \( S = \{ x \in X | \|x\| = 1 \} \) is the closed unit ball \( B_1(0) \). Indeed, since \( B_1(0) \) is convex the previous lemma shows \( \overline{S}^w \subseteq B_1(0) \). Conversely, if \( x \in B_1(0) \) is not in the weak closure, then there must be an open neighborhood \( x + \bigcup_{j=1}^n |\ell_j|^{-1}([0,\varepsilon)) \) not contained in the weak closure. Since \( X \) is infinite dimensional we can find a nonzero element \( x_0 \in \bigcap_{j=1}^n \ker(\ell_j) \) such that the affine line \( x + tx_0 \) is in this neighborhood and hence also avoids \( \overline{S}^w \). But this is impossible since by the intermediate value theorem there is some \( t_0 > 0 \) such that \( \|x + t_0x_0\| = 1 \). Hence \( B_1(0) \subseteq \overline{S}^w \).

Note that this example also shows that in an infinite dimensional the weak and norm topologies are always different! In a finite dimensional space both topologies of course agree.

**Corollary 4.40.** Suppose \( x_k \rightharpoonup x \), then there are convex combinations \( y_k = \sum_{j=1}^{n_k} \lambda_{k,j} x_j \) (with \( \sum_{j=1}^{n_k} \lambda_{k,j} = 1 \) and \( \lambda_{k,j} \geq 0 \)) such that \( y_k \rightharpoonup x \).

**Proof.** Let \( K = \{ \sum_{j=1}^n \lambda_j x_j | n \in \mathbb{N}, \sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0 \} \) be the convex hull of the points \( \{ x_n \} \). Then by the previous result \( x \in K \).

**Example.** Let \( H \) be a Hilbert space and \( \{ \varphi_j \} \) some infinite ONS. Then we already know \( \varphi_j \rightharpoonup 0 \). Moreover, the convex combination \( \psi_j = \frac{1}{j} \sum_{k=1}^j \varphi_k \rightharpoonup 0 \) since \( \|\psi_j\| = j^{-1/2} \).
Problem 4.21. Suppose $K \subseteq X$ is convex and $x$ is a boundary point of $K$. Then there is a supporting hyperplane at $x$. That is, there is some $\ell \in X^*$ such that $\ell(x) = 0$ and $K$ is contained in the closed half-plane \{\(y | \text{Re}(\ell(y - x)) \leq 0\}\}.

4.7. Beyond Banach spaces: Locally convex spaces

In the last two sections we have seen that it is often important to weaken the notion of convergence (i.e., to weaken the underlying topology) to get a larger class of converging sequences. It turns out that all of them fit within a general framework which we want to briefly discuss in this section. We start with an alternate definition of a locally convex vector space which we already briefly encountered in Corollary 4.26 (equivalence of both definitions will be established below).

A vector space $X$ together with a topology is called a \textbf{locally convex vector space} if there exists a family of seminorms \(\{q_\alpha\}_{\alpha \in A}\) which generates the topology in the sense that the topology is the weakest topology for which the family of functions \(\{q_\alpha(\cdot, x)\}_{\alpha \in A, x \in X}\) is continuous. Hence the topology is generated by sets of the form $x + q_\alpha^{-1}(I)$, where $I \subseteq [0, \infty)$ is open (in the relative topology). Moreover, sets of the form

$$x + \bigcap_{j=1}^{n} q_{\alpha_j}^{-1}([0, \varepsilon_j])$$

(4.33)

are a neighborhood base at $x$ and hence it is straightforward to check that a locally convex vector space is a topological vector space, that is, both vector addition and scalar multiplication are continuous. For example, if $z = x + y$ then the preimage of the open neighborhood $z + \bigcap_{j=1}^{n} q_{\alpha_j}^{-1}([0, \varepsilon_j])$ contains the open neighborhood $(x + \bigcap_{j=1}^{n} q_{\alpha_j}^{-1}([0, \varepsilon_j/2]), y + \bigcap_{j=1}^{n} q_{\alpha_j}^{-1}([0, \varepsilon_j/2]))$ by virtue of the triangle inequality. Similarly, if $z = \gamma x$ then the preimage of the open neighborhood $z + \bigcap_{j=1}^{n} q_{\alpha_j}^{-1}([0, \varepsilon_j])$ contains the open neighborhood $(B_\varepsilon(\gamma), x + \bigcap_{j=1}^{n} q_{\alpha_j}^{-1}([0, \frac{\varepsilon_j}{2(\gamma + \varepsilon)})]$ with $\varepsilon < \frac{\varepsilon_j}{2q_{\alpha_j}(x)}$.

Moreover, note that a sequence $x_n$ will converge to $x$ in this topology if and only if $q_\alpha(x_n - x) \to 0$ for all $\alpha$.

\textbf{Example.} Of course every Banach space equipped with the norm topology is a locally convex vector space if we choose the single seminorm $q(x) = \|x\|$.

\textbf{Example.} A Banach space $X$ equipped with the weak topology is a locally convex vector space. In this case we have used the continuous linear functionals $\ell \in X^*$ to generate the topology. However, note that the corresponding seminorms $q_\ell(x) = |\ell(x)|$ generate the same topology since
4.7. Beyond Banach spaces: Locally convex spaces

Let \( x + q_\epsilon^{-1}([0, \varepsilon]) = \ell^{-1}(B_\varepsilon(x)) \) in this case. The same is true for \( X^* \) equipped with the weak or the weak*-topology. \( \diamond \)

**Example.** The bounded linear operators \( \mathfrak{L}(X,Y) \) together with the seminorms \( q_x(A) = \|Ax\| \) for all \( x \in X \) (strong convergence) or the seminorms \( q_{\ell,x}(A) = |\ell(Ax)| \) for all \( x \in X, \ell \in Y^* \) (weak convergence) are locally convex vector spaces. \( \diamond \)

**Example.** The continuous functions \( C(I) \) together with the pointwise topology generated by the seminorms \( q_x(f) = |f(x)| \) for all \( x \in I \) is a locally convex vector space. \( \diamond \)

In all these examples we have one additional property which is often required as part of the definition: The seminorms are called **separated** if for every \( x \in X \) there is a seminorm with \( q_\alpha(x) \neq 0 \). In this case the corresponding locally convex space is Hausdorff since for \( x \neq y \) the neighborhoods \( U(x) = x + q_\alpha^{-1}([0, \varepsilon]) \) and \( U(y) = y + q_\alpha^{-1}([0, \varepsilon]) \) will be disjoint provided \( \varepsilon = \frac{1}{2}q_\alpha(x-y) > 0 \) (the converse is also true; Problem 4.28).

It turns out crucial to understand when a seminorm is continuous.

**Lemma 4.41.** Let \( X \) be a locally convex vector space with corresponding family of seminorms \( \{q_\alpha\}_{\alpha \in A} \). Then a seminorm \( q \) is continuous if and only if there are seminorms \( q_{\alpha_j} \) and constants \( c_j > 0 \), \( 1 \leq j \leq n \), such that \( q(x) \leq \sum_{j=1}^n c_jq_{\alpha_j}(x) \).

**Proof.** If \( q \) is continuous, then \( q^{-1}(B_1(0)) \) contains an open neighborhood of 0 of the form \( \bigcap_{j=1}^n q_{\alpha_j}^{-1}([0, \varepsilon_j]) \) and choosing \( c_j = \max_{1 \leq j \leq n} \varepsilon_j^{-1} \) we obtain that \( \sum_{j=1}^n c_jq_{\alpha_j}(x) < 1 \) implies \( q(x) < 1 \) and the claim follows from Problem 4.23. Conversely note that if \( q(x) = r \) then \( q^{-1}(B_r(0)) \) contains the set \( U(x) = x + \bigcap_{j=1}^n q_{\alpha_j}^{-1}([0, \varepsilon_j]) \) provided \( \sum_{j=1}^n c_j\varepsilon_j \leq \varepsilon \) since \( |q(y) - q(x)| \leq q(y-x) \leq \sum_{j=1}^n c_jq_{\alpha_j}(x-y) < \varepsilon \) for \( y \in U(x) \). \( \square \)

Moreover, note that the topology is translation invariant in the sense that \( U(x) \) is a neighborhood of \( x \) if and only if \( U(x) - x = \{y-x|y \in U(x)\} \) is a neighborhood of 0. Hence we can restrict our attention to neighborhoods of 0 (this is of course true for any topological vector space). Hence if \( X \) and \( Y \) are topological vector spaces then a linear map \( A : X \to Y \) will be continuous if and only if it is continuous at 0. Moreover, if \( Y \) is a locally convex space with respect to some seminorms \( p_\beta \), then \( A \) will be continuous if and only if \( p_\beta \circ A \) is continuous for every \( \beta \) (Lemma 1.7). Finally, since \( p_\beta \circ A \) is a seminorm the previous lemma implies:

**Corollary 4.42.** Let \( (X, \{q_\alpha\}) \) and \( (Y, \{p_\beta\}) \) be locally convex vector spaces. Then a linear map \( A : X \to Y \) is continuous if and only if for every \( \beta \)
There are some seminorms \( q_{\alpha_j} \) and constants \( c_j > 0 \), \( 1 \leq j \leq n \), such that

\[
p_\beta(Ax) \leq \sum_{j=1}^{n} c_j q_{\alpha_j}(x).
\]

It will shorten notation when sums of the type \( \sum_{j=1}^{n} c_j q_{\alpha_j}(x) \) which appeared in the last two results can be replaced by a single expression \( cq_{\alpha} \).

This can be done if the family of seminorms \( \{ q_\alpha \}_{\alpha \in A} \) is directed, that is, for given \( \alpha, \beta \in A \) there is a \( \gamma \in A \) such that \( q_\alpha(x) + q_\beta(x) \leq C q_\gamma(x) \) for some \( C > 0 \). Moreover, if \( \mathcal{F}(A) \) is the set of all finite subsets of \( A \), then \( \{ \tilde{q}_F = \sum_{\alpha \in F} q_\alpha \}_{F \in \mathcal{F}(A)} \) is a directed family which generates the same topology (since every \( \tilde{q}_F \) is continuous with respect to the original family we do not get any new open sets).

While the family of seminorms is in most cases more convenient to work with, it is important to observe that different families can give rise to the same topology and it is only the topology which matters for us. In fact, it is possible to characterize locally convex vector spaces as topological vector spaces which have a neighborhood basis at 0 of absolutely convex sets. Here a set \( U \) is called absolutely convex, if for \( |\alpha| + |\beta| \leq 1 \) we have \( \alpha U + \beta U \subseteq U \). Since the sets \( q^{-1}_\alpha([0, \varepsilon)) \) are absolutely convex we always have such a basis in our case. To see the converse note that such a neighborhood \( U \) of 0 is also absorbing (Problem 4.22) and hence the corresponding Minkowski functional (4.18) is a seminorm (Problem 4.27). By construction these seminorms generate the topology since if \( U_0 = \bigcap_{j=1}^{n} q^{-1}_{\alpha_j}([0, \varepsilon_j)) \subseteq U \) we have for the corresponding Minkowski functionals \( p_U(x) \leq p_{U_0}(x) \leq \varepsilon^{-1} \sum_{j=1}^{n} q_{\alpha_j}(x) \), where \( \varepsilon = \min \varepsilon_j \). With a little more work (Problem 4.26) one can even show that it suffices to assume to have a neighborhood basis at 0 of convex open sets.

Given a topological vector space \( X \) we can define its dual space \( X^* \) as the set of all continuous linear functionals. However, while it can happen in general that the dual space is empty, \( X^* \) will always be nontrivial for a locally convex space since the Hahn–Banach theorem can be used to construct linear functionals (using a continuous seminorm for \( \varphi \) in Theorem 4.11) and also the geometric Hahn–Banach theorem (Theorem 4.25) holds. Moreover, it is not difficult to see that \( X^* \) will be Hausdorff if and only if \( X \) is. In this respect note that for every continuous linear functional \( \ell \) in a topological vector space \( |\ell|^{-1}([0, \varepsilon)) \) is an absolutely convex open neighborhoods of 0 and hence existence of such sets is necessary for the existence of nontrivial continuous functionals. As a natural topology on \( X^* \) we could use the weak-* topology defined to be the weakest topology generated by the family of all point evaluations \( q_\ell(x) = |\ell(x)| \) for all \( x \in X \). Given a continuous linear operator \( A : X \to Y \) between locally convex spaces we can define its adjoint
4.7. Beyond Banach spaces: Locally convex spaces

\( A' : Y^* \to X^* \) as before,

\[
(A'y^*)(x) = y^*(Ax)
\]  

(4.34)

a brief calculation

\[
q_x(A'y^*) = |(A'y^*)(x)| = |y^*(Ax)| = q_{Ax}(y^*).
\]  

(4.35)

verifies that \( A' \) is continuous in the weak-\(^*\) topology by virtue of Corollary 4.42.

The remaining theorems we have established for Banach spaces were consequences of the Baire theorem and this leads us to the question when a locally convex space is a metric space. From our above analysis we see that a locally convex vector space will be first countable if and only if countably many seminorms suffice to determine the topology. In this case \( X \) turns out to be metrizable. As a preparation we show:

**Lemma 4.43.** Suppose we have a countable family of functions \( f_n : X \to Y_n \), where \( Y_n \) are (pseudo-)metric spaces. Then

\[
d(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{d_n(f_n(x), f_n(y))}{1 + d_n(f_n(x), f_n(y))}
\]

(4.36)

is a pseudometric on \( X \) which generates the initial topology of these functions.

Moreover, if \( Y_n \) are metric spaces, it will be a metric if and only if the family \( f_n \) separates points, that is, for \( x \neq y \) there is some \( n \) with \( f_n(x) \neq f_n(y) \).

**Proof.** The important observation is that \( d \) can be made small by making a (sufficiently large) finite part of the series small. More precisely let \( O \) be open with respect to this pseudometric and pick some \( x \in O \). Then there is some ball \( B_r(x) \subseteq O \) and we choose \( n \) and \( \varepsilon_j \) such that \( \sum_{j=1}^{\infty} 2^{-j} \frac{\varepsilon_j}{1+\varepsilon_j} + \sum_{j=n+1}^{\infty} 2^{-j} < r \). Then \( \bigcap_{j=1}^{n} f_{j}^{-1}(B_{\varepsilon_j}(f_j(x))) = \{ y | d_j(f_j(x), f_j(y)) < \varepsilon_j, j = 1, \cdots, n \} \subseteq B_r(x) \subseteq O \) shows that \( x \) is an interior point with respect to the initial topology. Since \( x \) was arbitrary, \( O \) is open with respect to the initial topology. Conversely, let \( O \) be open with respect to the initial topology and pick some \( x \in O \). Then there is some set \( \bigcap_{j=1}^{n} f_{n_j}^{-1}(O_{n_j}) \subseteq O \) containing \( x \). The same will continue to hold if we replace \( O_{n_j} \) by a ball of radius \( \varepsilon_j \) around \( f_{n_j}(x) \). Now choose \( r < 2^{-N} \frac{\varepsilon_j}{1+\varepsilon_j} \) where \( N = \max n_j \) and \( \varepsilon = \min \varepsilon_j \). Then \( d(x, y) < r \) implies \( d_{n_j}(f_{n_j}(x), f_{n_j}(y)) < \varepsilon < \varepsilon_j \) and thus \( B_r(x) \subseteq \bigcap_{j=1}^{n} f_{n_j}^{-1}(O_{n_j}) \subseteq O \). So again \( O \) is open with respect to our pseudometric.

The last claim is straightforward. \( \square \)
Theorem 4.44. A locally convex Hausdorff space is metrizable if and only if it is first countable. In this case there is a countable family of seminorms generating the topology and a metric is given by

\[ d(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{q_n(x - y)}{1 + q_n(x - y)}. \] (4.37)

Proof. If \( X \) is first countable there is a countable neighborhood base at 0 and hence also a countable neighborhood base of absolutely convex sets. The Minkowski functionals corresponding to the latter base are seminorms of the required type. The rest follows from the previous lemma. \( \square \)

In general, a locally convex vector space \( X \) which has a separated countable family of seminorms is called a Fréchet space if it is complete with respect to the metric (4.37). Note that the metric (4.37) is translation invariant

\[ d(f, g) = d(f - h, g - h). \] (4.38)

Example. The continuous functions \( C(\mathbb{R}) \) together with local uniform convergence are a Fréchet space. A countable family of seminorms is for example

\[ \|f\|_j = \sup_{|x| \leq j} |f(x)|, \quad j \in \mathbb{N}. \] (4.39)

Then \( f_k \to f \) if and only if \( \|f_k - f\|_j \to 0 \) for all \( j \in \mathbb{N} \) and it follows that \( C(\mathbb{R}) \) is complete. \( \diamond \)

Example. The space \( C^\infty(\mathbb{R}^m) \) together with the seminorms

\[ \|f\|_{j,k} = \sum_{|\alpha| \leq j} \sup_{|x| \leq k} |\partial_\alpha f(x)|, \quad j \in \mathbb{N}_0, \; k \in \mathbb{N}, \] (4.40)

is a Fréchet space. \( \diamond \)

Example. The Schwartz space

\[ \mathcal{S}(\mathbb{R}^m) = \{ f \in C^\infty(\mathbb{R}^m) | \sup_x |x^\alpha (\partial_\beta f)(x)| < \infty, \; \forall \alpha, \beta \in \mathbb{N}_0^m \} \] (4.41)

together with the seminorms

\[ q_{\alpha,\beta}(f) = \|x^\alpha (\partial_\beta f)(x)\|_\infty, \quad \alpha, \beta \in \mathbb{N}_0^m. \] (4.42)

To see completeness note that a Cauchy sequence \( f_n \) is in particular a Cauchy sequence in \( C^\infty(\mathbb{R}^m) \). Hence there is a limit \( f \in C^\infty(\mathbb{R}^m) \) such that all derivatives converge uniformly. Moreover, since Cauchy sequences are bounded \( \|x^\alpha (\partial_\beta f_n)(x)\|_\infty \leq C_{\alpha,\beta} \) we obtain \( \|x^\alpha (\partial_\beta f)(x)\|_\infty \leq C_{\alpha,\beta} \) and thus \( f \in \mathcal{S}(\mathbb{R}^m) \). The dual space \( \mathcal{S}^*(\mathbb{R}^m) \) is known as the space of tempered distributions. \( \diamond \)
Example. The space of all entire functions $f(z)$ (i.e. functions which are holomorphic on all of $C$) together with the seminorms $\|f\|_j = \sup_{|z| \leq j} |f(z)|$, $j \in \mathbb{N}$, is a Fréchet space. Completeness follows from the Weierstraß convergence theorem which states that a limit of holomorphic functions which is uniform on every compact subset is again holomorphic.

Of course the question which remains is if these spaces could be made into Banach spaces, that is, if we can find a single seminorm which describes the topology. To this end we call a set $B \subseteq X$ bounded if $\sup_{x \in B} q_\alpha(x) < \infty$ for every $\alpha$. By Corollary 4.42 this will then be true for any continuous seminorm on $X$.

**Theorem 4.45** (Kolmogorov). A locally convex vector space can be generated from a single seminorm if and only if it contains a bounded open set.

**Proof.** In a Banach space every open ball is bounded and hence only the converse direction is nontrivial. So let $U$ be a bounded open set. By shifting and decreasing $U$ if necessary we can assume $U$ to be an absolutely convex open neighborhood of 0 and consider the associated Minkowski functional $q = p_U$. Then since $U = \{x|q(x) < 1\}$ and $\sup_{x \in U} q_\alpha(x) = C_\alpha < \infty$ we infer $q_\alpha(x) \leq C_\alpha q(x)$ (Problem 4.23) and thus the single seminorm $q$ generates the topology. □

Example. The space $C(I)$ with the pointwise topology generated by $q_x(f) = |f(x)|$, $x \in I$ cannot be given by a norm. Otherwise there would be a bounded open neighborhood of zero. This neighborhood contains a neighborhood of the form $\bigcup_{j=1}^n q^{-1}_j((0, \varepsilon))$ which is also bounded. In particular, the set $U = \{f \in C(I) | |f(x_j)| < \varepsilon_j, 1 \leq j \leq n\}$ is bounded. Now choose $x_0 \in I \setminus \{x_j\}_{j=1}^n$, then $\sup_{f \in U} |f(x_0)| = \infty$ (take a piecewise linear function which vanishes at the points $x_j, 1 \leq j \leq n$ and takes arbitrary large values at $x_0$) contradicting our assumption. □

In a similar way one can show that the other spaces introduce above are no Banach spaces.

Finally, we mention that since the Baire category theorem holds for arbitrary complete metric spaces, the open mapping theorem (Theorem 4.4), the inverse mapping theorem (Theorem 4.5) and the closed graph (Theorem 4.6) hold for Fréchet spaces without modifications.

**Problem 4.22.** In a topological vector space every neighborhood $U$ of 0 is absorbing.

**Problem 4.23.** Let $p, q$ be two seminorms. Then $p(x) \leq Cq(x)$ if and only if $q(x) < 1$ implies $p(x) < C$. 
Problem 4.24. Let $X$ be a vector space. We call a set $U$ balanced if $\alpha U \subseteq U$ for every $|\alpha| \leq 1$. Show that a set is balanced and convex if and only if it is absolutely convex.

Problem 4.25. The intersection of arbitrary convex/balanced/absolutely convex sets is again convex/balanced/absolutely convex. Hence we can define the convex/balanced/absolutely convex hull of a set $U$ as the smallest convex/balanced/absolutely convex set containing $U$, that is, the intersection of all convex/balanced/absolutely convex sets containing $U$. Show that the convex hull is given by

$$\text{hull}(U) = \left\{ \sum_{j=1}^{n} \lambda_j x_j | n \in \mathbb{N}, x_j \in U, \sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0 \right\},$$

the balanced hull is given by

$$\text{bhull}(U) = \{ \alpha x | x \in U, |\alpha| \leq 1 \},$$

and the absolutely convex hull is given by

$$\text{ahull}(U) = \left\{ \sum_{j=1}^{n} \lambda_j x_j | n \in \mathbb{N}, x_j \in U, \sum_{j=1}^{n} |\lambda_j| \leq 1 \right\}.$$

Show that $\text{ahull}(U) = \text{hull}(\text{bhull}(U))$.

Problem 4.26. In a topological vector space every convex open neighborhood $U$ of zero contains an absolutely convex open neighborhood of zero. (Hint: By continuity of the scalar multiplication $U$ contains a set of the form $B_C^c(0) \cdot V$, where $V$ is an open neighborhood of zero.)

Problem 4.27. Let $X$ be a vector space. Show that the Minkowski functional of a balanced, convex, absorbing set is a seminorm.

Problem 4.28. If a locally convex space is Hausdorff then any corresponding family of seminorms is separated.

Problem 4.29. Let $X$ be some space together with a sequence of pseudometrics $d_j, j \in \mathbb{N}$. Show that

$$d(x, y) = \sum_{j \in \mathbb{N}} \frac{1}{2^j} \frac{d_j(x, y)}{1 + d_j(x, y)}$$

is again a pseudometric. It is a metric if and only if $d_j(x, y) = 0$ for all $j$ implies $x = y$.

Problem 4.30. Suppose $X$ is a metric vector space. Then balls are convex if and only if

$$d(\lambda x + (1 - \lambda)y, z) \leq \max\{d(x, z), d(y, z)\}, \quad \lambda \in (0, 1).$$
Problem 4.31. Consider the Fréchet space $C(\mathbb{R})$ with $q_n(f) = \sup_{[-n,n]} |f|$. Show that the metric balls are not convex.

Problem 4.32. Consider $\ell^p(\mathbb{N})$ for $p \in (0,1)$ — compare Problem 1.20. Show that $\|\cdot\|_p$ is not convex. Show that every convex open set is unbounded. Conclude that it is not a locally convex vector space. (Hint: Consider $B_R(0)$. Then for $r < R$ all vectors which have one entry equal to $r$ and all other entries zero are in this ball. By taking convex combinations all vectors which have $n$ entries equal to $r/n$ are in the convex hull. The quasinorm of such a vector is $n^{1/p-1} r$.)
Chapter 5

More on compact operators

5.1. Canonical form of compact operators

Our first aim is to find a generalization of Corollary 3.9 for general compact operators in a Hilbert space $H$. The key observation is that if $K$ is compact, then $K^*K$ is compact and symmetric and thus, by Corollary 3.9, there is a countable orthonormal set $\{u_j\}$ and nonzero numbers $s_j \neq 0$ such that

$$K^*Kf = \sum_j s_j^2 \langle u_j, f \rangle u_j.$$  \hfill (5.1)

Moreover, $\|Ku_j\|^2 = \langle u_j, K^*Ku_j \rangle = \langle u_j, s_j^2 u_j \rangle = s_j^2$ implies

$$s_j = \|Ku_j\| > 0.$$  \hfill (5.2)

The numbers $s_j = s_j(K)$ are called **singular values** of $K$. There are either finitely many singular values or they converge to zero.

**Theorem 5.1** (Canonical form of compact operators). Let $K$ be compact and let $s_j$ be the singular values of $K$ and $\{u_j\}$ corresponding orthonormal eigenvectors of $K^*K$. Then

$$K = \sum_j s_j \langle u_j, \cdot \rangle v_j,$$  \hfill (5.3)

where $v_j = s_j^{-1}Ku_j$. The norm of $K$ is given by the largest singular value

$$\|K\| = \max_j s_j(K).$$  \hfill (5.4)

Moreover, the vectors $v_j$ are again orthonormal and satisfy $K^*v_j = s_j u_j$. In particular, $v_j$ are eigenvectors of $KK^*$ corresponding to the eigenvalues $s_j^2$. 

125
Proof. For any $f \in \mathcal{H}$ we can write

$$f = \sum_j (u_j, f)u_j + f_\perp$$

with $f_\perp \in \text{Ker}(K^*K) = \text{Ker}(K)$ (Problem 5.1). Then

$$Kf = \sum_j (u_j, f)Ku_j = \sum_j s_j(u_j, f)v_j$$

as required. Furthermore,

$$\langle v_j, v_k \rangle = (s_js_k)^{-1}(Ku_j, Ku_k) = (s_js_k)^{-1}(K^*Ku_j, u_k) = s_js_k^{-1}(u_j, u_k)$$

shows that $\{v_j\}$ are orthonormal. By definition $K^*v_j = s_j^{-1}K^*Ku_j = s_jv_j$ which also shows $KK^*v_j = s_jKu_j = s_j^2v_j$.

Finally, (5.4) follows using Bessel’s inequality

$$\|Kf\|^2 = \|\sum_j s_j(u_j, f)v_j\|^2 = \sum_j s_j^2|\langle u_j, f \rangle|^2 \leq \left( \max_j s_j(K)^2 \right) \|f\|^2,$$

where equality holds for $f = u_j$ if $s_{j_0} = \max_j s_j(K)$.□

If $K$ is self-adjoint, then $u_j = \sigma_j v_j$, $\sigma_j^2 = 1$, are the eigenvectors of $K$ and $\sigma_js_j$ are the corresponding eigenvalues. The above theorem also gives rise to the polar decomposition

$$K = U|K| = |K^*|U,$$

where

$$|K| = \sqrt{K^*K} = \sum_j s_j(u_j, .)u_j, \quad |K^*| = \sqrt{KK^*} = \sum_j s_j(v_j, .)v_j$$

are self-adjoint (in fact nonnegative) and

$$U = \sum_j \langle u_j, . \rangle v_j$$

is an isometry from $\text{Ran}(K^*) = \text{span}\{u_j\}$ onto $\text{Ran}(K) = \text{span}\{v_j\}$.

From the max-min theorem (Theorem 3.12) we obtain:

**Lemma 5.2.** Let $K$ be compact; then

$$s_j(K) = \min_{f_1, \ldots, f_{j-1}} \sup_{f \in U(f_1, \ldots, f_{j-1})} \|Kf\|,$$

where $U(f_1, \ldots, f_j) = \{ f \in \mathcal{H} | \|f\| = 1, \ f \in \text{span}\{f_1, \ldots, f_j\}^\perp \}$.

In particular, note

$$s_j(AK) \leq \|A\|s_j(K), \quad s_j(KA) \leq \|A\|s_j(K)$$

whenever $K$ is compact and $A$ is bounded (the second estimate follows from the first by taking adjoints).
5.1. Canonical form of compact operators

An operator $K \in \mathcal{L}(\mathcal{H})$ is called a **finite rank operator** if its range is finite dimensional. The dimension

$$\text{rank}(K) = \dim \text{Ran}(K)$$

is called the **rank** of $K$. Since for a compact operator

$$\text{Ran}(K) = \text{span}\{v_j\}$$

we see that a compact operator is finite rank if and only if the sum in (5.3) is finite. Note that the finite rank operators form an ideal in $\mathcal{L}(\mathcal{H})$ just as the compact operators do. Moreover, every finite rank operator is compact by the Heine–Borel theorem (Theorem 1.17).

Now truncating the sum in the canonical form gives us a simple way to approximate compact operators by finite rank one. Moreover, this is in fact the best approximation within the class of finite rank operators:

**Lemma 5.3.** Let $K$ be compact and let its singular values be ordered. Then

$$s_j(K) = \min_{\text{rank}(F) < j} \|K - F\|,$$

with equality for

$$F_{j-1} = \sum_{k=1}^{j-1} s_k \langle u_k, . \rangle v_k.$$  

(5.11)

In particular, the closure of the ideal of finite rank operators in $\mathcal{L}(\mathcal{H})$ is the ideal of compact operators.

**Proof.** That there is equality for $F = F_{j-1}$ follows from (5.4). In general, the restriction of $F$ to $\text{span}\{u_1, \ldots, u_j\}$ will have a nontrivial kernel. Let $f = \sum_{k=1}^j \alpha_j u_j$ be a normalized element of this kernel, then $\|(K - F)f\|^2 = \|Kf\|^2 = \sum_{k=1}^j |\alpha_k s_k|^2 \geq s_j^2$.

In particular, every compact operator can be approximated by finite rank ones and since the limit of compact operators is compact, we cannot get more than the compact operators. □

Moreover, this also shows that the adjoint of a compact operator is again compact.

**Corollary 5.4.** An operator $K$ is compact (finite rank) if and only $K^*$ is. In fact, $s_j(K) = s_j(K^*)$ and

$$K^* = \sum_j s_j \langle v_j, . \rangle u_j.$$  

(5.13)

**Proof.** First of all note that (5.13) follows from (5.3) since taking adjoints is continuous and $(\langle u_j, . \rangle v_j)^* = \langle v_j, . \rangle u_j$ (cf. Problem 2.6). The rest is straightforward. □
From this last lemma one easily gets a number of useful inequalities for the singular values:

**Corollary 5.5.** Let $K_1$ and $K_2$ be compact and let $s_j(K_1)$ and $s_j(K_2)$ be ordered. Then

(i) $s_{j+k-1}(K_1 + K_2) \leq s_j(K_1) + s_k(K_2)$,

(ii) $s_{j+k-1}(K_1K_2) \leq s_j(K_1)s_k(K_2)$,

(iii) $|s_j(K_1) - s_j(K_2)| \leq \|K_1 - K_2\|$.

**Proof.** Let $F_1$ be of rank $j - 1$ and $F_2$ of rank $k - 1$ such that $\|K_1 - F_1\| = s_j(K_1)$ and $\|K_2 - F_2\| = s_k(K_2)$. Then $s_{j+k-1}(K_1 + K_2) \leq \|(K_1 + K_2) - (F_1 + F_2)\| = \|K_1 - F_1\| + \|K_2 - F_2\| = s_j(K_1) + s_k(K_2)$ since $F_1 + F_2$ is of rank at most $j + k - 2$.

Similarly $F = F_1(K_2 - F_2) + K_1F_2$ is of rank at most $j + k - 2$ and hence $s_{j+k-1}(K_1K_2) \leq \|K_1K_2 - F\| = \|(K_1 - F_1)(K_2 - F_2)\| \leq \|K_1 - F_1\|\|K_2 - F_2\| = s_j(K_1)s_k(K_2)$.

Next, choosing $k = 1$ and replacing $K_2 \to K_2 - K_1$ in (i) gives $s_j(K_2) \leq s_j(K_1) + \|K_2 - K_1\|$. Reversing the roles gives $s_j(K_1) \leq s_j(K_2) + \|K_1 - K_2\|$ and proves (iii). □

**Example.** On might hope that item (i) from the previous corollary can be improved to $s_j(K_1 + K_2) \leq s_j(K_1) + s_j(K_2)$. However, this is not the case as the following example shows:

$$K_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad K_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$  

Then $1 = s_2(K_1 + K_2) \nleq s_2(K_1) + s_2(K_2) = 0$.  

Finally, let me remark that there are a number of other equivalent definitions for compact operators.

**Lemma 5.6.** For $K \in \mathcal{L}(\mathfrak{H})$ the following statements are equivalent:

(i) $K$ is compact.

(ii) $A_n \in \mathcal{L}(\mathfrak{H})$ and $A_n \xrightarrow{\ast} A$ strongly implies $A_nK \to AK$.

(iii) $f_n \to f$ weakly implies $Kf_n \to Kf$ in norm.

**Proof.** Denote by (i*), (ii*), (iii*) the corresponding statements for $K^*$ and note that (i*) ⇔ (i) by Corollary 5.4.

(i) ⇒ (ii). Translating $A_n \to A_n - A$, it is no restriction to assume $A = 0$. Since $\|A_n\| \leq M$, it suffices to consider the case where $K$ is finite rank. Then
using (5.3) and applying the triangle plus Cauchy–Schwarz inequalities
\[ \| A_nK \|^2 \leq \sup_{\| f \|=1} \left( \sum_{j=1}^{N} s_j \| \langle u_j, f \rangle \| A_n v_j \| \right)^2 \leq \sum_{j=1}^{N} s_j^2 \| A_n v_j \|^2 \to 0. \]

(ii) ⇒ (iii*). Again, replace \( f_n \to f_n - f \) and assume \( f = 0 \). Choose \( A_n = \langle f_n, . \rangle u \), \( \| u \| = 1 \). Then \( \| K^* f_n \| = \| A_n K \| \to 0 \).

(iii) ⇒ (i). If \( f_n \) is bounded, it has a weakly convergent subsequence by Theorem 4.30. Now apply (iii) to this subsequence. □

Moreover, note that one cannot replace \( A_n K \to KA_n \to KA \) in (ii) as the following example shows.

**Example.** Let \( H = \ell^2(\mathbb{N}) \) and let \( S_n \) be the operator which shifts each sequence \( n \) places to the left and let \( K = \langle \delta_1, . \rangle \delta_1 \), where \( \delta_1 = (1, 0, \ldots) \). Then \( s\text{-lim} S_n = 0 \) but \( \| KS_n \| = 1 \).

Problem 5.1. Show that \( \ker(A^* A) = \ker(A) \) for any \( A \in \mathcal{L}(H) \).

Problem 5.2. Let \( K \) be multiplication by a sequence \( k \in c_0(\mathbb{N}) \) in the Hilbert space \( \ell^2(\mathbb{N}) \). What are the singular values of \( K \)?

### 5.2. Hilbert–Schmidt and trace class operators

We can further subdivide the class of compact operators \( \mathcal{C}(H) \) according to the decay of their singular values. We define
\[ \| K \|_p = \left( \sum_j s_j(K)^p \right)^{1/p} \] (5.14)

plus corresponding spaces
\[ J_p(H) = \{ K \in \mathcal{C}(H) | \| K \|_p < \infty \}, \] (5.15)

which are known as **Schatten p-classes**. Even though our notation hints at the fact that \( \| . \| \) is a norm we will not prove this here (the only nontrivial part is the triangle inequality). Note that by (5.4)
\[ \| K \| \leq \| K \|_p \] (5.16)

and that by \( s_j(K) = s_j(K^*) \) we have
\[ \| K \|_p = \| K^* \|_p. \] (5.17)

The two most important cases are \( p = 1 \) and \( p = 2 \): \( J_2(H) \) is the space of **Hilbert–Schmidt operators** and \( J_1(H) \) is the space of **trace class operators**.

We first prove an alternate definition for the Hilbert–Schmidt norm.
Lemma 5.7. A bounded operator $K$ is Hilbert–Schmidt if and only if
\[ \sum_{j \in J} \|Kw_j\|^2 < \infty \]  
(5.18)
for some orthonormal basis and
\[
\|K\|_2 = \left( \sum_{j \in J} \|Kw_j\|^2 \right)^{1/2},
\]  
(5.19)
for every orthonormal basis in this case.

Proof. First of all note that (5.18) implies that $K$ is compact. To see this, let $P_n$ be the projection onto the space spanned by the first $n$ elements of the orthonormal basis $\{w_j\}$. Then $K_n = KP_n$ is finite rank and converges to $K$ since
\[
\|(K - K_n)f\| = \left( \sum_{j \in J} \|c_j K w_j\| \right) \leq \left( \sum_{j \in J} \|K w_j\|^2 \right)^{1/2} \|f\|,
\]
where $f = \sum c_j w_j$.

The rest follows from (5.3) and
\[
\sum_j \|Kw_j\|^2 = \sum_{k,j} |\langle v_k, K w_j \rangle|^2 = \sum_{k,j} |\langle K^* v_k, w_j \rangle|^2 = \sum_k \|K^* v_k\|^2
\]
\[= \sum_k s_k(K)^2 = \|K\|_2^2.
\]
Here we have used $\text{span}\{v_k\} = \text{Ker}(K^*)^\perp = \text{Ran}(K)$ in the first step.

Now we can show

Lemma 5.8. The set of Hilbert–Schmidt operators forms an ideal in $\mathfrak{L}(\mathfrak{H})$ and
\[
\|KA\|_2 \leq \|A\| \|K\|_2, \quad \text{respectively,} \quad \|AK\|_2 \leq \|A\| \|K\|_2.
\]  
(5.20)

Proof. If $K_1$ and $K_2$ are Hilbert–Schmidt operators, then so is their sum since
\[
\|K_1 + K_2\|_2 = \left( \sum_{j \in J} \| (K_1 + K_2) w_j \|^2 \right)^{1/2} \leq \left( \sum_{j \in J} (\|K_1 w_j\|^2 + \|K_2 w_j\|^2) \right)^{1/2}
\]
\[\leq \|K_1\|_2 + \|K_2\|_2,
\]
where we have used the triangle inequality for $\ell^2(J)$.

Let $K$ be Hilbert–Schmidt and $A$ bounded. Then $AK$ is compact and
\[
\|AK\|_2^2 = \sum_j \|AK w_j\|^2 \leq \|A\|^2 \sum_j \|K w_j\|^2 = \|A\|^2 \|K\|_2^2.
\]
For $KA$ just consider adjoints. \(\square\)
Example. Consider $\ell^2(\mathbb{N})$ and let $K$ be some compact operator. Let $K_{jk} = \langle \delta^j, K\delta^k \rangle = (K\delta^j)_k$ be its matrix elements such that

$$(Ka)_j = \sum_{k=1}^{\infty} K_{jk}a_k.$$ 

Then, choosing $w_j = \delta^j$ in (5.19) we get

$$\|K\|_2 = \left( \sum_{j=1}^{\infty} \|K\delta^j\|^2 \right)^{1/2} = \left( \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} |K_{jk}|^2 \right)^{1/2}.$$ 

Hence $K$ is Hilbert–Schmidt if and only if its matrix elements are in $\ell^2(\mathbb{N} \times \mathbb{N})$ and the Hilbert–Schmidt norm coincides with the $\ell^2(\mathbb{N} \times \mathbb{N})$ norm of the matrix elements. Especially in the finite dimensional case the Hilbert–Schmidt norm is also known as Frobenius norm.

Of course the same calculation shows that a bounded operator is Hilbert–Schmidt if and only if its matrix elements $\langle w_j, KW_k \rangle$ with respect to some orthonormal basis $\{w_j\}_{j \in J}$ are in $\ell^2(J \times J)$ and the Hilbert–Schmidt norm coincides with the $\ell^2(J \times J)$ norm of the matrix elements.

Since Hilbert–Schmidt operators turn out easy to identify (cf. also Section 8.5), it is important to relate $J_1(\delta)$ with $J_2(\delta)$:

**Lemma 5.9.** An operator is trace class if and only if it can be written as the product of two Hilbert–Schmidt operators, $K = K_1K_2$, and in this case we have

$$\|K\|_1 \leq \|K_1\|_2\|K_2\|_2.$$ 

**Proof.** Using (5.3) (where we can extend $u_n$ and $v_n$ to orthonormal bases if necessary) and Cauchy–Schwarz we have

$$\|K\|_1 = \sum_n \langle v_n, Ku_n \rangle = \sum_n |\langle K_1^*v_n, K_2u_n \rangle| \leq \left( \sum_n \|K_1^*v_n\|^2 \sum_n \|K_2u_n\|^2 \right)^{1/2} = \|K_1\|_2\|K_2\|_2$$

and hence $K = K_1K_2$ is trace class if both $K_1$ and $K_2$ are Hilbert–Schmidt operators. To see the converse, let $K$ be given by (5.3) and choose $K_1 = \sum_j \sqrt{s_j(K)}\langle u_j, . \rangle v_j$, respectively, $K_2 = \sum_j \sqrt{s_j(K)}\langle u_j, . \rangle u_j$. 

Now we can also explain the name trace class:

**Lemma 5.10.** If $K$ is trace class, then for every orthonormal basis $\{w_n\}$ the trace

$$\text{tr}(K) = \sum_n \langle w_n, Kw_n \rangle$$ 

(5.22)
is finite

\[ |\text{tr}(K)| \leq \|K\|_1. \]  

(5.23)

and independent of the orthonormal basis.

\textbf{Proof.} We first compute the trace with respect to \( \{v_n\} \) using (5.3)

\[ \text{tr}(K) = \sum_k \langle v_k, Kv_k \rangle = \sum_{k,j} s_j \langle u_j, v_k \rangle \langle v_k, v_j \rangle = \sum_k s_k \langle u_k, v_k \rangle \]

which shows (5.23) upon taking absolute values on both sides.

Next, let \( \{w_n\} \) and \( \{\tilde{w}_n\} \) be two orthonormal bases. If we write \( K = K_1K_2 \) with \( K_1, K_2 \) Hilbert–Schmidt, we have

\[ \sum_n \langle w_n, K_1K_2w_n \rangle = \sum_n \langle K_1^*w_n, K_2w_n \rangle = \sum_{n,m} \langle K_1^*w_n, \tilde{w}_m \rangle \langle \tilde{w}_m, K_2w_n \rangle \]
\[ = \sum_{m,n} \langle K_2^*v_m, w_n \rangle \langle w_n, K_1v_m \rangle = \sum_m \langle K_2^*\tilde{w}_m, K_1\tilde{w}_m \rangle \]
\[ = \sum_m \langle \tilde{w}_m, K_2K_1\tilde{w}_m \rangle. \]

In the special case \( w = \tilde{w} \) we see \( \text{tr}(K_1K_2) = \text{tr}(K_2K_1) \) and the general case now shows that the trace is independent of the orthonormal basis. \( \square \)

Clearly for self-adjoint trace class operators, the trace is the sum over all eigenvalues (counted with their multiplicity). To see this, one just has to choose the orthonormal basis to consist of eigenfunctions. This is even true for all trace class operators and is known as Lidskij trace theorem (see [23] for an easy to read introduction).

We also note the following elementary properties of the trace:

\textbf{Lemma 5.11.} Suppose \( K, K_1, K_2 \) are trace class and \( U \) is unitary.

(i) The trace is linear.

(ii) \( \text{tr}(K^*) = \text{tr}(K)^* \).

(iii) If \( K_1 \leq K_2 \), then \( \text{tr}(K_1) \leq \text{tr}(K_2) \).

(iv) \( \text{tr}(UK) = \text{tr}(KU) \).

\textbf{Proof.} (i) and (ii) are straightforward. (iii) follows from \( K_1 \leq K_2 \) if and only if \( \langle f, K_1f \rangle \leq \langle f, K_2f \rangle \) for every \( f \in \mathcal{H} \). (iv) Let \( \{u_j\} \) be some ONB and note that \( \{v_j = Uu_j\} \) is also an ONB. Then

\[ \text{tr}(UK) = \sum_j \langle v_j, UKv_j \rangle = \sum_j \langle Uu_j, KUu_j \rangle \]
\[ = \sum_j \langle u_j, KUu_j \rangle = \text{tr}(KU) \]
and the claim follows. □

We also mention a useful criterion for $K$ to be trace class.

**Lemma 5.12.** An operator $K$ is trace class if and only if it can be written as

$$K = \sum_j \langle f_j, \cdot \rangle g_j$$

(5.24)

for some sequences $f_j$, $g_j$ satisfying

$$\sum_j \|f_j\|\|g_j\| < \infty.$$  

(5.25)

**Proof.** To see that a trace class operator (5.3) can be written in such a way choose $f_j = u_j$, $g_j = s_j v_j$. Conversely note that for every finite $N$ we have

$$\sum_{k=1}^N s_k = \sum_{k=1}^N \langle v_k, Ku_k \rangle = \sum_{k=1}^N \sum_j \langle v_k, g_j \rangle \langle f_j, u_k \rangle = \sum_j \sum_{k=1}^N \langle v_k, g_j \rangle \langle f_j, u_k \rangle$$

$$\leq \sum_j \left( \sum_{k=1}^N |\langle v_k, g_j \rangle|^2 \right)^{1/2} \left( \sum_{k=1}^N |\langle f_j, u_k \rangle|^2 \right)^{1/2} \leq \sum_j \|f_j\|\|g_j\|. \quad \Box$$

Finally, note that

$$\|K\|_2 = (\operatorname{tr}(K^*K))^{1/2}$$

(5.26)

which shows that $\mathcal{H}_2(\mathfrak{H})$ is in fact a Hilbert space with scalar product given by

$$\langle K_1, K_2 \rangle = \operatorname{tr}(K_1^*K_2).$$

(5.27)

**Problem 5.3.** Let $\mathfrak{H} = \ell^2(\mathbb{N})$ and let $A$ be multiplication by a sequence $a = (a_j)_{j=1}^{\infty}$. Show that $A$ is Hilbert–Schmidt if and only if $a \in \ell^2(\mathbb{N})$. Furthermore, show that $\|A\|_2 = \|a\|$ in this case.

**Problem 5.4.** An operator of the form $K : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$, $f_n \mapsto \sum_{j\in\mathbb{N}} k_{n+j} f_j$ is called Hankel operator.

- Show that $K$ is Hilbert–Schmidt if and only if $\sum_{j\in\mathbb{N}} j |k_j|^2 < \infty$ and this number equals $\|K\|_2$.
- Show that $K$ is Hilbert–Schmidt with $\|K\|_2 \leq \|c\|_1$ if $|k_j| \leq c_j$, where $c_j$ is decreasing and summable.

(Hint: For the first item use summation by parts.)
5. More on compact operators

5.3. Fredholm theory for compact operators

In this section we want to investigate solvability of the equation

\[ f = Kf + g \]  \hspace{1cm} (5.28)

for given \( g \). Clearly there exists a solution if \( g \in \text{Ran}(1 - K) \) and this solution is unique if \( \text{Ker}(1 - K) = \{0\} \). Hence these subspaces play a crucial role. Moreover, if the underlying Hilbert space is finite dimensional it is well-known that \( \text{Ker}(1 - K) = \{0\} \) automatically implies \( \text{Ran}(1 - K) = \mathcal{H} \) since

\[ \dim \text{Ker}(1 - K) + \dim \text{Ran}(1 - K) = \dim \mathcal{H}. \]  \hspace{1cm} (5.29)

Unfortunately this formula is of no use if \( \mathcal{H} \) is infinite dimensional, but if we rewrite it as

\[ \dim \text{Ker}(1 - K) = \dim \mathcal{H} - \dim \text{Ran}(1 - K) = \dim \text{Ran}(1 - K)\perp \]  \hspace{1cm} (5.30)

there is some hope. In fact, we will show that this formula (makes sense and) holds if \( K \) is a compact operator.

**Lemma 5.13.** Let \( K \in \mathfrak{C}(\mathcal{H}) \) be compact. Then \( \text{Ker}(1 - K) \) is finite dimensional and \( \text{Ran}(1 - K) \) is closed.

**Proof.** We first show \( \dim \text{Ker}(1 - K) < \infty \). If not we could find an infinite orthonormal system \( \{u_j\}_{j=1}^\infty \subset \text{Ker}(1 - K) \). By \( Ku_j = u_j \) compactness of \( K \) implies that there is a convergent subsequence \( u_{j_k} \). But this is impossible by \( \|u_j - u_k\|^2 = 2 \) for \( j \neq k \).

To see that \( \text{Ran}(1 - K) \) is closed we first claim that there is a \( \gamma > 0 \) such that

\[ \|(1 - K)f\| \geq \gamma \|f\|, \quad \forall f \in \text{Ker}(1 - K)\perp. \]  \hspace{1cm} (5.31)

In fact, if there were no such \( \gamma \), we could find a normalized sequence \( f_j \in \text{Ker}(1 - K)\perp \) with \( \|f_j - Kf_j\| < \frac{1}{j} \), that is, \( f_j - Kf_j \to 0 \). After passing to a subsequence we can assume \( Kf_j \to f \) by compactness of \( K \). Combining this with \( f_j - Kf_j \to 0 \) implies \( f_j \to f \) and \( f - Kf = 0 \), that is, \( f \in \text{Ker}(1 - K) \). On the other hand, since \( \text{Ker}(1 - K)\perp \) is closed, we also have \( f \in \text{Ker}(1 - K)\perp \) which shows \( f = 0 \). This contradicts \( \|f\| = \lim \|f_j\| = 1 \) and thus (5.31) holds.

Now choose a sequence \( g_j \in \text{Ran}(1 - K) \) converging to some \( g \). By assumption there are \( f_k \) such that \( (1 - K)f_k = g_k \) and we can even assume \( f_k \in \text{Ker}(1 - K)\perp \) by removing the projection onto \( \text{Ker}(1 - K) \). Hence (5.31) shows

\[ \|f_j - f_k\| \leq \gamma^{-1}\|1 - K)(f_j - f_k)\| = \gamma^{-1}\|g_j - g_k\| \]

that \( f_j \) converges to some \( f \) and \( (1 - K)f = g \) implies \( g \in \text{Ran}(1 - K) \). \( \square \)
Since
\[ \text{Ran}(1 - K)^\perp = \text{Ker}(1 - K^*) \] (5.32)
by (2.28) we see that the left and right hand side of (5.30) are at least finite for compact \( K \) and we can try to verify equality.

**Theorem 5.14.** Suppose \( K \) is compact. Then
\[ \dim \text{Ker}(1 - K) = \dim \text{Ran}(1 - K)^\perp, \] (5.33)
where both quantities are finite.

**Proof.** It suffices to show
\[ \dim \text{Ker}(1 - K) \geq \dim \text{Ran}(1 - K)^\perp, \] (5.34)
since replacing \( K \) by \( K^* \) in this inequality and invoking (2.28) provides the reversed inequality.

We begin by showing that \( \dim \text{Ker}(1 - K) = 0 \) implies \( \dim \text{Ran}(1 - K)^\perp = 0 \), that is \( \text{Ran}(1 - K) = \mathcal{S} \). To see this, suppose \( \mathcal{S}_1 = \text{Ran}(1 - K) = (1 - K)\mathcal{S} \) is not equal to \( \mathcal{S} \). Then \( \mathcal{S}_2 = (1 - K)\mathcal{S}_1 \) can also not be equal to \( \mathcal{S}_1 \). Otherwise we could choose \( f \in \mathcal{S}_1^\perp \) and since \( (1 - K)f \in \mathcal{S}_1 = (1 - K)\mathcal{S} \) we could find a \( g \in \mathcal{S}_2 \) with \( (1 - K)f = (1 - K)^2g \). By injectivity of \( 1 - K \) we must have \( f = (1 - K)g \in \mathcal{S}_1 \) contradicting our assumption. Proceeding inductively we obtain a sequence of subspaces \( \mathcal{S}_j = (1 - K)^j\mathcal{S} \) with \( \mathcal{S}_j \subset \mathcal{S}_{j+1} \). Now choose a normalized sequence \( f_j \in \mathcal{S}_j \cap \mathcal{S}_{j+1}^\perp \). Then for \( k > j \) we have
\[
\|Kf_j - Kf_k\|^2 = \|f_j - f_k - (1 - K)(f_j - f_k)\|^2 = \|f_j\|^2 + \|f_k + (1 - K)(f_j - f_k)\|^2 \geq 1
\]
since \( f_j \in \mathcal{S}_j^\perp \) and \( f_k + (1 - K)(f_j - f_k) \in \mathcal{S}_{j+1} \). But this contradicts the fact that \( Kf_j \) must have a convergent subsequence.

To show (5.34) in the general case, suppose \( \dim \text{Ker}(1 - K) < \dim \text{Ran}(1 - K)^\perp \) instead. Then we can find a bounded map \( A : \text{Ker}(1 - K) \to \text{Ran}(1 - K)^\perp \) which is injective but not onto. Extend \( A \) to a map on \( \mathcal{S} \) by setting \( Af = 0 \) for \( f \in \text{Ker}(1 - K)^\perp \). Since \( A \) is finite rank, the operator \( \hat{K} = K + A \) is again compact. We claim \( \text{Ker}(1 - \hat{K}) = \{0\} \). Indeed, if \( f \in \text{Ker}(1 - \hat{K}) \), then \( f - Kf = Af \in \text{Ran}(1 - K)^\perp \) implies \( f \in \text{Ker}(1 - K) \cap \text{Ker}(A) \). But \( A \) is injective on \( \text{Ker}(1 - K) \) and thus \( f = 0 \) as claimed. Thus the first step applied to \( \hat{K} \) implies \( \text{Ran}(1 - \hat{K}) = \mathcal{S} \). But this is impossible since the equation
\[
f - \hat{K}f = (1 - K)f - Af = g
\]
for \( g \in \text{Ran}(1 - K)^\perp \) reduces to \( (1 - K)f = 0 \) and \( Af = -g \) which has no solution if we choose \( g \not\in \text{Ran}(A) \).

As a special case we obtain the famous
Theorem 5.15 (Fredholm alternative). Suppose $K \in \mathcal{C}(\mathcal{H})$ is compact. Then either the inhomogeneous equation

$$f = Kf + g \quad (5.35)$$

has a unique solution for every $g \in \mathcal{H}$ or the corresponding homogeneous equation

$$f = Kf \quad (5.36)$$

has a nontrivial solution.

Note that (5.32) implies that in any case the inhomogeneous equation $f = Kf + g$ has a solution if and only if $g \in \text{Ker}(1 - K^*)^\perp$. Moreover, combining (5.33) with (5.32) also shows

$$\dim \text{Ker}(1 - K) = \dim \text{Ker}(1 - K^*) \quad (5.37)$$

for compact $K$. Furthermore, note that in the case where $\dim \text{Ker}(1 - K) = 0$ the solution is given by $(1 - K)^{-1}g$, where $(1 - K)^{-1}$ is bounded by the closed graph theorem (cf. Corollary 4.8).

This theory can be generalized to the case of operators where both $\text{Ker}(1 - K)$ and $\text{Ran}(1 - K)^\perp$ are finite dimensional. Such operators are called Fredholm operators (also Noether operators) and the number

$$\text{ind}(1 - K) = \dim \text{Ker}(1 - K) - \dim \text{Ran}(1 - K)^\perp \quad (5.38)$$

is the called the index of $K$. Theorem 5.14 now says that a compact operator is Fredholm of index zero.

Problem 5.5. Compute $\text{Ker}(1 - K)$ and $\text{Ran}(1 - K)^\perp$ for the operator $K = \langle v, \cdot \rangle u$, where $u, v \in \mathcal{H}$ satisfy $\langle u, v \rangle = 1$.

Problem 5.6. Let $M$ be multiplication by a sequence $m_j = \frac{1}{j}$ in the Hilbert space $\ell^2(\mathbb{N})$ and consider $K = MS^{-}$. Show that $K - z$ has a bounded inverse for every $z \in \mathbb{C}\setminus\{0\}$. Show that $K$ is invertible.
Chapter 6

Bounded linear operators

6.1. Banach algebras

In this section we want to have a closer look at the set of bounded linear operators \( \mathcal{L}(X) \) from a Banach space \( X \) into itself. We already know from Section 1.5 that they form a Banach space which has a multiplication given by composition. In this section we want to further investigate this structure.

A Banach space \( X \) together with a multiplication satisfying

\[
(x + y)z = xz + yz, \quad x(y + z) = xy + xz, \quad x, y, z \in X,
\]

and

\[
(xy)z = x(yz), \quad \alpha(xy) = (\alpha x)y = x(\alpha y), \quad \alpha \in \mathbb{C}.
\]

and

\[
\|xy\| \leq \|x\|\|y\|.
\]

is called a Banach algebra. In particular, note that (6.3) ensures that multiplication is continuous (Problem 6.1). An element \( e \in X \) satisfying

\[
ex = xe = x, \quad \forall x \in X
\]

is called identity (show that \( e \) is unique) and we will assume \( \|e\| = 1 \) in this case.

**Example.** The continuous functions \( C(I) \) over some compact interval form a commutative Banach algebra with identity 1.

\[\diamond\]

**Example.** The bounded linear operators \( \mathcal{L}(X) \) form a Banach algebra with identity \( I \).

\[\diamond\]
Example. The space $L^1(\mathbb{R}^n)$ together with the convolution
\[
(g * f)(x) = \int_{\mathbb{R}^n} g(x-y)f(y)dy = \int_{\mathbb{R}^n} g(y)f(x-y)dy
\] (6.5)
is a commutative Banach algebra (Problem 6.5) without identity. ♦

Let $X$ be a Banach algebra with identity $e$. Then $x \in X$ is called \textit{invertible} if there is some $y \in X$ such that
\[
xy = yx = e.
\] (6.6)
In this case $y$ is called the inverse of $x$ and is denoted by $x^{-1}$. It is straightforward to show that the inverse is unique (if one exists at all) and that
\[
(xy)^{-1} = y^{-1}x^{-1}.
\] (6.7)

Example. Let $X = \mathcal{L}(\ell^1(\mathbb{N}))$ and let $S^\pm$ be defined via
\[
S^-x_n = \begin{cases} 
0 & n = 1 \\
x_{n-1} & n > 1
\end{cases}, \quad S^+x_n = x_{n+1}
\] (6.8)
(i.e., $S^-$ shifts each sequence one place right (filling up the first place with a 0) and $S^+$ shifts one place left (dropping the first place)). Then $S^+S^- = I$ but $S^-S^+ \neq I$. So you really need to check both $xy = e$ and $yx = e$ in general. ♦

Lemma 6.1. Let $X$ be a Banach algebra with identity $e$. Suppose $\|x\| < 1$. Then $e - x$ is invertible and
\[
(e - x)^{-1} = \sum_{n=0}^{\infty} x^n.
\] (6.9)

Proof. Since $\|x\| < 1$ the series converges and
\[
(e - x) \sum_{n=0}^{\infty} x^n = \sum_{n=0}^{\infty} x^n - \sum_{n=1}^{\infty} x^n = e
\]
respectively
\[
\left( \sum_{n=0}^{\infty} x^n \right)(e - x) = \sum_{n=0}^{\infty} x^n - \sum_{n=1}^{\infty} x^n = e.
\]
\[\square\]

Corollary 6.2. Suppose $x$ is invertible and $\|x^{-1}y\| < 1$ or $\|yx^{-1}\| < 1$. Then $(x - y)$ is invertible as well and
\[
(x - y)^{-1} = \sum_{n=0}^{\infty} (x^{-1}y)^n x^{-1} \quad \text{or} \quad (x - y)^{-1} = \sum_{n=0}^{\infty} x^{-1}(yx^{-1})^n.
\] (6.10)

In particular, both conditions are satisfied if $\|y\| < \|x^{-1}\|^{-1}$ and the set of invertible elements is open.
Proof. Just observe $x - y = x(e - x^{-1}y) = (e - yx^{-1})x$. \hfill \Box

This last corollary implies that the set of invertible elements is open and that taking inverses is continuous, that is, if $x_n \to x$ and $x$ is invertible, then $x_n$ is also invertible for $n$ sufficiently large and $x_n^{-1} \to x^{-1}$.

The resolvent set is defined as
\begin{equation}
\rho(x) = \{ \alpha \in \mathbb{C} | \exists (x - \alpha)^{-1} \} \subseteq \mathbb{C},
\end{equation}
where we have used the shorthand notation $x - \alpha = x - \alpha e$. Its complement is called the spectrum
\begin{equation}
\sigma(x) = \mathbb{C} \setminus \rho(x).
\end{equation}

It is important to observe that the fact that the inverse has to exist as an element of $X$. That is, if $X$ are bounded linear operators, it does not suffice that $x - \alpha$ is bijective, the inverse must also be bounded!

Example. If $X = \mathcal{L}(\mathbb{C}^n)$ is the space of $n$ by $n$ matrices, then the spectrum is just the set of eigenvalues. \diamond

Example. If $X = C(I)$, then the spectrum of a function $x \in C(I)$ is just its range, $\sigma(x) = x(I)$. \diamond

The map $\alpha \mapsto (x - \alpha)^{-1}$ is called the resolvent of $x \in X$. If $\alpha_0 \in \rho(x)$ we can choose $x \to x - \alpha_0$ and $y \to \alpha - \alpha_0$ in (6.10) which implies
\begin{equation}
(x - \alpha)^{-1} = \sum_{n=0}^{\infty} (\alpha - \alpha_0)^n (x - \alpha_0)^{-n-1}, \quad |\alpha - \alpha_0| < \|(x - \alpha_0)^{-1}\|^{-1}. \tag{6.13}
\end{equation}

In particularly, since the radius of convergence cannot reach the spectrum (since everything within the radius of convergent must belong to the resolvent set), we see that the norm of the resolvent must diverge
\begin{equation}
\|(x - \alpha)^{-1}\| \geq \frac{1}{\text{dist}(\alpha, \sigma(x))} \tag{6.14}
\end{equation}
as $\alpha$ approaches the spectrum. Moreover, this shows that $(x - \alpha)^{-1}$ has a convergent power series with coefficients in $X$ around every point $\alpha_0 \in \rho(x)$. As in the case of coefficients in $\mathbb{C}$, such functions will be called analytic. In particular, $\ell((x - \alpha)^{-1})$ is a complex-valued analytic function for every $\ell \in X^*$ and we can apply well-known results from complex analysis:

Theorem 6.3. For every $x \in X$, the spectrum $\sigma(x)$ is compact, nonempty and satisfies
\begin{equation}
\sigma(x) \subseteq \{ \alpha | |\alpha| \leq \|x\| \}. \tag{6.15}
\end{equation}
Proof. Equation (6.13) already shows that $\rho(x)$ is open. Hence $\sigma(x)$ is closed. Moreover, $x - \alpha = -\alpha(e - \frac{1}{\alpha}x)$ together with Lemma 6.1 shows

$$
(x - \alpha)^{-1} = -\frac{1}{\alpha} \sum_{n=0}^{\infty} \left( \frac{1}{x} \right)^n, \quad |\alpha| > \|x\|,
$$

which implies $\sigma(x) \subseteq \{ \alpha \mid |\alpha| \leq \|x\| \}$ is bounded and thus compact. Moreover, taking norms shows

$$
\| (x - \alpha)^{-1} \| \leq \frac{1}{|\alpha|} \sum_{n=0}^{\infty} \frac{\|x\|^n}{|\alpha|^n} = \frac{1}{|\alpha| - \|x\|}, \quad |\alpha| > \|x\|,
$$

which implies $(x - \alpha)^{-1} \to 0$ as $\alpha \to \infty$. In particular, if $\sigma(x)$ is empty, then $\ell((x - \alpha)^{-1})$ is an entire analytic function which vanishes at infinity. By Liouville’s theorem we must have $\ell((x - \alpha)^{-1}) = 0$ in this case, and so $(x - \alpha)^{-1} = 0$, which is impossible. □

As another simple consequence we obtain:

**Theorem 6.4** (Gelfand–Mazur). Suppose $X$ is a Banach algebra in which every element except 0 is invertible. Then $X$ is isomorphic to $\mathbb{C}$.

**Proof.** Pick $x \in X$ and $\alpha \in \sigma(x)$. Then $x - \alpha$ is not invertible and hence $x - \alpha = 0$, that is $x = \alpha$. Thus every element is a multiple of the identity. □

Given a polynomial $p(\alpha) = \sum_{j=0}^{n} p_j \alpha^j$ we set

$$
p(x) = \sum_{j=0}^{n} p_j x^j.
$$

(6.16)

Then we have the following result:

**Theorem 6.5** (Spectral mapping). For every polynomial $p$ and $x \in X$ we have

$$
\sigma(p(x)) = p(\sigma(x)),
$$

where $p(\sigma(x)) = \{ p(\alpha) \mid \alpha \in \sigma(x) \}$.

**Proof.** Fix $\alpha_0 \in \mathbb{C}$ and observe

$$
p(x) - p(\alpha_0) = (x - \alpha_0)q_0(x).
$$

If $p(\alpha_0) \notin \sigma(p(x))$ we have

$$
(x - \alpha_0)^{-1} = q_0(x)((x - \alpha_0)q_0(x))^{-1} = ((x - \alpha_0)q_0(x))^{-1}q_0(x)
$$

(check this — since $q_0(x)$ commutes with $(x - \alpha_0)q_0(x)$ it also commutes with its inverse). Hence $\alpha_0 \notin \sigma(x)$.

Conversely, let $\alpha_0 \in \sigma(p(x))$. Then

$$
p(x) - \alpha_0 = a(x - \lambda_1) \cdots (x - \lambda_n)
$$
and at least one \( \lambda_j \in \sigma(x) \) since otherwise the right-hand side would be invertible. But then \( p(\lambda_j) = \alpha_0 \), that is, \( \alpha_0 \in p(\sigma(x)) \). \( \square \)

Next let us look at the convergence radius of the Neumann series for the resolvent
\[
(x - \alpha)^{-1} = -\frac{1}{\alpha} \sum_{n=0}^{\infty} \left( \frac{x}{\alpha} \right)^n
\]
(6.18)
encountered in the proof of Theorem 6.3 (which is just the Laurent expansion around infinity).

The number
\[
r(x) = \sup_{\alpha \in \sigma(x)} |\alpha|
\]
(6.19)
is called the spectral radius of \( x \). Note that by (6.15) we have
\[
r(x) \leq \|x\|.
\]
(6.20)

**Theorem 6.6.** The spectral radius satisfies
\[
r(x) = \inf_{n \in \mathbb{N}} \|x^n\|^{1/n} = \lim_{n \to \infty} \|x^n\|^{1/n}.
\]
(6.21)

**Proof.** By spectral mapping we have \( r(x)^n = r(x^n) \leq \|x^n\| \) and hence
\[
r(x) \leq \inf \|x^n\|^{1/n}.
\]
Conversely, fix \( \ell \in X^* \), and consider
\[
\ell((x - \alpha)^{-1}) = -\frac{1}{\alpha} \sum_{n=0}^{\infty} \frac{1}{\alpha^n} \ell(x^n).
\]
(6.22)
Then \( \ell((x - \alpha)^{-1}) \) is analytic in \( |\alpha| > r(x) \) and hence (6.22) converges absolutely for \( |\alpha| > r(x) \) by Cauchy’s integral formula for derivatives. Hence for fixed \( \alpha \) with \( |\alpha| > r(x) \), \( \ell(x^n/\alpha^n) \) converges to zero for every \( \ell \in X^* \). Since every weakly convergent sequence is bounded we have
\[
\frac{\|x^n\|}{|\alpha|^n} \leq C(\alpha)
\]
and thus
\[
\limsup_{n \to \infty} \|x^n\|^{1/n} \leq \limsup_{n \to \infty} C(\alpha)^{1/n} |\alpha| = |\alpha|.
\]
Since this holds for every \( |\alpha| > r(x) \) we have
\[
r(x) \leq \inf \|x^n\|^{1/n} \leq \liminf_{n \to \infty} \|x^n\|^{1/n} \leq \limsup_{n \to \infty} \|x^n\|^{1/n} \leq r(x),
\]
which finishes the proof. \( \square \)
To end this section let us look at two examples illustrating these ideas.

**Example.** Let \( X = \mathcal{L}(\mathbb{C}^2) \) be the space of two by two matrices and consider
\[
x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.
\] (6.23)

Then \( x^2 = 0 \) and consequently \( r(x) = 0 \). This is not surprising, since \( x \) has the only eigenvalue 0. In particular, the spectral radius can be strictly smaller than the norm (note that \( \|x\| = 1 \) in our example). The same is true for any nilpotent matrix.

**Example.** Consider the linear Volterra integral operator
\[
K(x)(t) = \int_0^t k(t,s)x(s)ds, \quad x \in C([0,1]),
\] (6.24)
then, using induction, it is not hard to verify (Problem 6.4)
\[
|K^nx(t)| \leq \frac{\|k\|_\infty t^n}{n!}\|x\|_\infty.
\] (6.25)

Consequently
\[
\|K^n x\|_\infty \leq \frac{\|k\|_\infty^n}{n!}\|x\|_\infty,
\]
that is \( \|K^n\| \leq \frac{\|k\|_\infty^n}{n!} \), which shows
\[
r(K) \leq \lim_{n \to \infty} \frac{\|k\|_\infty^n}{(n!)^{1/n}} = 0.
\]

Hence \( r(K) = 0 \) and for every \( \lambda \in \mathbb{C} \) and every \( y \in C(I) \) the equation
\[
x - \lambda K x = y
\] (6.26)
has a unique solution given by
\[
x = (I - \lambda K)^{-1}y = \sum_{n=0}^{\infty} \lambda^n K^n y.
\] (6.27)

**Problem 6.1.** Show that the multiplication in a Banach algebra \( X \) is continuous: \( x_n \to x \) and \( y_n \to y \) imply \( x_ny_n \to xy \).

**Problem 6.2.** Show \( \sigma(x^{-1}) = \sigma(x)^{-1} \) if \( x \) is invertible.

**Problem 6.3.** Suppose \( x \) has both a right inverse \( y \) (i.e., \( xy = e \)) and a left inverse \( z \) (i.e., \( zx = e \)). Show that \( y = z = x^{-1} \).

**Problem 6.4.** Show (6.25).

**Problem 6.5.** Show that \( L^1(\mathbb{R}^n) \) with convolution as multiplication is a commutative Banach algebra without identity (Hint: Lemma 8.14).
Problem 6.6. Show the first resolvent identity
\[(x - \alpha)^{-1} - (x - \beta)^{-1} = (\alpha - \beta)(x - \alpha)^{-1}(x - \beta)^{-1} = (\alpha - \beta)(x - \beta)^{-1}(x - \alpha)^{-1}, \tag{6.28}\]
for \(\alpha, \beta \in \rho(x)\).

6.2. The \(C^*\) algebra of operators and the spectral theorem

We begin by recalling that if \(\mathcal{H}\) is some Hilbert space, then for every \(A \in \mathcal{L}(\mathcal{H})\) we can define its adjoint \(A^* \in \mathcal{L}(\mathcal{H})\). Hence the Banach algebra \(\mathcal{L}(\mathcal{H})\) has an additional operation in this case. In general, a Banach algebra \(X\) together with an involution
\[(x + y)^* = x^* + y^*, \quad (\alpha x)^* = \alpha^* x^*, \quad x^{**} = x, \quad (xy)^* = y^* x^*, \tag{6.29}\]
satisfying
\[\|x\|^2 = \|x^* x\| \tag{6.30}\]
is called a \(C^*\) algebra. Any subalgebra (we do not require a subalgebra to contain the identity) which is also closed under involution, is called a \(*\)-subalgebra. Note that (6.30) implies \(\|x\|^2 \leq \|x^*\| \|x\|\) and hence \(\|x\| = \|x^*\|\).

Example. The continuous functions \(C(I)\) together with complex conjugation form a commutative \(C^*\) algebra.

Example. The Banach algebra \(\mathcal{L}(\mathcal{H})\) is a \(C^*\) algebra by Lemma 2.13.

If \(X\) has an identity \(e\), we clearly have \(e^* = e\), \((x^{-1})^* = (x^*)^{-1}\) (show this), and
\[\sigma(x^*) = \sigma(x)^*. \tag{6.32}\]

We will always assume that we have an identity. In fact, if \(X\) has no identity we can always add one by considering \(\tilde{X} = X \oplus \text{span} e\) and define the product via \((x + \alpha e)(y + \beta e) = xy + \alpha y + \beta x + (\alpha \beta) e\).

If \(X\) is a \(C^*\) algebra, then \(x \in X\) is called normal if \(x^* x = xx^*\), self-adjoint if \(x^* = x\), and unitary if \(x^* = x^{-1}\). Moreover, \(x\) is called positive if \(x = y^2\) for some \(y = y^* \in X\). Clearly both self-adjoint and unitary elements are normal and positive elements are self-adjoint.

If \(x\) is normal (self-adjoint), then so is any polynomial \(p(x)\). Moreover, if \(x\) is self-adjoint, then (6.30) implies \(\|x^2\| = \|x\|^2\). This even holds for normal elements. For unitary elements we have \(\|x\| = \sqrt{\|x^* x\|} = \sqrt{\|e\|} = 1\).

Lemma 6.7. If \(x \in X\) is normal, then \(\|x^2\| = \|x\|^2\) and \(r(x) = \|x\|\).
Proof. Using (6.30) three times we have
\[ \|x^2\| = \|(x^2)^*(x^2)\|^{1/2} = \|(xx^*)^*(xx^*)\|^{1/2} = \|x^*x\| = \|x\|^2 \]
and hence \( r(x) = \lim_{k \to \infty} \|x^2k\|^{1/2k} = \|x\| \). \qed

Example. If \( X = \mathcal{L}(\mathbb{C}^2) \) and \( x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \) such that \( x^2 = 0 \), then \( 0 = \|x^2\| \neq \|x\|^2 = 1 \). Hence the above result does not hold for arbitrary elements. \( \Diamond \)

Lemma 6.8. If \( x \) is self-adjoint, then \( \sigma(x) \subseteq \mathbb{R} \). If \( x \) is positive, then \( \sigma(x) \subseteq [0, \infty) \).

Proof. Suppose \( \alpha + i\beta \in \sigma(x) \), \( \lambda \in \mathbb{R} \). Then \( \alpha + i(\beta + \lambda) \in \sigma(x + i\lambda) \) and
\[ \alpha^2 + (\beta + \lambda)^2 \leq \|x + i\lambda\|^2 = \|(x + i\lambda)(x - i\lambda)\| = \|x^2 + \lambda^2\| \leq \|x\|^2 + \lambda^2. \]
Hence \( \alpha^2 + \beta^2 + 2\beta\lambda \leq \|x\|^2 \) which gives a contradiction if we let \( |\lambda| \to \infty \) unless \( \beta = 0 \).

The second claim follows from the first using spectral mapping (Theorem 6.5). \( \Box \)

Example. If \( X = \mathcal{L}(\mathbb{C}^2) \) and \( x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \) then \( \sigma(x) = \{0\} \). Hence the converse of the above lemma is not true in general. \( \Diamond \)

Given \( x \in X \) we can consider the \( C^* \) algebra \( C^*(x) \) (with identity) generated by \( x \) (i.e., the smallest closed \( * \)-subalgebra containing \( e \) and \( x \)). If \( x \) is normal we explicitly have
\[ C^*(x) = \{ p(x, x^*) \mid p : \mathbb{C}^2 \to \mathbb{C} \text{ polynomial} \}, \quad xx^* = x^*x, \quad (6.33) \]
and, in particular, \( C^*(x) \) is commutative (Problem 6.8). In the self-adjoint case this simplifies to
\[ C^*(x) = \{ p(x) \mid p : \mathbb{C} \to \mathbb{C} \text{ polynomial} \}, \quad x = x^*. \quad (6.34) \]
Moreover, in this case \( C^*(x) \) is isomorphic to \( C(\sigma(x)) \) (the continuous functions on the spectrum).

Theorem 6.9 (Spectral theorem). If \( X \) is a \( C^* \) algebra and \( x \) is self-adjoint, then there is an isometric isomorphism \( \Phi : C(\sigma(x)) \to C^*(x) \) such that \( f(t) = t \) maps to \( \Phi(t) = x \) and \( f(t) = 1 \) maps to \( \Phi(1) = e \).

Moreover, for every \( f \in C(\sigma(x)) \) we have
\[ \sigma(f(x)) = f(\sigma(x)), \quad (6.35) \]
where \( f(x) = \Phi(f(t)) \).
Proof. First of all, $\Phi$ is well defined for polynomials $p$ and given by $\Phi(p) = p(x)$. Moreover, since $p(x)$ is normal spectral mapping implies

$$\|p(x)\| = r(p(x)) = \sup_{\alpha \in \sigma(p(x))} |\alpha| = \sup_{\alpha \in \sigma(x)} |p(\alpha)| = \|p\|_{\infty}$$

for every polynomial $p$. Hence $\Phi$ is isometric. Since the polynomials are dense by the Stone–Weierstraß theorem (see the next section) $\Phi$ uniquely extends to a map on all of $C(\sigma(x))$ by Theorem 1.35. By continuity of the norm this extension is again isometric. Similarly, we have $\Phi(fg) = \Phi(f)\Phi(g)$ and $\Phi(f)^* = \Phi(f^*)$ since both relations hold for polynomials.

To show $\sigma(f(x)) = f(\sigma(x))$ fix some $\alpha \in \mathbb{C}$. If $\alpha \notin f(\sigma(x))$, then $g(t) = \frac{1}{f(t) - \alpha} \in C(\sigma(x))$ and $\Phi(g) = (f(x) - \alpha)^{-1} \in X$ shows $\alpha \notin \sigma(f(x))$. Conversely, if $\alpha \notin \sigma(f(x))$ then $g = \Phi^{-1}((f(x) - \alpha)^{-1}) = \frac{1}{f - \alpha}$ is continuous, which shows $\alpha \notin f(\sigma(x))$. □

In particular, this last theorem tells us that we have a functional calculus for self-adjoint operators, that is, if $A \in \mathcal{L}(\mathcal{H})$ is self-adjoint, then $f(A)$ is well defined for every $f \in C(\sigma(A))$. If $f$ is given by a power series, $f(A)$ defined via $\Phi$ coincides with $f(A)$ defined via its power series.

Problem 6.7. Let $X$ be a $C^*$ algebra and $Y$ a $*$-subalgebra. Show that if $Y$ is commutative, then so is $\overline{Y}$.

Problem 6.8. Show that the map $\Phi$ from the spectral theorem is positivity preserving, that is, $f \geq 0$ if and only if $\Phi(f)$ is positive.

Problem 6.9. Let $x$ be self-adjoint. Show that $\sigma(x) \subseteq \{\alpha \in \mathbb{R} | \alpha \geq 0\}$ if and only if $x$ is positive.

Problem 6.10. Let $A \in \mathcal{L}(\mathcal{H})$. Show that $A$ is normal if and only if

$$\|Au\| = \|A^*u\|, \quad \forall u \in \mathcal{H}.$$  \hfill (6.36)

(Hint: Problem 1.26.)

Problem 6.11. Show that the Cayley transform of a self-adjoint element $x$,

$$y = (x - i)(x + i)^{-1}$$

is unitary. Show that $1 \notin \sigma(y)$ and

$$x = i(1 + y)(1 - y)^{-1}.$$  

Problem 6.12. Show if $x$ is unitary then $\sigma(x) \subseteq \{\alpha \in \mathbb{C} | ||\alpha|| = 1\}$.

Problem 6.13. Suppose $x$ is self-adjoint. Show that

$$\|(x - \alpha)^{-1}\| = \frac{1}{\text{dist}(\alpha, \sigma(x))}.$$
6.3. Spectral measures

Note: This section requires familiarity with measure theory.

Using the Riesz representation theorem we get another formulation in terms of spectral measures:

**Theorem 6.10.** Let \( \mathcal{H} \) be a Hilbert space, and let \( A \in \mathfrak{L}(\mathcal{H}) \) be self-adjoint. For every \( u, v \in \mathcal{H} \) there is a corresponding complex Borel measure \( \mu_{u,v} \) supported on \( \sigma(A) \) (the *spectral measure*) such that

\[
\langle u, f(A)v \rangle = \int_{\sigma(A)} f(t) d\mu_{u,v}(t), \quad f \in C(\sigma(A)).
\]  

(6.37)

We have

\[
\mu_{u,v_1+v_2} = \mu_{u,v_1} + \mu_{u,v_2}, \quad \mu_{u,\alpha v} = \alpha \mu_{u,v}, \quad \mu_{v,u} = \mu_{u,v}^* \quad \text{(6.38)}
\]

and \( |\mu_{u,v}(\sigma(A))| \leq \|u\| \|v\| \). Furthermore, \( \mu_u = \mu_{u,u} \) is a positive Borel measure with \( \mu_u(\sigma(A)) = \|u\|^2 \).

**Proof.** Consider the continuous functions on \( I = [-\|A\|, \|A\|] \) and note that every \( f \in C(I) \) gives rise to some \( f \in C(\sigma(A)) \) by restricting its domain. Clearly \( \ell_{u,v}(f) = \langle u, f(A)v \rangle \) is a bounded linear functional and the existence of a corresponding measure \( \mu_{u,v} \) with \( |\mu_{u,v}(I)| = \|\ell_{u,v}\| \leq \|u\| \|v\| \) follows from Theorem 10.4. Since \( \ell_{u,v}(f) \) depends only on the value of \( f \) on \( \sigma(A) \subseteq I \), \( \mu_{u,v} \) is supported on \( \sigma(A) \).

Moreover, if \( f \geq 0 \) we have \( \ell_u(f) = \langle u, f(A)u \rangle = \langle f(A)^{1/2}u, f(A)^{1/2}u \rangle = \|f(A)^{1/2}u\|^2 \geq 0 \) and hence \( \ell_u \) is positive and the corresponding measure \( \mu_u \) is positive. The rest follows from the properties of the scalar product. \( \square \)

It is often convenient to regard \( \mu_{u,v} \) as a complex measure on \( \mathbb{R} \) by using \( \mu_{u,v}(\Omega) = \mu_{u,v}(\Omega \cap \sigma(A)) \). If we do this, we can also consider \( f \) as a function on \( \mathbb{R} \). However, note that \( f(A) \) depends only on the values of \( f \) on \( \sigma(A) \)!

Moreover, it suffices to consider \( \mu_u \) since using the polarization identity (1.58) we have

\[
\mu_{u,v}(\Omega) = \frac{1}{4}(\mu_{u+v}(\Omega) - \mu_{u-v}(\Omega) + i\mu_{u-iv}(\Omega) - i\mu_{u+iv}(\Omega)).
\]  

(6.39)

Now the last theorem can be used to define \( f(A) \) for every bounded measurable function \( f \in B(\sigma(A)) \) via Lemma 2.11 and extend the functional calculus from continuous to measurable functions:

**Theorem 6.11** (Spectral theorem). If \( \mathcal{H} \) is a Hilbert space and \( A \in \mathfrak{L}(\mathcal{H}) \) is self-adjoint, then there is an homomorphism \( \Phi : B(\sigma(A)) \to \mathfrak{L}(\mathcal{H}) \) given by

\[
\langle u, f(A)v \rangle = \int_{\sigma(A)} f(t) d\mu_{u,v}(t), \quad f \in B(\sigma(A)).
\]  

(6.40)
Moreover, if \( f_n(t) \to f(t) \) pointwise and \( \sup_n \| f_n \|_\infty \) is bounded, then \( f_n(A)u \to f(A)u \) for every \( u \in \mathcal{H} \).

**Proof.** The map \( \Phi \) is a well-defined linear operator by Lemma 2.11 since we have

\[
\left| \int_{\sigma(A)} f(t) d\mu_{u,v}(t) \right| \leq \| f \|_\infty |\mu_{u,v}|(\sigma(A)) \leq \| f \|_\infty \| u \| \| v \|
\]

and (6.38). Next, observe that \( \Phi(f^*) = \Phi(f)^* \) and \( \Phi(fg) = \Phi(f)\Phi(g) \) holds at least for continuous functions. To obtain it for arbitrary bounded functions, choose a (bounded) sequence \( f_n \) converging to \( f \) in \( L^2(\sigma(A), d\mu_u) \) and observe

\[
\left\| (f_n(A) - f(A))u \right\|^2 = \int |f_n(t) - f(t)|^2 d\mu_u(t)
\]

(use \( \| h(A)u \|^2 = \langle h(A)u, h(A)u \rangle = \langle u, h(A)^*h(A)u \rangle \)). Thus \( f_n(A)u \to f(A)u \) and for bounded \( g \) we also have that \( (gf_n)(A)u \to (gf)(A)u \) and \( g(A)f_n(A)u \to g(A)f(A)u \). This establishes the case where \( f \) is bounded and \( g \) is continuous. Similarly, approximating \( g \) removes the continuity requirement from \( g \).

The last claim follows since \( f_n \to f \) in \( L^2 \) by dominated convergence in this case. \( \square \)

Our final aim is to generalize Corollary 3.9 to bounded self-adjoint operators. Since the spectrum of an arbitrary self-adjoint might contain more than just eigenvalues we need to replace the sum by an integral. To this end we begin by defining the **spectral projections**

\[
P_A(\Omega) = \chi_\Omega(A), \quad \Omega \in \mathcal{B} (\mathbb{R})
\]

such that

\[
\mu_{u,v}(\Omega) = \langle u, P_A(\Omega)v \rangle.
\]

By \( \chi_\Omega^2 = \chi_\Omega \) and \( \chi_\Omega^* = \chi_\Omega \) they are **orthogonal projections**, that is \( P^2 = P \) and \( P^* = P \). Recall that any orthogonal projection \( P \) decomposes \( \mathcal{H} \) decomposes into an orthogonal sum

\[
\mathcal{H} = \text{Ker}(P) \oplus \text{Ran}(P)
\]

where \( \text{Ker}(P) = (I - P)\mathcal{H}, \text{Ran}(P) = P\mathcal{H} \).

In addition, the spectral projections satisfy

\[
P_A(\mathbb{R}) = I, \quad P_A(\bigcup_{n=1}^\infty \Omega_n)u = \sum_{n=1}^\infty P_A(\Omega_n)u, \quad \Omega_n \cap \Omega_m = \emptyset, m \neq n,
\]

for every \( u \in \mathcal{H} \). Such a family of projections is called a **projection-valued measure**. Indeed the first claim follows since \( \chi_\mathbb{R} = 1 \) and by \( \chi_{\Omega_1 \cup \Omega_2} = \chi_{\Omega_1} + \chi_{\Omega_2} \) if \( \Omega_1 \cap \Omega_2 = \emptyset \) the second claim follows at least for finite unions.
The case of countable unions follows from the last part of the previous theorem since \( \sum_{n=1}^{N} \chi_{\Omega_n} = \chi_{\bigcup_{n=1}^{N} \Omega_n} \rightarrow \chi_{\bigcup_{n=1}^{\infty} \Omega_n} \) pointwise (note that the limit will not be uniform unless the \( \Omega_n \) are eventually empty and hence there is no chance that this series will converge in the operator norm). In fact, since all spectral measures are supported on \( \sigma(A) \) the same is true for \( P_A \) in the sense that
\[
P_A(\sigma(A)) = \mathbb{I}. \tag{6.45}
\]
I also remark that in this connection the corresponding distribution function
\[
P_A(t) = P_A((-\infty, t]) \tag{6.46}
\]
is called a **resolution of the identity**.

Using our projection-valued measure we can define an operator-valued integral as follows: For every simple function \( f = \sum_{j=1}^{n} \alpha_j \chi_{\Omega_j} \) (where \( \Omega_j = f^{-1}(\alpha_j) \)), we set
\[
\int_{\mathbb{R}} f(t) dP_A(t) = \sum_{j=1}^{n} \alpha_j P_A(\Omega_j) \tag{6.47}
\]
By (6.42) we conclude that this definition agrees with \( f(A) \) from Theorem 6.11:
\[
\int_{\mathbb{R}} f(t) dP_A(t) = f(A). \tag{6.48}
\]
Extending this integral to functions from \( B(\sigma(A)) \) by approximating such functions with simple functions we get an alternative way of defining \( f(A) \) for such functions. This can in fact be done by just using the definition of a projection-valued measure and hence there is a one-to-one correspondence between projection-valued measures (with bounded support) and (bounded) self-adjoint operators such that
\[
A = \int t \, dP_A(t). \tag{6.49}
\]
If \( P_A(\{\alpha\}) \neq 0 \), then \( \alpha \) is an eigenvalue and \( \text{Ran}(P_A(\{\alpha\})) \) is the corresponding eigenspace (Problem 6.15). The fact that eigenspaces to different eigenvalues are orthogonal now generalizes to

**Lemma 6.12.** Suppose \( \Omega_1 \cap \Omega_2 = \emptyset \). Then
\[
\text{Ran}(P_A(\Omega_1)) \perp \text{Ran}(P_A(\Omega_2)). \tag{6.50}
\]

**Proof.** Clearly \( \chi_{\Omega_1} \chi_{\Omega_2} = \chi_{\Omega_1 \cap \Omega_2} \) and hence
\[
P_A(\Omega_1) P_A(\Omega_2) = P_A(\Omega_1 \cap \Omega_2).
\]
Now if \( \Omega_1 \cap \Omega_2 = \emptyset \), then
\[
\langle P_A(\Omega_1)u, P_A(\Omega_2)v \rangle = \langle u, P_A(\Omega_1) P_A(\Omega_2) v \rangle = \langle u, P_A(\emptyset) v \rangle = 0,
\]
which shows that the ranges are orthogonal to each other. \( \square \)
Example. Let $A \in \mathfrak{L}(\mathbb{C}^n)$ be some symmetric matrix and let $\alpha_1, \ldots, \alpha_m$ be its (distinct) eigenvalues. Then

$$A = \sum_{j=1}^{m} \alpha_j P_A(\{ \alpha_j \}),$$

where

$$P_A(\{ \alpha_j \}) = \text{Ker}(A - \alpha_j)$$

is the projection onto the eigenspace corresponding to the eigenvalue $\alpha_j$ by Problem 6.15. In fact, using that $P_A$ is supported on the spectrum, $P_A(\sigma(A)) = 1$, we see

$$P(\Omega) = P_A(\sigma(A))P(\Omega) = P(\sigma(A) \cap \Omega) = \sum_{\alpha_j \in \Omega} P_A(\{ \alpha_j \}).$$

Hence using that any $f \in B(\sigma(A))$ is given as a simple function $f = \sum_{j=1}^{m} f(\alpha_j) \chi_{\{ \alpha_j \}}$ we obtain

$$f(A) = \int f(t) dP_A(t) = \sum_{j=1}^{m} f(\alpha_j) P_A(\{ \alpha_j \}).$$

In particular, for $f(t) = t$ we recover the above representation for $A$. ◦

Problem 6.14. Suppose $A$ is self-adjoint. Let $\alpha$ be an eigenvalue and $u$ a corresponding normalized eigenvector. Show $\int f(t) d\mu_u(t) = f(\alpha)$, that is, $\mu_u$ is the Dirac delta measure (with mass one) centered at $\alpha$.

Problem 6.15. Suppose $A$ is self-adjoint. Show

$$\text{Ran}(P_A(\{ \alpha \})) = \text{Ker}(A - \alpha).$$

(Hint: Start by verifying $\text{Ran}(P_A(\{ \alpha \})) \subseteq \text{Ker}(A - \alpha)$. To see the converse, let $u \in \text{Ker}(A - \alpha)$ and use the previous example.)

6.4. The Stone–Weierstraß theorem

In the last section we have seen that the $C^*$ algebra of continuous functions $C(K)$ over some compact set $K \subseteq \mathbb{C}$ plays a crucial role and that it is important to be able to identify dense sets. We will be slightly more general and assume that $K$ is some compact metric space. Then it is straightforward to check that the same proof as in the case $K = [a, b]$ (Section 1.2) shows that $C(K, \mathbb{R})$ and $C(K, \mathbb{C})$ are Banach spaces when equipped with the maximum norm $\| f \|_\infty = \max_{x \in K} | f(x) |$.

Theorem 6.13 (Stone–Weierstraß, real version). Suppose $K$ is a compact metric space and let $C(K, \mathbb{R})$ be the Banach algebra of continuous functions (with the maximum norm).
If $F \subset C(K, \mathbb{R})$ contains the identity 1 and separates points (i.e., for every $x_1 \neq x_2$ there is some function $f \in F$ such that $f(x_1) \neq f(x_2)$), then the algebra generated by $F$ is dense.

**Proof.** Denote by $A$ the algebra generated by $F$. Note that if $f \in A$, we have $|f| \in A$. By the Weierstraß approximation theorem (Theorem 1.26) there is a polynomial $p_n(t)$ such that $||t| - p_n(t)|| < \frac{1}{n}$ for $t \in f(K)$ and hence $p_n(f) \to |f|$.

In particular, if $f, g$ are in $A$, we also have

$$\max\{f, g\} = \frac{(f + g) + |f - g|}{2}, \quad \min\{f, g\} = \frac{(f + g) - |f - g|}{2}$$

in $A$.

Now fix $f \in C(K, \mathbb{R})$. We need to find some $f^\varepsilon \in A$ with $\|f - f^\varepsilon\|_\infty < \varepsilon$.

First of all, since $A$ separates points, observe that for given $y, z \in K$ there is a function $f_{y, z} \in A$ such that $f_{y, z}(y) = f(y)$ and $f_{y, z}(z) = f(z)$ (show this). Next, for every $y \in K$ there is a neighborhood $U(y)$ such that $f_{y, z}(x) > f(x) - \varepsilon$, $x \in U(y)$, and since $K$ is compact, finitely many, say $U(y_1), \ldots, U(y_j)$, cover $K$. Then

$$f_z = \max\{f_{y_1, z}, \ldots, f_{y_j, z}\} \in A$$

and satisfies $f_z > f - \varepsilon$ by construction. Since $f_z(z) = f(z)$ for every $z \in K$, there is a neighborhood $V(z)$ such that

$$f_z(x) < f(x) + \varepsilon, \quad x \in V(z),$$

and a corresponding finite cover $V(z_1), \ldots, V(z_k)$. Now

$$f^\varepsilon = \min\{f_{z_1}, \ldots, f_{z_k}\} \in A$$

satisfies $f^\varepsilon < f + \varepsilon$. Since $f - \varepsilon < f_{z_1}$ we also have $f - \varepsilon < f^\varepsilon$ and we have found a required function. □

**Theorem 6.14** (Stone–Weierstraß). Suppose $K$ is a compact metric space and let $C(K)$ be the $C^*$ algebra of continuous functions (with the maximum norm).

If $F \subset C(K)$ contains the identity 1 and separates points, then the $*$-subalgebra generated by $F$ is dense.

**Proof.** Just observe that $\tilde{F} = \{\text{Re}(f), \text{Im}(f)\} | f \in F\}$ satisfies the assumption of the real version. Hence every real-valued continuous function can be approximated by elements from the subalgebra generated by $\tilde{F}$; in particular, this holds for the real and imaginary parts for every given complex-valued function. Finally, note that the subalgebra spanned by $\tilde{F}$ contains the $*$-subalgebra spanned by $F$. □
Note that the additional requirement of being closed under complex conjugation is crucial: The functions holomorphic on the unit ball and continuous on the boundary separate points, but they are not dense (since the uniform limit of holomorphic functions is again holomorphic).

**Corollary 6.15.** Suppose $K$ is a compact metric space and let $C(K)$ be the $C^*$ algebra of continuous functions (with the maximum norm).

If $F \subset C(K)$ separates points, then the closure of the $*$-subalgebra generated by $F$ is either $C(K)$ or $\{ f \in C(K) | f(t_0) = 0 \}$ for some $t_0 \in K$.

**Proof.** There are two possibilities: either all $f \in F$ vanish at one point $t_0 \in K$ (there can be at most one such point since $F$ separates points) or there is no such point.

If there is no such point, then the identity can be approximated by elements in $\overline{A}$: First of all note that $|f| \in \overline{A}$ if $f \in \overline{A}$, since the polynomials $p_n(t)$ used to prove this fact can be replaced by $p_n(t) - p_n(0)$ which contain no constant term. Hence for every point $y$ we can find a nonnegative function in $\overline{A}$ which is positive at $y$ and by compactness we can find a finite sum of such functions which is positive everywhere, say $m \leq f(t) \leq M$. Now approximate $\min(m^{-1}t, t^{-1})$ by polynomials $q_n(t)$ (again a constant term is not needed) to conclude that $q_n(f) \to f^{-1} \in \overline{A}$. Hence $1 = f : f^{-1} \in \overline{A}$ as claimed and so $\overline{A} = C(K)$ by the Stone–Weierstraß theorem.

If there is such a $t_0$ we have $\overline{A} \subseteq \{ f \in C(K) | f(t_0) = 0 \}$ and the identity is clearly missing from $\overline{A}$. However, adding the identity to $\overline{A}$ we get $\overline{A} + \mathbb{C} = C(K)$ by the Stone–Weierstraß theorem. Moreover, if $f \in C(K)$ with $f(t_0) = 0$ we get $f = \tilde{f} + \alpha$ with $\tilde{f} \in \overline{A}$ and $\alpha \in \mathbb{C}$. But $0 = f(t_0) = \tilde{f}(t_0) + \alpha = \alpha$ implies $f = \tilde{f} \in \overline{A}$, that is, $\overline{A} = \{ f \in C(K) | f(t_0) = 0 \}$.

**Problem 6.16.** Show that the functions $\varphi_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}$, $n \in \mathbb{Z}$, form an orthonormal basis for $\mathcal{S} = L^2(0, 2\pi)$. (Hint: Start with $K = [0, 2\pi]$ where $0$ and $\pi$ are identified.)

**Problem 6.17.** Let $k \in \mathbb{N}$ and $I \subseteq \mathbb{R}$. Show that the $*$-subalgebra generated by $f_{z_0}(t) = \frac{1}{(t - z_0)^k}$ for one $z_0 \in \mathbb{C}$ and $k \in \mathbb{N}$ is dense in the $C^*$ algebra $C_0(I)$ of continuous functions vanishing at infinity:

- for $I = \mathbb{R}$ if $z_0 \in \mathbb{C} \setminus \mathbb{R}$ and $k = 1$ or $k = 2$,
- for $I = [a, \infty)$ if $z_0 \in (-\infty, a)$ and $k$ arbitrary,
- for $I = (-\infty, a) \cup [b, \infty)$ if $z_0 \in (a, b)$ and $k$ odd.

(Hint: Add $\infty$ to $\mathbb{R}$ to make it compact.)

**Problem 6.18.** Let $K \subseteq \mathbb{C}$ be a compact set. Show that the set of all functions $f(z) = p(x, y)$, where $p : \mathbb{R}^2 \to \mathbb{C}$ is polynomial and $z = x + iy$, is dense in $C(K)$. 

Part 2

Real Analysis
Almost everything about Lebesgue integration

7.1. Borel measures in a nut shell

The first step in defining the Lebesgue integral is extending the notion of size from intervals to arbitrary sets. Unfortunately, this turns out to be too much, since a classical paradox by Banach and Tarski shows that one can break the unit ball in \( \mathbb{R}^3 \) into a finite number of (wild – choosing the pieces uses the Axiom of Choice and cannot be done with a jigsaw;-) pieces, rotate and translate them, and reassemble them to get two copies of the unit ball (compare Problem 7.5). Hence any reasonable notion of size (i.e., one which is translation and rotation invariant) cannot be defined for all sets!

A collection of subsets \( \mathcal{A} \) of a given set \( X \) such that

- \( X \in \mathcal{A} \),
- \( \mathcal{A} \) is closed under finite unions,
- \( \mathcal{A} \) is closed under complements

is called an algebra. Note that \( \emptyset \in \mathcal{A} \) and that \( \mathcal{A} \) is also closed under finite intersections and relative complements: \( \emptyset = X' \), \( A \cap B = (A' \cup B')' \) (de Morgan), and \( A \setminus B = A \cap B' \), where \( A' = X \setminus A \) denotes the complement. If an algebra is closed under countable unions (and hence also countable intersections), it is called a \( \sigma \)-algebra.

Example. Let \( X = \{1, 2, 3\} \), then \( \mathcal{A} = \{\emptyset, \{1\}, \{2, 3\}, X\} \) is an algebra.
Moreover, the intersection of any family of \((\sigma\)-algebras \(\{A_\alpha\}\) is again a \((\sigma\)-algebra and for any collection \(S\) of subsets there is a unique smallest \((\sigma\)-algebra \(\Sigma(S)\) containing \(S\) (namely the intersection of all \((\sigma\)-algebras containing \(S\)). It is called the \((\sigma\)-algebra generated by \(S\).

**Example.** For a given set \(X\) the power set \(\mathcal{P}(X)\) is clearly the largest \(\sigma\)-algebra and \(\{\emptyset, X\}\) is the smallest.

**Example.** Let \(X\) be some set with a \(\sigma\)-algebra \(\Sigma\). Then every subset \(Y \subseteq X\) has a natural \(\sigma\)-algebra \(\Sigma \cap Y\) (show that this is indeed a \(\sigma\)-algebra) known as the relative \(\sigma\)-algebra.

Note that if \(S\) generates \(\Sigma\), then \(S \cap Y\) generates \(\Sigma \cap Y\): \(\Sigma(S) \cap Y = \Sigma(S \cap Y)\). Indeed, since \(\Sigma \cap Y\) is a \(\sigma\)-algebra containing \(S \cap Y\), we have \(\Sigma(S \cap Y) \subseteq \Sigma(S) \cap Y = \Sigma \cap Y\). Conversely, consider \(\{A \in \Sigma | A \cap Y \in \Sigma(S \cap Y)\}\) which is a \(\sigma\)-algebra (check this). Since this last \(\sigma\)-algebra contains \(S\) it must be equal to \(\Sigma = \Sigma(S)\) and thus \(\Sigma \cap Y \subseteq \Sigma(S \cap Y)\).

If \(X\) is a topological space, the Borel \(\sigma\)-algebra \(\mathcal{B}(X)\) of \(X\) is defined to be the \(\sigma\)-algebra generated by all open (respectively, all closed) sets. In fact, if \(X\) is second countable, any countable base will suffice to generate the Borel \(\sigma\)-algebra (recall Lemma 1.1).

Sets in the Borel \(\sigma\)-algebra are called Borel sets.

**Example.** In the case \(X = \mathbb{R}^n\) the Borel \(\sigma\)-algebra will be denoted by \(\mathcal{B}^n\) and we will abbreviate \(\mathcal{B} = \mathcal{B}^1\). Note that in order to generate \(\mathcal{B}\), open (or closed) intervals with rational boundary points suffice.

**Example.** If \(X\) is a topological space, then any Borel set \(Y \subseteq X\) is also a topological space equipped with the relative topology and its Borel \(\sigma\)-algebra is given by \(\mathcal{B}(Y) = \mathcal{B}(X) \cap Y = \{A | A \in \mathcal{B}(X), A \subseteq Y\}\) (show this).

Now let us turn to the definition of a measure: A set \(X\) together with a \(\sigma\)-algebra \(\Sigma\) is called a measurable space. A measure \(\mu\) is a map \(\mu : \Sigma \to [0, \infty]\) on a \(\sigma\)-algebra \(\Sigma\) such that

- \(\mu(\emptyset) = 0\),
- \(\mu(\bigcup_{j=1}^\infty A_j) = \sum_{j=1}^\infty \mu(A_j)\) if \(A_j \cap A_k = \emptyset\) for all \(j \neq k\) (\(\sigma\)-additivity).

Here the sum is set equal to \(\infty\) if one of the summands is \(\infty\) or if it diverges.

The measure \(\mu\) is called \(\sigma\)-finite if there is a countable cover \(\{X_j\}_{j=1}^\infty\) of \(X\) such that \(X_j \in \Sigma\) and \(\mu(X_j) < \infty\) for all \(j\). (Note that it is no restriction to assume \(X_j \subseteq X_{j+1}\).) It is called finite if \(\mu(X) < \infty\) and a probability measure if \(\mu(X) = 1\). The sets in \(\Sigma\) are called measurable sets and the triple \((X, \Sigma, \mu)\) is referred to as a measure space.
Example. Take a set $X$ and $\Sigma = \mathcal{P}(X)$ and set $\mu(A)$ to be the number of elements of $A$ (respectively, $\infty$ if $A$ is infinite). This is the so-called counting measure. It will be finite if and only if $X$ is finite and $\sigma$-finite if and only if $X$ is countable.

Example. Take a set $X$ and $\Sigma = \mathcal{P}(X)$. Fix a point $x \in X$ and set $\mu(A) = 1$ if $x \in A$ and $\mu(A) = 0$ else. This is the Dirac measure centered at $x$.

Example. Let $\mu_1, \mu_2$ be two measures on $(X, \Sigma)$ and $\alpha_1, \alpha_2 \geq 0$. Then $\mu = \alpha_1 \mu_1 + \alpha_2 \mu_2$ defined via

$$\mu(A) = \alpha_1 \mu_1(A) + \alpha_2 \mu_2(A)$$

is again a measure. Furthermore, given a countable number of measures $\mu_n$ and numbers $\alpha_n \geq 0$, then $\mu = \sum_n \alpha_n \mu_n$ is again a measure (show this).

Example. Let $\mu$ be a measure on $(X, \Sigma)$ and $Y \subseteq X$ a measurable subset. Then

$$\nu(A) = \mu(A \cap Y)$$

is again a measure on $(X, \Sigma)$ (show this).

Example. If $Y \in \Sigma$ we can restrict the $\sigma$-algebra $\Sigma|_Y = \{ A \in \Sigma | A \subseteq Y \}$ such that $(Y, \Sigma|_Y, \mu|_Y)$ is again a measurable space. It will be $\sigma$-finite if $(X, \Sigma, \mu)$ is.

If we replace the $\sigma$-algebra by an algebra $\mathcal{A}$, then $\mu$ is called a premeasure. In this case $\sigma$-additivity clearly only needs to hold for disjoint sets $A_n$ for which $\bigcup_n A_n \in \mathcal{A}$.

We will write $A_n \not\subset A$ if $A_n \subseteq A_{n+1}$ with $A = \bigcup_n A_n$ and $A_n \not\subset A$ if $A_{n+1} \subseteq A_n$ with $A = \bigcap_n A_n$.

**Theorem 7.1.** Any measure $\mu$ satisfies the following properties:

1. $A \subseteq B$ implies $\mu(A) \leq \mu(B)$ (monotonicity).
2. $\mu(A_n) \rightarrow \mu(A)$ if $A_n \not\subset A$ (continuity from below).
3. $\mu(A_n) \rightarrow \mu(A)$ if $A_n \not\subset A$ and $\mu(A_1) < \infty$ (continuity from above).

**Proof.** The first claim is obvious from $\mu(B) = \mu(A) + \mu(B \setminus A)$. To see the second define $\tilde{A}_1 = A_1$, $\tilde{A}_n = A_n \setminus A_{n-1}$ and note that these sets are disjoint and satisfy $A_n = \bigcup_{j=1}^{n} \tilde{A}_j$. Hence $\mu(A_n) = \sum_{j=1}^{n} \mu(\tilde{A}_j) \rightarrow \sum_{j=1}^{\infty} \mu(\tilde{A}_j) = \mu(\bigcup_{j=1}^{\infty} \tilde{A}_j) = \mu(A)$ by $\sigma$-additivity. The third follows from the second using $\tilde{A}_n = A_1 \setminus A_n \not\subset A_1 \setminus A$ implying $\mu(\tilde{A}_n) = \mu(A_1) - \mu(A_n) \rightarrow \mu(A_1 \setminus A) = \mu(A_1) - \mu(A)$. \qed
**Example.** Consider the counting measure on $X = \mathbb{N}$ and let $A_n = \{j \in \mathbb{N} | j \geq n\}$, then $\mu(A_n) = \infty$, but $\mu(\bigcap_n A_n) = \mu(\emptyset) = 0$ which shows that the requirement $\mu(A_1) < \infty$ in item (iii) of Theorem 7.1 is not superfluous.

A measure on the Borel $\sigma$-algebra is called a **Borel measure** if $\mu(K) < \infty$ for every compact set $K$. Note that some authors do not require this last condition.

**Example.** Let $X = \mathbb{R}$ and $\Sigma = \mathfrak{B}$. The Dirac measure is a Borel measure. The counting measure is no Borel measure since $\mu([a, b]) = \infty$ for $a < b$.

A measure on the Borel $\sigma$-algebra is called **outer regular** if

$$\mu(A) = \inf_{O \supseteq A, O \text{ open}} \mu(O)$$

and **inner regular** if

$$\mu(A) = \sup_{K \subseteq A, K \text{ compact}} \mu(K).$$

It is called **regular** if it is both outer and inner regular.

**Example.** Let $X = \mathbb{R}$ and $\Sigma = \mathfrak{B}$. The counting measure is inner regular but not outer regular (every nonempty open set has infinite measure). The Dirac measure is a regular Borel measure.

But how can we obtain some more interesting Borel measures? We will restrict ourselves to the case of $X = \mathbb{R}$ for simplicity, in which case Borel measures are also known as **Lebesgue–Stieltjes measures**. Then the strategy is as follows: Start with the algebra of finite unions of disjoint intervals and define $\mu$ for those sets (as the sum over the intervals). This yields a premeasure. Extend this to an **outer measure** for all subsets of $\mathbb{R}$. Show that the restriction to the Borel sets is a measure.

Let us first show how we should define $\mu$ for intervals: To every Borel measure on $\mathfrak{B}$ we can assign its **distribution function**

$$\mu(x) = \begin{cases} -\mu((x, 0]], & x < 0, \\ 0, & x = 0, \\ \mu((0, x]], & x > 0, \end{cases}$$

which is right continuous and nondecreasing.

**Example.** The distribution function of the Dirac measure centered at 0 is

$$\mu(x) = \begin{cases} -1, & x < 0, \\ 0, & x \geq 0. \end{cases}$$
For a finite measure the alternate normalization \( \tilde{\mu}(x) = \mu((-\infty, x]) \) can be used. The resulting distribution function differs from our above definition by a constant \( \mu(x) = \tilde{\mu}(x) - \mu((-\infty, 0]) \). In particular, this is the normalization used in probability theory.

Conversely, to obtain a measure from a nondecreasing function \( m : \mathbb{R} \to \mathbb{R} \) we proceed as follows: Recall that an interval is a subset of the real line of the form

\[
I = (a, b], \quad I = [a, b], \quad I = (a, b), \quad \text{or} \quad I = [a, b),
\]

with \( a \leq b, \ a, b \in \mathbb{R} \cup \{-\infty, \infty\} \). Note that \((a, a), [a, a), \) and \((a, a)\) denote the empty set, whereas \([a, a]\) denotes the singleton \{a\}. For any proper interval with different endpoints (i.e. \( a < b \)) we can define its measure to be

\[
\mu(I) = \begin{cases} 
m(b+) - m(a+), & I = (a, b], 
m(b+) - m(a-), & I = [a, b], 
m(b-) - m(a+), & I = (a, b), 
m(b-) - m(a-), & I = [a, b),
\end{cases}
\]

(7.5)

where \( m(a\pm) = \lim_{\varepsilon \to 0} m(a \pm \varepsilon) \) (which exist by monotonicity). If one of the endpoints is infinite we agree to use \( m(\pm\infty) = \lim_{x \to \pm\infty} m(x) \). For the empty set we of course set \( \mu(\emptyset) = 0 \) and for the singletons we set

\[
\mu\{a\} = m(a+) - m(a-)
\]

(7.6)

(which agrees with (7.5) except for the case \( I = (a, a) \) which would give a negative value for the empty set if \( \mu \) jumps at \( a \)). Note that \( \mu\{a\} = 0 \) if and only if \( m(x) \) is continuous at \( a \) and that there can be only countably many points with \( \mu\{a\} > 0 \) since a nondecreasing function can have at most countably many jumps. Moreover, observe that the definition of \( \mu \) does not involve the actual value of \( m \) at a jump. Hence any function \( \tilde{m} \) with \( m(x-) \leq \tilde{m}(x) \leq m(x+) \) gives rise to the same \( \mu \). We will frequently assume that \( m \) is right continuous such that it coincides with the distribution function up to a constant, \( \mu(x) = m(x+) - m(0+) \). In particular, \( \mu \) determines \( m \) up to a constant and the value at the jumps.

Now we can consider the algebra of finite unions of disjoint intervals (check that this is indeed an algebra) and extend (7.5) to finite unions of disjoint intervals by summing over all intervals. It is straightforward to verify that \( \mu \) is well defined (one set can be represented by different unions of intervals) and by construction additive. In fact, it is even a premeasure.

**Lemma 7.2.** The interval function \( \mu \) defined in (7.5) gives rise to a unique \( \sigma \)-finite premeasure on the algebra \( \mathcal{A} \) of finite unions of disjoint intervals.
Proof. It remains to verify $\sigma$-additivity. We need to show that for any disjoint union

$$\mu(\bigcup_k I_k) = \sum_k \mu(I_k)$$

whenever $I_k \in \mathcal{A}$ and $I = \bigcup_k I_k \in \mathcal{A}$. Since each $I_k$ is a finite union of intervals, we can as well assume each $I_k$ is just one interval (just split $I_k$ into its subintervals and note that the sum does not change by additivity). Similarly, we can assume that $I$ is just one interval (just treat each subinterval separately).

By additivity $\mu$ is monotone and hence

$$\sum_{k=1}^{n} \mu(I_k) = \mu(\bigcup_{k=1}^{n} I_k) \leq \mu(I)$$

which shows

$$\sum_{k=1}^{\infty} \mu(I_k) \leq \mu(I).$$

To get the converse inequality, we need to work harder:

We can cover each $I_k$ by some slightly larger open interval $J_k$ such that $\mu(J_k) \leq \mu(I_k) + \frac{\varepsilon}{2^k}$ (only closed endpoints need extension). First suppose $I$ is compact. Then finitely many of the $J_k$, say the first $n$, cover $I$ and we have

$$\mu(I) \leq \mu\left(\bigcup_{k=1}^{n} J_k\right) \leq \sum_{k=1}^{n} \mu(J_k) \leq \sum_{k=1}^{\infty} \mu(I_k) + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this shows $\sigma$-additivity for compact intervals. By additivity we can always add/subtract the endpoints of $I$ and hence $\sigma$-additivity holds for any bounded interval. If $I$ is unbounded we can again assume that it is closed by adding an endpoint if necessary. Then for any $x > 0$ we can find an $n$ such that $\{J_k\}_{k=1}^{n}$ cover at least $I \cap [-x, x]$ and hence

$$\sum_{k=1}^{\infty} \mu(I_k) \geq \sum_{k=1}^{n} \mu(I_k) \geq \sum_{k=1}^{n} \mu(J_k) \geq \mu([-x, x] \cap I) - \varepsilon.$$

Since $x > 0$ and $\varepsilon > 0$ are arbitrary, we are done. \qed

In particular, this is a premeasure on the algebra of finite unions of intervals which can be extended to a measure:

**Theorem 7.3.** For every nondecreasing function $m : \mathbb{R} \to \mathbb{R}$ there exists a unique Borel measure $\mu$ which extends (7.5). Two different functions generate the same measure if and only if the difference is a constant away from the discontinuities.
Since the proof of this theorem is rather involved, we defer it to the next section and look at some examples first.

**Example.** Suppose \( \Theta(x) = 0 \) for \( x < 0 \) and \( \Theta(x) = 1 \) for \( x \geq 0 \). Then we obtain the so-called **Dirac measure** at 0, which is given by \( \Theta(A) = 1 \) if \( 0 \in A \) and \( \Theta(A) = 0 \) if \( 0 \not\in A \).

**Example.** Suppose \( \lambda(x) = x \). Then the associated measure is the ordinary **Lebesgue measure** on \( \mathbb{R} \). We will abbreviate the Lebesgue measure of a Borel set \( A \) by \( \lambda(A) = |A| \).

A set \( A \in \Sigma \) is called a **support** for \( \mu \) if \( \mu(X \setminus A) = 0 \). Note that a support is not unique (see the examples below). If \( X \) is a topological space and \( \Sigma = \mathcal{B}(X) \), one defines the **support** (also **topological support**) of \( \mu \) via

\[
\text{supp}(\mu) = \{ x \in X | \mu(O) > 0 \text{ for every open neighborhood } O \text{ of } x \}. \tag{7.7}
\]

Equivalently one obtains \( \text{supp}(\mu) \) by removing all points which have an open neighborhood of measure zero. In particular, this shows that \( \text{supp}(\mu) \) is closed. If \( X \) is second countable, then \( \text{supp}(\mu) \) is indeed a support for \( \mu \): For every point \( x \not\in \text{supp}(\mu) \) let \( O_x \) be an open neighborhood of measure zero. These sets cover \( X \setminus \text{supp}(\mu) \) and by the Lindelöf theorem there is a countable subcover, which shows that \( X \setminus \text{supp}(\mu) \) has measure zero.

**Example.** Let \( X = \mathbb{R}, \Sigma = \mathcal{B} \). The support of the Lebesgue measure \( \lambda \) is all of \( \mathbb{R} \). However, every single point has Lebesgue measure zero and so has every countable union of points (by \( \sigma \)-additivity). Hence any set whose complement is countable is a support. There are even uncountable sets of Lebesgue measure zero (see the Cantor set below) and hence a support might even lack an uncountable number of points.

The support of the Dirac measure centered at 0 is the single point 0. Any set containing 0 is a support of the Dirac measure.

In general, the support of a Borel measure on \( \mathbb{R} \) is given by

\[
\text{supp}(d\mu) = \{ x \in \mathbb{R} | \mu(x - \varepsilon) < \mu(x + \varepsilon), \forall \varepsilon > 0 \}.
\]

Here we have used \( d\mu \) to emphasize that we are interested in the support of the measure \( d\mu \) which is different from the support of its distribution function \( \mu(x) \).

A property is said to hold \( \mu \)-**almost everywhere** (a.e.) if it holds on a support for \( \mu \) or, equivalently, if the set where it does not hold is contained in a set of measure zero.

**Example.** The set of rational numbers is countable and hence has Lebesgue measure zero, \( \lambda(\mathbb{Q}) = 0 \). So, for example, the characteristic function of the rationals \( \mathbb{Q} \) is zero almost everywhere with respect to Lebesgue measure.
Any function which vanishes at 0 is zero almost everywhere with respect

to the Dirac measure centered at 0.

Example. The Cantor set is an example of a closed uncountable set of

Lebesgue measure zero. It is constructed as follows: Start with $C_0 = [0,1]$ and

remove the middle third to obtain $C_1 = [0,\frac{1}{3}] \cup [\frac{2}{3},1]$. Next, again remove

the middle third’s of the remaining sets to obtain $C_2 = [0,\frac{1}{9}] \cup [\frac{2}{9},\frac{1}{3}] \cup [\frac{2}{3},\frac{7}{9}] \cup [\frac{8}{9},1]$:

\[
\begin{array}{cccc}
\hline
\text{C} & \text{C} & \text{C} & \text{C} \\
\hline
C_0 & C_1 & C_2 & \vdots \\
\hline
\end{array}
\]

Proceeding like this, we obtain a sequence of nesting sets $C_n$ and the limit $C = \bigcap_n C_n$ is the Cantor set. Since $C_n$ is compact, so is $C$. Moreover, $C_n$ consists of $2^n$ intervals of length $3^{-n}$, and thus its Lebesgue measure is $\lambda(C_n) = (2/3)^n$. In particular, $\lambda(C) = \lim_{n \to \infty} \lambda(C_n) = 0$. Using the ternary expansion, it is extremely simple to describe: $C$ is the set of all $x \in [0,1]$ whose ternary expansion contains no one’s, which shows that $C$ is uncountable (why?). It has some further interesting properties: it is totally disconnected (i.e., it contains no subintervals) and perfect (it has no isolated points).

Problem 7.1. Find all algebras over $X = \{1,2,3\}$.

Problem 7.2. Show that $\mathcal{A} = \{A \subseteq X | A \text{ or } X \setminus A \text{ is finite}\}$ is an algebra (with $X$ some fixed set). Show that $\Sigma = \{A \subseteq X | A \text{ or } X \setminus A \text{ is countable}\}$ is a $\sigma$-algebra. (Hint: To verify closedness under unions consider the cases were all sets are finite and where one set has finite complement.)

Problem 7.3. Take some set $X$ and $\Sigma = \{A \subseteq X | A \text{ or } X \setminus A \text{ is countable}\}$. Show that

$$\nu(A) = \begin{cases} 0, & \text{if } A \text{ is countable}, \\ 1, & \text{else}. \end{cases}$$

is a measure.

Problem 7.4. Show that if $X$ is finite, then every algebra is a $\sigma$-algebra. Show that this is not true in general if $X$ is countable.

Problem 7.5 (Vitali set). Call two numbers $x,y \in [0,1)$ equivalent if $x - y$ is rational. Construct the set $V$ by choosing one representative from each equivalence class. Show that $V$ cannot be measurable with respect to any nontrivial finite translation invariant measure on $[0,1)$. (Hint: How can you build up $[0,1)$ from translations of $V$?)
7.2. Extending a premeasure to a measure

The purpose of this section is to prove Theorem 7.3. It is rather technical and can be skipped on first reading.

In order to prove Theorem 7.3, we need to show how a premeasure can be extended to a measure. To show that the extension is unique we need a better criterion to check when a given system of sets is in fact a $\sigma$-algebra. In many situations it is easy to show that a given set is closed under complements and under countable unions of disjoint sets. Hence we call a collection of sets $D$ with these properties a Dynkin system (also $\lambda$-system) if it also contains $X$.

Note that a Dynkin system is closed under proper relative complements since $A, B \in D$ implies $B \setminus A = (B' \cup A)' \in D$ provided $A \subseteq B$. Moreover, if it is also closed under finite intersections (or arbitrary finite unions) then it is an algebra and hence also a $\sigma$-algebra. To see the last claim note that if $A = \bigcup_j A_j$ then also $A = \bigcup_j B_j$ where the sets $B_j = A_j \setminus \bigcup_{k<j} A_k$ are disjoint.

As with $\sigma$-algebras, the intersection of Dynkin systems is a Dynkin system and every collection of sets $S$ generates a smallest Dynkin system $D(S)$. The important observation is that if $S$ is closed under finite intersections (in which case it is sometimes called a $\pi$-system), then so is $D(S)$ and hence will be a $\sigma$-algebra.

**Lemma 7.4** (Dynkin’s $\pi$-$\lambda$ theorem). Let $S$ be a collection of subsets of $X$ which is closed under finite intersections (or unions). Then $D(S) = \Sigma(S)$.

**Proof.** It suffices to show that $D = D(S)$ is closed under finite intersections. To this end consider the set $D(A) = \{ B \in D | A \cap B \in D \}$ for $A \in D$. I claim that $D(A)$ is a Dynkin system.

First of all $X \in D(A)$ since $A \cap X = A \in D$. Next, if $B \in D(A)$ then $A \cap B' = A \setminus (B \cap A) \in D$ (since $D$ is closed under proper relative complements) implying $B' \in D(A)$. Finally if $B = \bigcup_j B_j$ with $B_j \in D(A)$ disjoint, then $A \cap B = \bigcup_j (A \cap B_j) \in D$ with $B_j \in D$ disjoint, implying $B \in D(A)$.

Now if $A \in S$ we have $S \subseteq D(A)$ implying $D(A) = D$. Consequently $A \cap B \in D$ if at least one of the sets is in $S$. But this shows $S \subseteq D(A)$ and hence $D(A) = D$ for every $A \in D$. So $D$ is closed under finite intersections and thus a $\sigma$-algebra. The case of unions is analogous.  

The typical use of this lemma is as follows: First verify some property for sets in a set $S$ which is closed under finite intersections and generates
the \( \sigma \)-algebra. In order to show that it holds for every set in \( \Sigma(S) \), it suffices to show that the collection of sets for which it holds is a Dynkin system. As an application we show that a premeasure determines the corresponding measure \( \mu \) uniquely (if there is one at all):

**Theorem 7.5 (Uniqueness of measures).** Let \( S \subseteq \Sigma \) be a collection of sets which generates \( \Sigma \) and which is closed under finite intersections and contains a sequence of increasing sets \( X_n \uparrow X \) of finite measure \( \mu(X_n) < \infty \). Then \( \mu \) is uniquely determined by the values on \( S \).

**Proof.** Let \( \tilde{\mu} \) be a second measure and note \( \mu(X) = \lim_{n \to \infty} \mu(X_n) = \lim_{n \to \infty} \tilde{\mu}(X_n) = \tilde{\mu}(X) \). We first suppose \( \mu(X) < \infty \).

Then
\[
\mathcal{D} = \{ A \in \Sigma | \mu(A) = \tilde{\mu}(A) \}
\]
is a Dynkin system. In fact, by \( \mu(A') = \mu(X) - \mu(A) = \tilde{\mu}(X) - \tilde{\mu}(A) = \tilde{\mu}(A') \) for \( A \in \mathcal{D} \) we see that \( \mathcal{D} \) is closed under complements. Furthermore, by continuity of measures from below it is also closed under countable disjoint unions. Since \( \mathcal{D} \) contains \( S \) by assumption, we conclude \( \mathcal{D} = \Sigma(S) = \Sigma \) from Lemma 7.4. This finishes the finite case.

To extend our result to the general case observe that the finite case implies \( \mu(A \cap X_j) = \tilde{\mu}(A \cap X_j) \) (just restrict \( \mu, \tilde{\mu} \) to \( X_j \)). Hence
\[
\mu(A) = \lim_{j \to \infty} \mu(A \cap X_j) = \lim_{j \to \infty} \tilde{\mu}(A \cap X_j) = \tilde{\mu}(A)
\]
and we are done. \( \square \)

**Corollary 7.6.** Let \( \mu \) be a \( \sigma \)-finite premeasure on an algebra \( \mathcal{A} \). Then there is at most one extension to \( \Sigma(\mathcal{A}) \).

So it remains to ensure that there is an extension at all. For any premeasure \( \mu \) we define
\[
\mu^*(A) = \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) \left| A \subseteq \bigcup_{n=1}^{\infty} A_n, A_n \in \mathcal{A} \right. \right\} \quad (7.8)
\]
where the infimum extends over all countable covers from \( \mathcal{A} \). Then the function \( \mu^* : \mathcal{P}(X) \to [0, \infty] \) is an **outer measure**; that is, it has the properties (Problem 7.6)
- \( \mu^*(\emptyset) = 0 \),
- \( A_1 \subseteq A_2 \Rightarrow \mu^*(A_1) \leq \mu^*(A_2) \), and
- \( \mu^*(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu^*(A_n) \) (subadditivity).

Note that \( \mu^*(A) = \mu(A) \) for \( A \in \mathcal{A} \) (Problem 7.7).
Theorem 7.7 (Extensions via outer measures). Let $\mu^*$ be an outer measure. Then the set $\Sigma$ of all sets $A$ satisfying the Carathéodory condition

$$\mu^*(E) = \mu^*(A \cap E) + \mu^*(A' \cap E), \quad \forall E \subseteq X$$

(7.9)

(where $A' = X \setminus A$ is the complement of $A$) forms a $\sigma$-algebra and $\mu^*$ restricted to this $\sigma$-algebra is a measure.

Proof. We first show that $\Sigma$ is an algebra. It clearly contains $X$ and is closed under complements. Concerning unions let $A, B \in \Sigma$. Applying Carathéodory’s condition twice shows

$$\mu^*(E) = \mu^*(A \cap B \cap E) + \mu^*(A' \cap B \cap E) + \mu^*(A \cap B' \cap E) + \mu^*(A' \cap B' \cap E) \geq \mu^*((A \cup B) \cap E) + \mu^*((A \cup B)' \cap E),$$

where we have used de Morgan and

$$\mu^*(A \cap B \cap E) + \mu^*(A' \cap B \cap E) + \mu^*(A \cap B' \cap E) \geq \mu^*((A \cup B) \cap E)$$

which follows from subadditivity and $(A \cup B) \cap E = (A \cap B \cap E) \cup (A' \cap B \cap E) \cup (A \cap B' \cap E)$. Since the reverse inequality is just subadditivity, we conclude that $\Sigma$ is an algebra.

Next, let $A_n$ be a sequence of sets from $\Sigma$. Without restriction we can assume that they are disjoint (compare the argument for item (ii) in the proof of Theorem 7.1). Abbreviate $\hat{A}_n = \bigcup_{k \leq n} A_k$, $A = \bigcup_{n} A_n$. Then for every set $E$ we have

$$\mu^*(\hat{A}_n \cap E) = \mu^*(A_n \cap \hat{A}_n \cap E) + \mu^*(A'_n \cap \hat{A}_n \cap E)$$

$$= \mu^*(A_n \cap E) + \mu^*(\hat{A}_{n-1} \cap E)$$

$$= \ldots = \sum_{k=1}^{n} \mu^*(A_k \cap E).$$

Using $\hat{A}_n \in \Sigma$ and monotonicity of $\mu^*$, we infer

$$\mu^*(E) = \mu^*(\hat{A}_n \cap E) + \mu^*(\hat{A}_n' \cap E)$$

$$\geq \sum_{k=1}^{n} \mu^*(A_k \cap E) + \mu^*(A'_n \cap E).$$

Letting $n \to \infty$ and using subadditivity finally gives

$$\mu^*(E) \geq \sum_{k=1}^{\infty} \mu^*(A_k \cap E) + \mu^*(A'_n \cap E)$$

$$\geq \mu^*(A \cap E) + \mu^*(A' \cap E) \geq \mu^*(E)$$

(7.10)

and we infer that $\Sigma$ is a $\sigma$-algebra.
Finally, setting $E = A$ in (7.10), we have
\[
\mu^*(A) = \sum_{k=1}^{\infty} \mu^*(A_k \cap A) + \mu^*(A' \cap A) = \sum_{k=1}^{\infty} \mu^*(A_k)
\]
and we are done. \qed

Remark: The constructed measure $\mu$ is complete; that is, for every measurable set $A$ of measure zero, every subset of $A$ is again measurable (Problem 7.8).

The only remaining question is whether there are any nontrivial sets satisfying the Carathéodory condition.

**Lemma 7.8.** Let $\mu$ be a premeasure on $\mathcal{A}$ and let $\mu^*$ be the associated outer measure. Then every set in $\mathcal{A}$ satisfies the Carathéodory condition.

**Proof.** Let $A_n \in \mathcal{A}$ be a countable cover for $E$. Then for every $A \in \mathcal{A}$ we have
\[
\sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} \mu(A_n \cap A) + \sum_{n=1}^{\infty} \mu(A_n \cap A') \geq \mu^*(E \cap A) + \mu^*(E \cap A')
\]
since $A_n \cap A \in \mathcal{A}$ is a cover for $E \cap A$ and $A_n \cap A' \in \mathcal{A}$ is a cover for $E \cap A'$. Taking the infimum, we have $\inf_{A \in \mathcal{A}} \mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A')$, which finishes the proof. \qed

Concerning regularity we note:

**Lemma 7.9.** Suppose outer regularity (7.1) holds for every set in the algebra, then $\mu$ is outer regular.

**Proof.** By assumption we can replace each set $A_n$ in (7.8) by a possibly slightly larger open set and hence the infimum in (7.8) can be realized with open sets. \qed

Thus, as a consequence we obtain Theorem 7.3 except for regularity. Outer regularity is easy to see for a finite union of intervals since we can replace each interval by a possibly slightly larger open interval with only slightly larger measure. Inner regularity will be postponed until Lemma 7.14.

**Problem 7.6.** Show that $\mu^*$ defined in (7.8) is an outer measure. (Hint for the last property: Take a cover $\{B_{nk}\}_{k=1}^{\infty}$ for $A_n$ such that $\mu^*(A_n) = \frac{\epsilon}{2^n} + \sum_{k=1}^{\infty} \mu(B_{nk})$ and note that $\{B_{nk}\}_{n,k=1}^{\infty}$ is a cover for $\bigcup_n A_n$.)

**Problem 7.7.** Show that $\mu^*$ defined in (7.8) extends $\mu$. (Hint: For the cover $A_n$ it is no restriction to assume $A_n \cap A_m = \emptyset$ and $A_n \subseteq A$.)
Problem 7.8. Show that the measure constructed in Theorem 7.7 is complete.

Problem 7.9. Let \( \mu \) be a finite measure. Show that
\[
d(A, B) = \mu(A \Delta B), \quad A \Delta B = (A \cup B) \setminus (A \cap B) \quad (7.11)
\]
is a metric on \( \Sigma \) if we identify sets of measure zero. Show that if \( \mathcal{A} \) is an algebra, then it is dense in \( \Sigma(\mathcal{A}) \). (Hint: Show that the sets which can be approximated by sets in \( \mathcal{A} \) form a Dynkin system.)

7.3. Measurable functions

The Riemann integral works by splitting the \( x \) coordinate into small intervals and approximating \( f(x) \) on each interval by its minimum and maximum. The problem with this approach is that the difference between maximum and minimum will only tend to zero (as the intervals get smaller) if \( f(x) \) is sufficiently nice. To avoid this problem, we can force the difference to go to zero by considering, instead of an interval, the set of \( x \) for which \( f(x) \) lies between two given numbers \( a < b \). Now we need the size of the set of these \( x \), that is, the size of the preimage \( f^{-1}((a, b)) \). For this to work, preimages of intervals must be measurable.

Let \((X, \Sigma_X)\) and \((Y, \Sigma_Y)\) be measurable spaces. A function \( f : X \to Y \) is called measurable if \( f^{-1}(A) \in \Sigma_X \) for every \( A \in \Sigma_Y \). Clearly it suffices to check this condition for every set \( A \) in a collection of sets which generate \( \Sigma_Y \), since the collection of sets for which it holds forms a \( \sigma \)-algebra by \( f^{-1}(Y \setminus A) = X \setminus f^{-1}(A) \) and \( f^{-1}(\bigcup_j A_j) = \bigcup_j f^{-1}(A_j) \).

We will be mainly interested in the case where \((Y, \Sigma_Y) = (\mathbb{R}^n, \mathcal{B}^n)\).

Lemma 7.10. A function \( f : X \to \mathbb{R}^n \) is measurable if and only if
\[
f^{-1}(I) \in \Sigma \quad \forall I = \prod_{j=1}^n (a_j, \infty). \quad (7.12)
\]
In particular, a function \( f : X \to \mathbb{R}^n \) is measurable if and only if every component is measurable and a complex-valued function \( f : X \to \mathbb{C}^n \) is measurable if and only if both its real and imaginary parts are.

Proof. We need to show that \( \mathcal{B}^n \) is generated by rectangles of the above form. The \( \sigma \)-algebra generated by these rectangles also contains all open rectangles of the form \( I = \prod_{j=1}^n (a_j, b_j) \), which form a base for the topology.

Clearly the intervals \((a_j, \infty)\) can also be replaced by \([a_j, \infty), (-\infty, a_j), \) or \((-\infty, a_j] \).
7. Almost everything about Lebesgue integration

If \( X \) is a topological space and \( \Sigma \) the corresponding Borel \( \sigma \)-algebra, we will also call a measurable function \textbf{Borel function}. Note that, in particular,

\textbf{Lemma 7.11.} Let \((X, \Sigma_X), (Y, \Sigma_Y), (Z, \Sigma_Z)\) be topological spaces with their corresponding Borel \( \sigma \)-algebras. Any continuous function \( f : X \to Y \) is measurable. Moreover, if \( f : X \to Y \) and \( g : Y \to Z \) are measurable functions, then the composition \( g \circ f \) is again measurable.

The set of all measurable functions forms an algebra.

\textbf{Lemma 7.12.} Let \( X \) be a topological space and \( \Sigma \) its Borel \( \sigma \)-algebra. Suppose \( f, g : X \to \mathbb{R} \) are measurable functions. Then the sum \( f + g \) and the product \( fg \) are measurable.

\textbf{Proof.} Note that addition and multiplication are continuous functions from \( \mathbb{R}^2 \to \mathbb{R} \) and hence the claim follows from the previous lemma.

Sometimes it is also convenient to allow ±\( \infty \) as possible values for \( f \), that is, functions \( f : X \to \mathbb{R}, \mathbb{R} = \mathbb{R} \cup \{-\infty, \infty\} \). In this case \( A \subseteq \mathbb{R} \) is called Borel if \( A \cap \mathbb{R} \) is. This implies that \( f : X \to \mathbb{R} \) will be Borel if and only if \( f^{-1}(\pm \infty) \) are Borel and \( f : X \setminus f^{-1}(\{-\infty, \infty\}) \to \mathbb{R} \) is Borel. Since

\[
\{+\infty\} = \bigcap_n (n, +\infty], \quad \{-\infty\} = \bigcup_n (-n, +\infty],
\]

we see that \( f : X \to \mathbb{R} \) is measurable if and only if

\[
f^{-1}((a, \infty]) \in \Sigma \quad \forall a \in \mathbb{R}.
\]

Again the intervals \((a, \infty]\) can also be replaced by \([a, \infty], [-\infty, a], \) or \([-\infty, a] \).

Hence it is not hard to check that the previous lemma still holds if one either avoids undefined expressions of the type \( \infty - \infty \) and \( \pm \infty \cdot 0 \) or makes a definite choice, e.g., \( \infty - \infty = 0 \) and \( \pm \infty \cdot 0 = 0 \).

Moreover, the set of all measurable functions is closed under all important limiting operations.

\textbf{Lemma 7.13.} Suppose \( f_n : X \to \overline{\mathbb{R}} \) is a sequence of measurable functions. Then

\[
\inf_{n \in \mathbb{N}} f_n, \quad \sup_{n \in \mathbb{N}} f_n, \quad \liminf_{n \to \infty} f_n, \quad \limsup_{n \to \infty} f_n
\]

are measurable as well.

\textbf{Proof.} It suffices to prove that \( \sup f_n \) is measurable since the rest follows from \( \inf f_n = -\sup(-f_n), \liminf f_n = \sup_n \inf_{k \geq n} f_k, \) and \( \limsup f_n = \inf_n \sup_{k \geq n} f_k \). But \( \sup f_n \) is Borel and we are done.
A few immediate consequences are worthwhile noting: It follows that if \( f \) and \( g \) are measurable functions, so are \( \min(f, g) \), \( \max(f, g) \), \( |f| = \max(f, -f) \), and \( f^\pm = \max(\pm f, 0) \). Furthermore, the pointwise limit of measurable functions is again measurable.

Sometimes the case of arbitrary suprema and infima is also of interest. In this respect the following observation is useful: Recall that a function \( f : X \to \mathbb{R} \) is lower semicontinuous if the set \( f^{-1}((a, \infty]) \) is open for every \( a \in \mathbb{R} \). Then it follows from the definition that the sup over an arbitrary collection of lower semicontinuous functions

\[
\bar{f}(x) = \sup_{\alpha} f_\alpha(x) \tag{7.16}
\]

is again lower semicontinuous. Similarly, \( f \) is upper semicontinuous if the set \( f^{-1}([-\infty, a)) \) is open for every \( a \in \mathbb{R} \). In this case the infimum

\[
f(x) = \inf_{\alpha} f_\alpha(x) \tag{7.17}
\]

is again upper semicontinuous. Note that \( f \) is lower semicontinuous if and only if \( -f \) is upper semicontinuous.

**Problem 7.10.** Show that the supremum over lower semicontinuous functions is again lower semicontinuous.

**Problem 7.11.** Show that \( f \) is lower semicontinuous if and only if

\[
\liminf_{x \to x_0} f(x) \geq f(x_0), \quad x_0 \in X.
\]

Similarly, \( f \) is upper semicontinuous if and only if

\[
\limsup_{x \to x_0} f(x) \leq f(x_0), \quad x_0 \in X.
\]

Show a lower semicontinuous function is also sequentially lower semicontinuous

\[
\liminf_{n \to \infty} f(x_n) \geq f(x_0), \quad x_n \to x_0, \ x_0 \in X.
\]

Show that the converse is also true if \( X \) is a metric space. (Hint: Problem 1.9.)

### 7.4. How wild are measurable objects

In this section we want to investigate how far measurable objects are away from well-understood ones. As our first task we want to show that measurable sets can be well approximated by using closed sets from the inside and open sets from the outside in nice spaces like \( \mathbb{R}^n \).
Lemma 7.14. Let $X$ be a metric space and $\mu$ a Borel measure which is finite on finite balls. Then $\mu$ is $\sigma$-finite and for every $A \in \mathcal{B}(X)$ and any given $\varepsilon > 0$ there exists an open set $O$ and a closed set $F$ such that

$$F \subseteq A \subseteq O \quad \text{and} \quad \mu(O \setminus F) \leq \varepsilon. \quad (7.18)$$

Proof. That $\mu$ is $\sigma$-finite is immediate from the definitions since for any fixed $x_0 \in X$ the open balls $X_n = B_n(x_0)$ have finite measure and satisfy $X_n \not\supset X$.

To see that $(7.18)$ holds we begin with the case when $\mu$ is finite. Denote by $\mathcal{A}$ the set of all Borel sets satisfying $(7.18)$. Then $\mathcal{A}$ contains every closed set $F$: Given $F$ define $O_n = \{x \in X | d(x, F) < 1/n\}$ and note that $O_n$ are open sets which satisfy $O_n \subseteq F$. Thus by Theorem 7.1 (iii) $\mu(O_n \setminus F) \to 0$ and hence $F \in \mathcal{A}$.

Moreover, $\mathcal{A}$ is even a $\sigma$-algebra. That it is closed under complements is easy to see (note that $\bar{O} = X \setminus F$ and $\bar{F} = X \setminus O$ are the required sets for $A = X \setminus A$). To see that it is closed under countable unions consider $A = \bigcup_{n=1}^{\infty} A_n$ with $A_n \in \mathcal{A}$. Then there are $F_n$, $O_n$ such that $\mu(O_n \setminus F_n) \leq \varepsilon 2^{-n-1}$. Now $O = \bigcup_{n=1}^{\infty} O_n$ is open and $F = \bigcup_{n=1}^{N} F_n$ is closed for any finite $N$. Since $\mu(A)$ is finite we can choose $N$ sufficiently large such that $\mu(\bigcup_{n=N+1}^{\infty} F_n \setminus F) \leq \varepsilon/2$. Then we have found two sets of the required type: $\mu(O \setminus F) \leq \sum_{n=1}^{\infty} \mu(O_n \setminus F_n) + \mu(\bigcup_{n=N+1}^{\infty} F_n \setminus F) \leq \varepsilon$. Thus $\mathcal{A}$ is a $\sigma$-algebra containing the open sets, hence it is the entire Borel algebra.

Now suppose $\mu$ is not finite. Pick some $x_0 \in X$ and set $X_0 = B_{2/3}(x_0)$ and $X_n = B_{n+2/3}(x_0) \setminus B_{n-2/3}(x_0)$, $n \in \mathbb{N}$. Let $A_n = A \cap X_n$ and note that $A = \bigcup_{n=1}^{\infty} A_n$. By the finite case we can choose $F_n \subseteq A_n \subseteq O_n \subseteq X_n$ such that $\mu(O_n \setminus F_n) \leq \varepsilon 2^{-n-1}$. Now set $F = \bigcup_{n} F_n$ and $O = \bigcup_{n} O_n$ and observe that $F$ is closed. Indeed, let $x \in F$ and let $x_j$ be some sequence from $F$ converging to $x$. Since $d(x_0, x_j) \to d(x_0, x)$ this sequence must eventually lie in $F_n \cup F_{n+1}$ for some fixed $n$ implying $x \in F_n \cup F_{n+1} = F_n \cup F_{n+1} \subseteq F$. Finally, $\mu(O \setminus F) \leq \sum_{n=0}^{\infty} \mu(O_n \setminus F_n) \leq \varepsilon$ as required. \hfill $\square$

This result immediately gives us outer regularity and, if we strengthen our assumption, also inner regularity.

Corollary 7.15. Under the assumptions of the previous lemma

$$\mu(A) = \inf_{O \supseteq A, O \text{ open}} \mu(O) = \sup_{F \subseteq A, F \text{ closed}} \mu(F) \quad (7.19)$$

and $\mu$ is outer regular. If $X$ is proper (i.e., every closed ball is compact), then $\mu$ is also inner regular.

$$\mu(A) = \sup_{K \subseteq A, K \text{ compact}} \mu(K). \quad (7.20)$$
7.4. How wild are measurable objects

Proof. Finally, (7.19) follows from 
\[ \mu(A) = \mu(O) - \mu(O \setminus A) = \mu(F) + \mu(A \setminus F) \]
and if every finite ball is compact, for every sequence of closed sets \( F_n \)
with \( \mu(F_n) \to \mu(A) \) we also have compact sets \( K_n = F_n \cap B_n(x_0) \) with
\( \mu(K_n) \to \mu(A) \). \( \square \)

By the Heine–Borel theorem every bounded closed ball in \( \mathbb{R}^n \) (or \( \mathbb{C}^n \))
is compact and thus has finite measure by the very definition of a Borel
measure. Hence every Borel measure on \( \mathbb{R}^n \) (or \( \mathbb{C}^n \)) satisfies the assumptions
of Lemma 7.14.

An inner regular measure on a Hausdorff space which is locally finite
(every point has a neighborhood of finite measure) is called a Radon mea-
sure. Accordingly every Borel measure on \( \mathbb{R}^n \) (or \( \mathbb{C}^n \)) is automatically a
Radon measure.

Example. Since Lebesgue measure on \( \mathbb{R} \) is regular, we can cover the rational
numbers by an open set of arbitrary small measure (it is also not hard to find
such a set directly) but we cannot cover it by an open set of measure zero
(since any open set contains an interval and hence has positive measure).
However, if we slightly extend the family of admissible sets, this will be
possible. \( \diamond \)

Looking at the Borel \( \sigma \)-algebra the next general sets after open sets are
countable intersections of open sets, known as \( G_\delta \) sets (here \( G \) and \( \delta \) stand for
the german words Gebiet and Durchschnitt, respectively). The next general
sets after closed sets are countable unions of closed sets, known as \( F_\sigma \) sets
(here \( F \) and \( \sigma \) stand for the french words fermé and somme, respectively).

Example. The irrational numbers are a \( G_\delta \) set in \( \mathbb{R} \). To see this, let \( x_n \) be
an enumeration of the rational numbers and consider the intersection of the
open sets \( O_n = \mathbb{R} \setminus \{x_n\} \). The rational numbers are hence an \( F_\sigma \) set. \( \diamond \)

Corollary 7.16. A set in \( \mathbb{R}^n \) is Borel if and only if it differs from a \( G_\delta \) set
by a Borel set of measure zero. Similarly, a set in \( \mathbb{R}^n \) is Borel if and only if
it differs from an \( F_\sigma \) set by a Borel set of measure zero.

Proof. Since \( G_\delta \) sets are Borel, only the converse direction is nontrivial.
By Lemma 7.14 we can find open sets \( O_n \) such that \( \mu(O_n \setminus A) \leq 1/n \). Now
let \( G = \bigcap_n O_n \). Then \( \mu(G \setminus A) \leq \mu(O_n \setminus A) \leq 1/n \) for any \( n \) and thus
\( \mu(G \setminus A) = 0 \) The second claim is analogous. \( \square \)

Problem 7.12. Show directly (without using regularity) that for every \( \varepsilon > 0 \)
there is an open set \( O \) of Lebesgue measure \( |O| < \varepsilon \) which covers the rational
numbers.
Problem 7.13. Show that a Borel set \( A \subseteq \mathbb{R} \) has Lebesgue measure zero if and only if for every \( \varepsilon \) there exists a countable set of Intervals \( I_j \) which cover \( A \) and satisfy \( \sum_j |I_j| < \varepsilon \).

7.5. Integration — Sum me up, Henri

Throughout this section \((X, \Sigma, \mu)\) will be a measure space. A measurable function \( s : X \to \mathbb{R} \) is called \textbf{simple} if its image is finite; that is, if

\[
s = \sum_{j=1}^{p} \alpha_j \chi_{A_j}, \quad s(X) = \{\alpha_j\}_{j=1}^{p}, \quad A_j = s^{-1}(\alpha_j) \in \Sigma. \tag{7.21}
\]

Here \( \chi_A \) is the \textbf{characteristic function} of \( A \); that is, \( \chi_A(x) = 1 \) if \( x \in A \) and \( \chi_A(x) = 0 \) otherwise. Note that the set of simple functions is a vector space and while there are different ways of writing a simple function as a linear combination of characteristic functions, the representation (7.21) is unique.

For a nonnegative simple function \( s \) as in (7.21) we define its \textbf{integral} as

\[
\int_A s \, d\mu = \sum_{j=1}^{p} \alpha_j \mu(A_j \cap A). \tag{7.22}
\]

Here we use the convention \( 0 \cdot \infty = 0 \).

Lemma 7.17. The integral has the following properties:

(i) \( \int_A s \, d\mu = \int_X \chi_A s \, d\mu \).
(ii) \( \int_{\bigcup_{j=1}^{\infty} A_j} s \, d\mu = \sum_{j=1}^{\infty} \int_{A_j} s \, d\mu, \quad A_j \cap A_k = \emptyset \text{ for } j \neq k. \)
(iii) \( \int_A \alpha s \, d\mu = \alpha \int_A s \, d\mu, \quad \alpha \geq 0. \)
(iv) \( \int_A (s + t) \, d\mu = \int_A s \, d\mu + \int_A t \, d\mu. \)
(v) \( A \subseteq B \Rightarrow \int_A s \, d\mu \leq \int_B s \, d\mu. \)
(vi) \( s \leq t \Rightarrow \int_A s \, d\mu \leq \int_A t \, d\mu. \)

Proof. (i) is clear from the definition. (ii) follows from \( \sigma \)-additivity of \( \mu \). (iii) is obvious. (iv) Let \( s = \sum_j \alpha_j \chi_{A_j}, \quad t = \sum_j \beta_j \chi_{B_j} \) as in (7.21) and abbreviate \( C_{jk} = (A_j \cap B_k) \cap A \). Then, by (ii),

\[
\int_A (s + t) \, d\mu = \sum_{j,k} \int_{C_{jk}} (s + t) \, d\mu = \sum_{j,k} (\alpha_j + \beta_k) \mu(C_{jk})
\]

\[
= \sum_{j,k} \left( \int_{C_{jk}} s \, d\mu + \int_{C_{jk}} t \, d\mu \right) = \int_A s \, d\mu + \int_A t \, d\mu.
\]

(v) follows from monotonicity of \( \mu \). (vi) follows since by (iv) we can write \( s = \sum_j \alpha_j \chi_{C_j}, \quad t = \sum_j \beta_j \chi_{C_j} \) where, by assumption, \( \alpha_j \leq \beta_j \).

\[\square\]
Our next task is to extend this definition to nonnegative measurable functions by
\[ \int_A f \, d\mu = \sup_{s \leq f} \int_A s \, d\mu, \tag{7.23} \]
where the supremum is taken over all simple functions \( s \leq f \). Note that, except for possibly (ii) and (iv), Lemma 7.17 still holds for arbitrary nonnegative functions \( s, t \).

**Theorem 7.18** (Monotone convergence, Beppo Levi’s theorem). Let \( f_n \) be a monotone nondecreasing sequence of nonnegative measurable functions, \( f_n \nearrow f \). Then
\[ \int_A f_n \, d\mu \to \int_A f \, d\mu. \tag{7.24} \]

**Proof.** By property (vi), \( \int_A f_n \, d\mu \) is monotone and converges to some number \( \alpha \). By \( f_n \leq f \) and again (vi) we have
\[ \alpha \leq \int_A f \, d\mu. \]
To show the converse, let \( s \) be simple such that \( s \leq f \) and let \( \theta \in (0, 1) \). Put \( A_n = \{ x \in A | f_n(x) \geq \theta s(x) \} \) and note \( A_n \nearrow A \) (show this). Then
\[ \int_A f_n \, d\mu \geq \int_{A_n} f_n \, d\mu \geq \theta \int_{A_n} s \, d\mu. \]
Letting \( n \to \infty \), we see
\[ \alpha \geq \theta \int_A s \, d\mu. \]
Since this is valid for every \( \theta < 1 \), it still holds for \( \theta = 1 \). Finally, since \( s \leq f \) is arbitrary, the claim follows. \( \square \)

In particular
\[ \int_A f \, d\mu = \lim_{n \to \infty} \int_A s_n \, d\mu, \tag{7.25} \]
for every monotone sequence \( s_n \nearrow f \) of simple functions. Note that there is always such a sequence, for example,
\[ s_n(x) = \sum_{k=0}^{2^n} \frac{k}{2^n} \chi_{f^{-1}(A_k)}(x), \quad A_k = \left[ \frac{k}{2^n}, \frac{k + 1}{2^n} \right), A_{n2^n} = [n, \infty). \tag{7.26} \]
By construction \( s_n \) converges uniformly if \( f \) is bounded, since \( 0 \leq f(x) - s_n(x) < \frac{1}{2^n} \) if \( f(x) \leq n \).

Now what about the missing items (ii) and (iv) from Lemma 7.17? Since limits can be spread over sums, item (iv) holds, and (ii) also follows directly from the monotone convergence theorem. We even have the following result:
Lemma 7.19. If \( f \geq 0 \) is measurable, then \( d\nu = f \, d\mu \) defined via
\[
\nu(A) = \int_A f \, d\mu
\]
is a measure such that
\[
\int g \, d\nu = \int g \, f \, d\mu
\]
for every measurable function \( g \).

**Proof.** As already mentioned, additivity of \( \nu \) is equivalent to linearity of the integral and \( \sigma \)-additivity follows from Theorem 7.18:
\[
\nu\left( \bigcup_{n=1}^{\infty} A_n \right) = \int \left( \sum_{n=1}^{\infty} \chi_{A_n} \right) f \, d\mu = \sum_{n=1}^{\infty} \int \chi_{A_n} f \, d\mu = \sum_{n=1}^{\infty} \nu(A_n).
\]
The second claim holds for simple functions and hence for all functions by construction of the integral. \( \Box \)

If \( f_n \) is not necessarily monotone, we have at least

**Theorem 7.20** (Fatou’s lemma). If \( f_n \) is a sequence of nonnegative measurable function, then
\[
\int_A \liminf_{n \to \infty} f_n \, d\mu \leq \liminf_{n \to \infty} \int_A f_n \, d\mu.
\]

**Proof.** Set \( g_n = \inf_{k \geq n} f_k \). Then \( g_n \leq f_n \) implying
\[
\int_A g_n \, d\mu \leq \int_A f_n \, d\mu.
\]
Now take the \( \liminf \) on both sides and note that by the monotone convergence theorem
\[
\liminf_{n \to \infty} \int_A g_n \, d\mu = \lim_{n \to \infty} \int_A g_n \, d\mu = \int_A \lim_{n \to \infty} g_n \, d\mu = \int_A \liminf_{n \to \infty} f_n \, d\mu,
\]
proving the claim. \( \Box \)

**Example.** Consider \( f_n = \chi_{[n,n+1]} \). Then \( \lim_{n \to \infty} f_n(x) = 0 \) for every \( x \in \mathbb{R} \). However, \( \int_{\mathbb{R}} f_n(x) \, dx = 1 \). This shows that the inequality in Fatou’s lemma cannot be replaced by equality in general. \( \diamond \)

If the integral is finite for both the positive and negative part \( f^\pm = \max(\pm f, 0) \) of an arbitrary measurable function \( f \), we call \( f \) **integrable** and set
\[
\int_A f \, d\mu = \int_A f^+ \, d\mu - \int_A f^- \, d\mu.
\]
Similarly, we handle the case where $f$ is complex-valued by calling $f$ integrable if both the real and imaginary part are and setting

$$
\int_A f \, d\mu = \int_A \text{Re}(f) \, d\mu + i \int_A \text{Im}(f) \, d\mu.
$$

(7.31)

Clearly $f$ is integrable if and only if $|f|$ is. The set of all integrable functions is denoted by $L^1(X, d\mu)$.

**Lemma 7.21.** The integral is linear and Lemma 7.17 holds for integrable functions $s, t$.

Furthermore, for all integrable functions $f, g$ we have

$$
|\int_A f \, d\mu| \leq \int_A |f| \, d\mu
$$

(7.32)

and (triangle inequality)

$$
\int_A |f + g| \, d\mu \leq \int_A |f| \, d\mu + \int_A |g| \, d\mu.
$$

(7.33)

**Proof.** Linearity and Lemma 7.17 are straightforward to check. To see (7.32) put $\alpha = \frac{z}{|z|}$, where $z = \int_A f \, d\mu$ (without restriction $z \neq 0$). Then

$$
|\int_A f \, d\mu| = \alpha \int_A f \, d\mu = \int_A \alpha f \, d\mu = \int_A \text{Re}(\alpha f) \, d\mu \leq \int_A |f| \, d\mu,
$$

proving (7.32). The last claim follows from $|f + g| \leq |f| + |g|$.

**Lemma 7.22.** Let $f$ be measurable. Then

$$
\int_X |f| \, d\mu = 0 \iff f(x) = 0 \text{ $\mu$ - a.e.}
$$

(7.34)

Moreover, suppose $f$ is nonnegative or integrable. Then

$$
\mu(A) = 0 \implies \int_A f \, d\mu = 0.
$$

(7.35)

**Proof.** Observe that we have $A = \{x|f(x) \neq 0\} = \bigcup_n A_n$, where $A_n = \{x| |f(x)| > \frac{1}{n}\}$. If $\int_X |f| \, d\mu = 0$ we must have $\mu(A_n) = 0$ for every $n$ and hence $\mu(A) = \lim_{n \to \infty} \mu(A_n) = 0$.

The converse will follow from (7.35) since $\mu(A) = 0$ (with $A$ as before) implies $\int_X |f| \, d\mu = \int_A |f| \, d\mu = 0$.

Finally, to see (7.35) note that by our convention $0 \cdot \infty = 0$ it holds for any simple function and hence for any nonnegative $f$ by definition of the integral (7.23). Since any function can be written as a linear combination of four nonnegative functions this also implies the case when $f$ is integrable. 

$\Box$
Note that the proof also shows that if $f$ is not 0 almost everywhere, there is an $\varepsilon > 0$ such that $\mu(\{x \mid |f(x)| \geq \varepsilon\}) > 0$.

In particular, the integral does not change if we restrict the domain of integration to a support of $\mu$ or if we change $f$ on a set of measure zero. In particular, functions which are equal a.e. have the same integral.

Finally, our integral is well behaved with respect to limiting operations. We first state a simple generalization of Fatou’s lemma.

**Lemma 7.23** (generalized Fatou lemma). If $f_n$ is a sequence of real-valued measurable function and $g$ some integrable function. Then

$$\int_A \liminf_{n \to \infty} f_n \, d\mu \leq \liminf_{n \to \infty} \int_A f_n \, d\mu \quad (7.36)$$

if $g \leq f_n$ and

$$\limsup_{n \to \infty} \int_A f_n \, d\mu \leq \int_A \limsup_{n \to \infty} f_n \, d\mu \quad (7.37)$$

if $f_n \leq g$.

**Proof.** To see the first apply Fatou’s lemma to $f_n - g$ and subtract $\int_A g \, d\mu$ on both sides of the result. The second follows from the first using $\lim \inf (-f_n) = - \lim \sup f_n$. \qed

If in the last lemma we even have $|f_n| \leq g$, we can combine both estimates to obtain

$$\int_A \liminf_{n \to \infty} f_n \, d\mu \leq \liminf_{n \to \infty} \int_A f_n \, d\mu \leq \limsup_{n \to \infty} \int_A f_n \, d\mu \leq \int_A \limsup_{n \to \infty} f_n \, d\mu, \quad (7.38)$$

which is known as Fatou–Lebesgue theorem. In particular, in the special case where $f_n$ converges we obtain

**Theorem 7.24** (Dominated convergence). Let $f_n$ be a convergent sequence of measurable functions and set $f = \lim_{n \to \infty} f_n$. Suppose there is an integrable function $g$ such that $|f_n| \leq g$. Then $f$ is integrable and

$$\lim_{n \to \infty} \int_A f_n d\mu = \int_A f d\mu. \quad (7.39)$$

**Proof.** The real and imaginary parts satisfy the same assumptions and hence it suffices to prove the case where $f_n$ and $f$ are real-valued. Moreover, since $\lim \inf f_n = \lim \sup f_n = f$ equation (7.38) establishes the claim. \qed

Remark: Since sets of measure zero do not contribute to the value of the integral, it clearly suffices if the requirements of the dominated convergence theorem are satisfied almost everywhere (with respect to $\mu$).

**Example.** Note that the existence of $g$ is crucial: The functions $f_n(x) = \frac{1}{2n} \chi_{[-n,n]}(x)$ on $\mathbb{R}$ converge uniformly to 0 but $\int_{\mathbb{R}} f_n(x) dx = 1$. \diamond
Example. If \( \mu(x) = \Theta(x) \) is the Dirac measure at 0, then
\[
\int_{\mathbb{R}} f(x) d\mu(x) = f(0).
\]
In fact, the integral can be restricted to any support and hence to \( \{0\} \).

If \( \mu(x) = \sum_n \alpha_n \Theta(x - x_n) \) is a sum of Dirac measures, \( \Theta(x) \) centered at \( x = 0 \), then (Problem 7.14)
\[
\int_{\mathbb{R}} f(x) d\mu(x) = \sum_n \alpha_n f(x_n).
\]
Hence our integral contains sums as special cases.

Finally, let me conclude this section with a remark on how to compute Lebesgue integrals in the classical case of Lebesgue measure on some interval \((a,b) \subseteq \mathbb{R}\). Given a continuous function \( f \in C(a,b) \) which is integrable over \((a,b)\) we can introduce
\[
F(x) = \int_{(a,x]} f(y) dy, \quad x \in (a,b).
\]
(7.40)
Then one has
\[
\frac{F(x + \varepsilon) - F(x)}{\varepsilon} = f(x) + \frac{1}{\varepsilon} \int_{(x,x+\varepsilon]} (f(y) - f(x)) dy
\]
(where \((x,x+\varepsilon]\) has to be understood as \((x+\varepsilon,x]\) if \( \varepsilon < 0 \)) and
\[
\limsup_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_{(x,x+\varepsilon]} |f(y) - f(x)| dy \leq \limsup_{\varepsilon \to 0} \sup_{y \in (x,x+\varepsilon]} |f(y) - f(x)| = 0
\]
by the continuity of \( f \) at \( x \). Thus \( F \in C^1(a,b) \) and
\[
F'(x) = f(x),
\]
which is a variant of the fundamental theorem of calculus. This tells us that the integral of a continuous function \( f \) can be computed in terms of its antiderivative and, in particular, all tools from calculus like integration by parts or integration by substitution are readily available for the Lebesgue integral on \( \mathbb{R} \). Moreover, the Lebesgue integral must coincide with the Riemann integral for (piecewise) continuous functions. More on the connection with the Riemann integral will be given in Section 7.9. A generalization of the fundamental theorem of calculus will be given in Theorem 9.25.

Problem 7.14. Consider a countable set of measures \( \mu_n \) and numbers \( \alpha_n \geq 0 \). Let \( \mu = \sum_n \alpha_n \mu_n \) and show
\[
\int_A f \, d\mu = \sum_n \alpha_n \int_A f \, d\mu_n
\]
for any measurable function which is either nonnegative or integrable.
Problem 7.15. Show that the set $B(X)$ of bounded measurable functions with the sup norm is a Banach space. Show that the set $S(X)$ of simple functions is dense in $B(X)$. Show that the integral is a bounded linear functional on $B(X)$ if $\mu(X) < \infty$. (Hence Theorem 1.35 could be used to extend the integral from simple to bounded measurable functions.)

Problem 7.16. Show that the monotone convergence holds for nondecreasing sequences of real-valued measurable functions $f_n \nearrow f$ provided $f_1$ is integrable.

Problem 7.17. Show that the dominated convergence theorem implies (under the same assumptions)

$$\lim_{n \to \infty} \int |f_n - f| d\mu = 0.$$  

Problem 7.18. Let $f$ be an integrable function satisfying $f(x) \leq M$. Show that

$$\int_A f \, d\mu \leq M \mu(A)$$

with equality if and only if $f(x) = M$ for a.e. $x \in A$.

Problem 7.19. Let $X \subseteq \mathbb{R}$, $Y$ be some measure space, and $f : X \times Y \to \mathbb{C}$. Suppose $y \mapsto f(x, y)$ is measurable for every $x$ and $x \mapsto f(x, y)$ is continuous for every $y$. Show that

$$F(x) = \int_A f(x, y) \, d\mu(y)$$

is continuous if there is an integrable function $g(y)$ such that $|f(x, y)| \leq g(y)$.

Problem 7.20. Let $X \subseteq \mathbb{R}$, $Y$ be some measure space, and $f : X \times Y \to \mathbb{C}$. Suppose $y \mapsto f(x, y)$ is integrable for all $x$ and $x \mapsto f(x, y)$ is differentiable for a.e. $y$. Show that

$$F(x) = \int_A f(x, y) \, d\mu(y)$$

is differentiable if there is an integrable function $g(y)$ such that $|\frac{\partial}{\partial x} f(x, y)| \leq g(y)$. Moreover, $y \mapsto \frac{\partial}{\partial x} f(x, y)$ is measurable and

$$F'(x) = \int_A \frac{\partial}{\partial x} f(x, y) \, d\mu(y)$$

in this case. (See Problem 9.27 for an extension.)
7.6. Product measures

Let $\mu_1$ and $\mu_2$ be two measures on $\Sigma_1$ and $\Sigma_2$, respectively. Let $\Sigma_1 \otimes \Sigma_2$ be the $\sigma$-algebra generated by rectangles of the form $A_1 \times A_2$.

**Example.** Let $\mathcal{B}$ be the Borel sets in $\mathbb{R}$. Then $\mathcal{B}^2 = \mathcal{B} \otimes \mathcal{B}$ are the Borel sets in $\mathbb{R}^2$ (since the rectangles are a basis for the product topology). $\diamond$

Any set in $\Sigma_1 \otimes \Sigma_2$ has the **section property**; that is,

**Lemma 7.25.** Suppose $A \in \Sigma_1 \otimes \Sigma_2$. Then its sections

$$A_1(x_2) = \{x_1|(x_1, x_2) \in A\} \quad \text{and} \quad A_2(x_1) = \{x_2|(x_1, x_2) \in A\}$$

(7.45)

are measurable.

**Proof.** Denote all sets $A \in \Sigma_1 \otimes \Sigma_2$ with the property that $A_1(x_2) \in \Sigma_1$ by $S$. Clearly all rectangles are in $S$ and it suffices to show that $S$ is a $\sigma$-algebra. Now, if $A \in S$, then $(A')_1(x_2) = (A_1(x_2))' \in \Sigma_1$ and thus $S$ is closed under complements. Similarly, if $A_n \in S$, then $(\bigcup_n A_n)_1(x_2) = \bigcup_n (A_n)_1(x_2)$ shows that $S$ is closed under countable unions.

This implies that if $f$ is a measurable function on $X_1 \times X_2$, then $f(., x_2)$ is measurable on $X_1$ for every $x_2$ and $f(x_1, .)$ is measurable on $X_2$ for every $x_1$ (observe $A_1(x_2) = \{x_1|f(x_1, x_2) \in B\}$, where $A = \{(x_1, x_2)|f(x_1, x_2) \in B\}$).

Given two measures $\mu_1$ on $\Sigma_1$ and $\mu_2$ on $\Sigma_2$, we now want to construct the **product measure** $\mu_1 \otimes \mu_2$ on $\Sigma_1 \otimes \Sigma_2$ such that

$$\mu_1 \otimes \mu_2(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2), \quad A_j \in \Sigma_j, \quad j = 1, 2.$$ (7.46)

Since the rectangles are closed under intersection, Theorem 7.5 implies that there is at most one measure on $\Sigma_1 \otimes \Sigma_2$ provided $\mu_1$ and $\mu_2$ are $\sigma$-finite.

**Theorem 7.26.** Let $\mu_1$ and $\mu_2$ be two $\sigma$-finite measures on $\Sigma_1$ and $\Sigma_2$, respectively. Let $A \in \Sigma_1 \otimes \Sigma_2$. Then $\mu_2(A_2(x_1))$ and $\mu_1(A_1(x_2))$ are measurable and

$$\int_{X_1} \mu_2(A_2(x_1))d\mu_1(x_1) = \int_{X_2} \mu_1(A_1(x_2))d\mu_2(x_2).$$ (7.47)

**Proof.** As usual, we begin with the case where $\mu_1$ and $\mu_2$ are finite. Let $\mathcal{D}$ be the set of all subsets for which our claim holds. Note that $\mathcal{D}$ contains at least all rectangles. Thus it suffices to show that $\mathcal{D}$ is a Dynkin system by Lemma 7.4. To see this, note that measurability and equality of both integrals follow from $A_1(x_2)' = A_1'(x_2)$ (implying $\mu_1(A_1'(x_2)) = \mu_1(X_1) - \mu_1(A_1(x_2))$) for complements and from the monotone convergence theorem for disjoint unions of sets.
If $\mu_1$ and $\mu_2$ are $\sigma$-finite, let $X_{i,j} \not
st X_1$ with $\mu_i(X_{i,j}) < \infty$ for $i = 1, 2$. Now $\mu_2((A \cap X_{i,j} \times X_{2,j})2(x_1)) = \mu_2(A_{2}(x_1) \cap X_{2,j})\chi_{X_{1,j}}(x_1)$ and similarly with 1 and 2 exchanged. Hence by the finite case

$$\int_{X_1} \mu_2(A_{2} \cap X_{2,j})\chi_{X_{1,j}} d\mu_1 = \int_{X_2} \mu_1(A_{1} \cap X_{1,j})\chi_{X_{2,j}} d\mu_2$$

(7.48)

and the $\sigma$-finite case follows from the monotone convergence theorem. \qed

Hence for given $A \in \Sigma_1 \otimes \Sigma_2$ we can define

$$\mu_1 \otimes \mu_2(A) = \int_{X_1} \mu_2(A_{2}(x_1))d\mu_1(x_1) = \int_{X_2} \mu_1(A_{1}(x_2))d\mu_2(x_2)$$

(7.49)

or equivalently, since $\chi_{A_{1}(x_2)}(x_1) = \chi_{A_{2}(x_1)}(x_2) = \chi_A(x_1, x_2)$,

$$\mu_1 \otimes \mu_2(A) = \int_{X_1} \left( \int_{X_2} \chi_A(x_1, x_2)d\mu_2(x_2) \right) d\mu_1(x_1)$$

$$= \int_{X_2} \left( \int_{X_1} \chi_A(x_1, x_2)d\mu_1(x_1) \right) d\mu_2(x_2).$$

(7.50)

Then $\mu_1 \otimes \mu_2$ gives rise to a unique measure on $A \in \Sigma_1 \otimes \Sigma_2$ since $\sigma$-additivity follows from the monotone convergence theorem.

Finally we have

**Theorem 7.27** (Fubini). Let $f$ be a measurable function on $X_1 \times X_2$ and let $\mu_1$, $\mu_2$ be $\sigma$-finite measures on $X_1$, $X_2$, respectively.

(i) If $f \geq 0$, then $\int f(., x_2)d\mu_2(x_2)$ and $\int f(x_1,.d\mu_1(x_1)$ are both measurable and

$$\iint f(x_1, x_2)d\mu_1 \otimes \mu_2(x_1, x_2) = \int \left( \int f(x_1, x_2)d\mu_1(x_1) \right) d\mu_2(x_2)$$

$$= \int \left( \int f(x_1, x_2)d\mu_2(x_2) \right) d\mu_1(x_1).$$

(7.51)

(ii) If $f$ is complex-valued, then

$$\int |f(x_1, x_2)|d\mu_1(x_1) \in L^1(X_2, d\mu_2),$$

(7.52)

respectively,

$$\int |f(x_1, x_2)|d\mu_2(x_2) \in L^1(X_1, d\mu_1),$$

(7.53)

if and only if $f \in L^1(X_1 \times X_2, d\mu_1 \otimes d\mu_2)$. In this case (7.51) holds.

**Proof.** By Theorem 7.26 and linearity the claim holds for simple functions. To see (i), let $s_n \not
st f$ be a sequence of nonnegative simple functions. Then it
follows by applying the monotone convergence theorem (twice for the double integrals).

For (ii) we can assume that \( f \) is real-valued by considering its real and imaginary parts separately. Moreover, splitting \( f = f^+ - f^- \) into its positive and negative parts, the claim reduces to (i). \( \square \)

In particular, if \( f(x_1, x_2) \) is either nonnegative or integrable, then the order of integration can be interchanged. The case of nonnegative functions is also called Tonelli’s theorem. In the general case the integrability condition is crucial, as the following example shows.

**Example.** Let \( X = [0, 1] \times [0, 1] \) with Lebesgue measure and consider

\[
    f(x, y) = \frac{x - y}{(x + y)^3}.
\]

Then

\[
    \int_0^1 \int_0^1 f(x, y) \, dx \, dy = -\int_0^1 \frac{1}{(1 + y)^2} \, dy = -\frac{1}{2}
\]

but (by symmetry)

\[
    \int_0^1 \int_0^1 f(x, y) \, dy \, dx = \int_0^1 \frac{1}{(1 + x)^2} \, dx = \frac{1}{2}.
\]

Consequently \( f \) cannot be integrable over \( X \) (verify this directly). \( \diamond \)

**Lemma 7.28.** If \( \mu_1 \) and \( \mu_2 \) are outer regular measures, then so is \( \mu_1 \otimes \mu_2 \).

**Proof.** Outer regularity holds for every rectangle and hence also for the algebra of finite disjoint unions of rectangles (Problem 7.21). Thus the claim follows from Lemma 7.9. \( \square \)

In connection with Theorem 7.5 the following observation is of interest:

**Lemma 7.29.** If \( S_1 \) generates \( \Sigma_1 \) and \( S_2 \) generates \( \Sigma_2 \), then \( S_1 \times S_2 = \{A_1 \times A_2 | A_j \in S_j, j = 1, 2\} \) generates \( \Sigma_1 \otimes \Sigma_2 \).

**Proof.** Denote the \( \sigma \)-algebra generated by \( S_1 \times S_2 \) by \( \Sigma \). Consider the set \( \{A_1 \in \Sigma_1 | A_1 \times X_2 \in \Sigma\} \) which is clearly a \( \sigma \)-algebra containing \( S_1 \) and thus equal to \( \Sigma_1 \). In particular, \( \Sigma_1 \times X_2 \subset \Sigma \) and similarly \( X_1 \times \Sigma_2 \subset \Sigma \). Hence also \((\Sigma_1 \times X_2) \cap (X_1 \times \Sigma_2) = \Sigma_1 \times \Sigma_2 \subset \Sigma \). \( \square \)

Finally, note that we can iterate this procedure.

**Lemma 7.30.** Suppose \( (X_j, \Sigma_j, \mu_j), j = 1, 2, 3, \) are \( \sigma \)-finite measure spaces. Then \( (\Sigma_1 \otimes \Sigma_2) \otimes \Sigma_3 = \Sigma_1 \otimes (\Sigma_2 \otimes \Sigma_3) \) and

\[
    (\mu_1 \otimes \mu_2) \otimes \mu_3 = \mu_1 \otimes (\mu_2 \otimes \mu_3). \tag{7.54}
\]
Proof. First of all note that $(\Sigma_1 \otimes \Sigma_2) \otimes \Sigma_3 = \Sigma_1 \otimes (\Sigma_2 \otimes \Sigma_3)$ is the sigma algebra generated by the rectangles $A_1 \times A_2 \times A_3$ in $X_1 \times X_2 \times X_3$. Moreover, since
\[(\mu_1 \otimes \mu_2 \otimes \mu_3)(A_1 \times A_2 \times A_3) = \mu_1(A_1)\mu_2(A_2)\mu_3(A_3) = (\mu_1 \otimes (\mu_2 \otimes \mu_3))(A_1 \times A_2 \times A_3),\]
the two measures coincide on rectangles and hence everywhere by Theorem 7.5.

Example. If $\lambda$ is Lebesgue measure on $\mathbb{R}$, then $\lambda^n = \lambda \otimes \cdots \otimes \lambda$ is Lebesgue measure on $\mathbb{R}^n$. Since $\lambda$ is outer regular, so is $\lambda^n$. Of course regularity also follows from Corollary 7.15.

Moreover, Lebesgue measure is translation invariant and up to normalization the only measure with this property. To see this, let $\mu$ be a second translation invariant measure. Denote by $Q_r$ a cube with side length $r > 0$. Without loss we can assume $\mu(Q_1) = 1$. Since we can split $Q_1$ into $m^n$ cubes of side length $1/m$, we see that $\mu(Q_1/m) = m^{-n}$ by translation invariance and additivity. Hence we obtain $\mu(Q_r) = r^n$ for every rational $r$ and thus for every $r$ by continuity from below. Proceeding like this we see that $\lambda^n$ and $\mu$ coincide on all rectangles which are products of bounded open intervals. Since this set is closed under intersections and generates the Borel algebra $\mathcal{B}^n$ by Lemma 7.29 the claim follows again from Theorem 7.5.

Problem 7.21. Show that the set of all finite union of rectangles $A_1 \times A_2$ forms an algebra. Moreover, every set in this algebra can be written a finite union of disjoint rectangles.

Problem 7.22. Let $U \subseteq \mathbb{C}$ be a domain, $Y$ be some measure space, and $f : U \times Y \to \mathbb{C}$. Suppose $y \mapsto f(z, y)$ is measurable for every $z$ and $z \mapsto f(z, y)$ is holomorphic for every $y$. Show that
\[F(z) = \int_A f(z, y) \, d\mu(y)\]
(7.55)
is holomorphic if for every compact subset $V \subseteq U$ there is an integrable function $g(y)$ such that $|f(z, y)| \leq g(y)$, $z \in V$. (Hint: Use Fubini and Morera.)

7.7. Transformation of measures and integrals

Finally we want to transform measures. Let $f : X \to Y$ be a measurable function. Given a measure $\mu$ on $X$ we can introduce a measure $f_*\mu$ on $Y$ via
\[(f_*\mu)(A) = \mu(f^{-1}(A)).\]
(7.56)
It is straightforward to check that $f_*\mu$ is indeed a measure. Moreover, note that $f_*\mu$ is supported on the range of $f$.

**Theorem 7.31.** Let $f : X \to Y$ be measurable and let $g : Y \to \mathbb{C}$ be a Borel function. Then the Borel function $g \circ f : X \to \mathbb{C}$ is a.e. nonnegative or integrable if and only if $g$ is and in both cases

\[
\int_Y g \, d(f_*\mu) = \int_X g \circ f \, d\mu. \tag{7.57}
\]

**Proof.** In fact, it suffices to check this formula for simple functions $g$, which follows since $\chi_A \circ f = \chi_{f^{-1}(A)}$. $\Box$

**Example.** Suppose $f : X \to Y$ and $g : Y \to Z$. Then

\[
(g \circ f)_*\mu = g_*(f_*\mu),
\]

since $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. $\diamond$

**Example.** Let $f(x) = Mx + a$ be an affine transformation, where $M : \mathbb{R}^n \to \mathbb{R}^n$ is some invertible matrix. Then Lebesgue measure transforms according to

\[
f_*\lambda^n = \frac{1}{|\det(M)|}\lambda^n.
\]

To see this, note that $f_*\lambda^n$ is translation invariant and hence must be a multiple of $\lambda^n$. Moreover, for an orthogonal matrix this multiple is one (since an orthogonal matrix leaves the unit ball invariant) and for a diagonal matrix it must be the absolute value of the product of the diagonal elements (consider a rectangle). Finally since every matrix can be written as $M = O_1 DO_2$, where $O_j$ are orthogonal and $D$ is diagonal (Problem 7.24), the claim follows.

As a consequence we obtain

\[
\int_A g(Mx + a) d^n x = \frac{1}{|\det(M)|} \int_{MA + a} g(y) d^n y,
\]

which applies, for example, to shifts $f(x) = x + a$ or scaling transforms $f(x) = \alpha x$. $\diamond$

This result can be generalized to diffeomorphisms (one-to-one $C^1$ maps with inverse again $C^1$):

**Lemma 7.32.** Let $U, V \subseteq \mathbb{R}^n$ and suppose $f \in C^1(U, V)$ is a diffeomorphism. Then

\[
(f^{-1})_*d^n x = |J_f(x)| d^n x, \tag{7.58}
\]

where $J_f = \det\frac{\partial f}{\partial x}$ is the Jacobi determinant of $f$. In particular,

\[
\int_U g(f(x)) |J_f(x)| d^n x = \int_V g(y) d^n y. \tag{7.59}
\]
Proof. It suffices to show

$$\int_{f(R)} d^n y = \int_R |J_f(x)| d^n x$$

for every bounded open rectangle $R \subseteq U$. By Theorem 7.5 it will then follow for characteristic functions and thus for arbitrary functions by the very definition of the integral.

To this end we consider the integral

$$I_\varepsilon = \int_{f(R)} \int_R |J_f(f^{-1}(y)))| \varphi_\varepsilon(f(z) - y) d^n z d^n y$$

Here $\varphi = V_n^{-1} \chi_{B_1(0)}$ and $\varphi_\varepsilon(y) = \varepsilon^{-n} \varphi(\varepsilon^{-1} y)$, where $V_n$ is the volume of the unit ball (cf. below), such that $\int \varphi_\varepsilon(x) d^n x = 1$.

We will evaluate this integral in two ways. To begin with we consider the inner integral

$$h_\varepsilon(y) = \int_{W_\varepsilon(x)} \varphi \left( \frac{f(x + \varepsilon w) - f(x)}{\varepsilon} \right) d^n w, \quad W_\varepsilon(x) = \frac{1}{\varepsilon}(K - x).$$

By

$$\left| \frac{f(x + \varepsilon w) - f(x)}{\varepsilon} \right| \geq \frac{1}{C} |w|, \quad C = \sup_K ||df^{-1}||$$

the integrand is nonzero only for $w \in B_C(0)$. Hence, as $\varepsilon \to 0$ the domain $W_\varepsilon(x)$ will eventually cover all of $B_C(0)$ and dominated convergence implies

$$\lim_{\varepsilon \downarrow 0} h_\varepsilon(y) = \int_{B_C(0)} \varphi(df(x) w) dw = |J_f(x)|^{-1}.$$

Consequently, $\lim_{\varepsilon \downarrow 0} I_\varepsilon = |f(R)|$ again by dominated convergence. Now we use Fubini to interchange the order of integration

$$I_\varepsilon = \int_R \int_{f(R)} |J_f(f^{-1}(y)))| \varphi_\varepsilon(f(z) - y) d^n y d^n z.$$

Since $f(z)$ is an interior point of $f(R)$ continuity of $|J_f(f^{-1}(y)))|$ implies

$$\lim_{\varepsilon \downarrow 0} \int_{f(R)} |J_f(f^{-1}(y)))| \varphi_\varepsilon(f(z) - y) d^n y = |J_f(f^{-1}(f(z)))| = |J_f(z)|$$

and hence dominated convergence shows $\lim_{\varepsilon \downarrow 0} I_\varepsilon = \int_R |J_f(z)| d^n z$. □
Example. For example, we can consider polar coordinates \( T_2 : [0, \infty) \times [0, 2\pi) \to \mathbb{R}^2 \) defined by
\[
T_2(\rho, \varphi) = (\rho \cos(\varphi), \rho \sin(\varphi)).
\]
Then
\[
\det \frac{\partial T_2}{\partial (\rho, \varphi)} = \det \begin{vmatrix} \cos(\varphi) & -\rho \sin(\varphi) \\ \sin(\varphi) & \rho \cos(\varphi) \end{vmatrix} = \rho
\]
and one has
\[
\int_U f(\rho \cos(\varphi), \rho \sin(\varphi)) \rho \ d(\rho, \varphi) = \int_{T_2(U)} f(x) d^2 x.
\]
Note that \( T_2 \) is only bijective when restricted to \((0, \infty) \times [0, 2\pi)\). However, since the set \( \{0\} \times [0, 2\pi) \) is of measure zero, it does not contribute to the integral on the left. Similarly, its image \( T_2(\{0\} \times [0, 2\pi)) = \{0\} \) does not contribute to the integral on the right. ⊳

Example. We can use the previous example to obtain the transformation formula for spherical coordinates in \( \mathbb{R}^n \) by induction. We illustrate the process for \( n = 3 \). To this end let \( x = (x_1, x_2, x_3) \) and start with spherical coordinates in \( \mathbb{R}^2 \) (which are just polar coordinates) for the first two components:
\[
x = (\rho \cos(\varphi), \rho \sin(\varphi), x_3), \quad \rho \in [0, \infty), \ \varphi \in [0, 2\pi).
\]
Next use polar coordinates for \((\rho, x_3)\):
\[
(\rho, x_3) = (r \sin(\theta), r \cos(\theta)), \quad r \in [0, \infty), \ \theta \in [0, \pi].
\]
Note that the range for \( \theta \) follows since \( \rho \geq 0 \). Moreover, observe that
\[
r^2 = \rho^2 + x_3^2 = x_1^2 + x_2^2 + x_3^2 = |x|^2
\]
as already anticipated by our notation. In summary,
\[
x = T_3(r, \varphi, \theta) = (r \sin(\theta) \cos(\varphi), r \sin(\theta) \sin(\varphi), r \cos(\theta)).
\]
Furthermore, since \( T_3 \) is the composition with \( T_2 \) acting on the first two coordinates with the last unchanged and polar coordinates \( P \) acting on the first and last coordinate, the chain rule implies
\[
\det \frac{\partial T_3}{\partial (r, \varphi, \theta)} = \det \frac{\partial T_2}{\partial (\rho, \varphi, x_3)} \left. \frac{\partial P}{\partial (r, \varphi, \theta)} \right|_{\rho = r \sin(\theta), \ \varphi = r \cos(\theta)} = r^2 \sin(\theta).
\]
Hence one has
\[
\int_U f(T_3(r, \varphi, \theta)) r^2 \sin(\theta) d(r, \varphi, \theta) = \int_{T_2(T_3(U))} f(x) d^3 x.
\]
Again \( T_3 \) is only bijective on \((0, \infty) \times [0, 2\pi) \times (0, \pi)\).

It is left as an exercise to check that the extension \( T_n : [0, \infty) \times [0, 2\pi) \times [0, \pi]^{n-2} \to \mathbb{R}^n \) is given by
\[
x = T_n(r, \varphi, \theta_1, \ldots, \theta_{n-2})
\]
7. Almost everything about Lebesgue integration

with

\[
\begin{align*}
x_1 &= r \cos(\varphi) \sin(\theta_1) \sin(\theta_2) \cdots \sin(\theta_{n-2}), \\
x_2 &= r \sin(\varphi) \sin(\theta_1) \sin(\theta_2) \cdots \sin(\theta_{n-2}), \\
x_3 &= r \cos(\theta_1) \sin(\theta_2) \cdots \sin(\theta_{n-2}), \\
x_4 &= r \cos(\theta_2) \sin(\theta_3) \cdots \sin(\theta_{n-2}), \\
&\vdots \\
x_{n-1} &= r \cos(\theta_{n-3}) \sin(\theta_{n-2}), \\
x_n &= r \cos(\theta_{n-2}).
\end{align*}
\]

The Jacobi determinant is given by

\[
\det \frac{\partial T_n}{\partial (r, \varphi, \theta_1, \ldots, \theta_{n-2})} = r^{n-1} \sin(\theta_1) \sin(\theta_2)^2 \cdots \sin(\theta_{n-2})^{n-2}.
\]

Another useful consequence of Theorem 7.31 is the following rule for integrating radial functions.

**Lemma 7.33.** There is a measure \( \sigma^{n-1} \) on the unit sphere \( S^{n-1} = \partial B_1(0) = \{ x \in \mathbb{R}^n \mid |x| = 1 \} \), which is rotation invariant and satisfies

\[
\int_{\mathbb{R}^n} g(x) d^n x = \int_0^\infty \int_{S^{n-1}} g(r\omega) r^{n-1} d\sigma^{n-1}(\omega) dr,
\]

(7.60)

for every integrable (or positive) function \( g \).

Moreover, the surface area of \( S^{n-1} \) is given by

\[
S_n = \sigma^{n-1}(S^{n-1}) = nV_n,
\]

(7.61)

where \( V_n = \lambda^n(B_1(0)) \) is the volume of the unit ball in \( \mathbb{R}^n \), and if \( g(x) = \tilde{g}(|x|) \) is radial we have

\[
\int_{\mathbb{R}^n} \tilde{g}(|x|) d^n x = S_n \int_0^\infty \tilde{g}(r) r^{n-1} dr.
\]

(7.62)

**Proof.** Consider the transformation \( f : \mathbb{R}^n \to [0, \infty) \times S^{n-1}, x \mapsto (|x|, \frac{x}{|x|}) \) (with \( \frac{0}{|0|} = 1 \)). Let \( d\mu(r) = r^{n-1} dr \) and

\[
\sigma^{n-1}(A) = n\lambda^n(f^{-1}([0,1) \times A))
\]

(7.63)

for every \( A \in \mathcal{B}(S^{n-1}) = \mathcal{B}^n \cap S^{n-1} \). Note that \( \sigma^{n-1} \) inherits the rotation invariance from \( \lambda^n \). By Theorem 7.31 it suffices to show \( f_* \lambda^n = \mu \otimes \sigma^{n-1} \). This follows from

\[
(f_* \lambda^n)([0, r) \times A) = \lambda^n(f^{-1}([0, r) \times A)) = r^n \lambda^n(f^{-1}([0,1) \times A)) = \mu([0, r)) \sigma^{n-1}(A).
\]

since these sets determine the measure uniquely. \( \square \)
Example. Let us compute the volume of a ball in $\mathbb{R}^n$:

$$V_n(r) = \int_{\mathbb{R}^n} \chi_{B_r(0)} d^n x.$$ 

By the simple scaling transform $f(x) = rx$ we obtain $V_n(r) = V_n(1)r^n$ and hence it suffices to compute $V_n = V_n(1)$.

To this end we use (Problem 7.26)

$$\pi^n = \int_{\mathbb{R}^n} e^{-|x|^2} d^n x = n V_n \int_0^\infty e^{-r^2 r^{n-1}} dr = \frac{n V_n}{2} \int_0^\infty e^{-s s^{n/2 - 1}} ds$$

$$= \frac{n V_n}{2} \Gamma\left(\frac{n}{2}\right) = \frac{V_n}{2} \Gamma\left(\frac{n}{2} + 1\right)$$

where $\Gamma$ is the gamma function (Problems 7.27). Hence

$$V_n = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)}. \quad (7.64)$$

By $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ (see Problem 7.28) this coincides with the well-known values for $n = 1, 2, 3$.

Example. The above lemma can be used to determine when a radial function is integrable. For example, we obtain

$$|x|^\alpha \in L^1(B_1(0)) \iff \alpha > -n, \quad |x|^\alpha \in L^1(\mathbb{R}^n \setminus B_1(0)) \iff \alpha < -n.$$ 

\\ 

Problem 7.23. Let $\lambda$ be Lebesgue measure on $\mathbb{R}$. Show that if $f \in C^1(\mathbb{R})$ with $f' > 0$, then

$$d(f_\ast \lambda)(x) = \frac{1}{f'(f^{-1}(x))} dx.$$ 

Problem 7.24. Show that every invertible matrix $M$ can be written as $M = O_1 D O_2$, where $D$ is diagonal and $O_j$ are orthogonal. (Hint: The matrix $M^* M$ is nonnegative and hence there is an orthogonal matrix $U$ which diagonalizes $M^* M = U^* D^2 U$. Then one can choose $O_1 = MUD^{-1}$ and $O_2 = U^*$.)

Problem 7.25. Compute $V_n$ using spherical coordinates. (Hint: $\int \sin(x)^n dx = -\frac{1}{n} \sin(x)^{n-1} \cos(x) + \frac{n-1}{n} \int \sin(x)^{n-2} dx$.)

Problem 7.26. Show

$$I_n = \int_{\mathbb{R}^n} e^{-|x|^2} d^n x = \pi^{n/2}.$$ 

(Hint: Use Fubini to show $I_n = I_1^n$ and compute $I_2$ using polar coordinates.)
Problem 7.27. The gamma function is defined via
\[ \Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx, \quad \text{Re}(z) > 0. \] (7.65)
Verify that the integral converges and defines an analytic function in the indicated half-plane (cf. Problem 7.22). Use integration by parts to show
\[ \Gamma(z + 1) = z\Gamma(z), \quad \Gamma(1) = 1. \] (7.66)
Conclude \( \Gamma(n) = (n-1)! \) for \( n \in \mathbb{N} \).

Problem 7.28. Show that \( \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \). (Hint: Use the change of coordinates \( x = t^2 \) and then use Problem 7.26.)

Problem 7.29. Let \( U \subseteq \mathbb{R}^m \) be open and let \( f : U \to \mathbb{R}^n \) be locally Lipschitz (i.e., for every compact set \( K \subset U \) there is some constant \( L \) such that \( |f(x) - f(y)| \leq L|x - y| \) for all \( x, y \in K \)). Show that if \( A \subset U \) has Lebesgue measure zero, then \( f(A) \) is contained in a set of Lebesgue measure zero. (Hint: By Lindelöf it is no restriction to assume that \( A \) is contained in a compact ball contained in \( U \).)

7.8. Appendix: Transformation of Lebesgue–Stieltjes integrals

In this section we will look at Borel measures on \( \mathbb{R} \). In particular, we want to derive a generalized substitution rule.

As a preparation we will need a generalization of the usual inverse which works for arbitrary nondecreasing functions. Such a generalized inverse arises, for example, as quantile functions in probability theory.

So we look at nondecreasing functions \( f : \mathbb{R} \to \mathbb{R} \). By monotonicity the limits from left and right exist at every point and we will denote them by
\[ f(x\pm) = \lim_{\varepsilon \downarrow 0} f(x \pm \varepsilon). \] (7.67)
Clearly we have \( f(x-) \leq f(x+) \) and a strict inequality can occur only at a countable number of points. By monotonicity the value of \( f \) has to lie between these two values \( f(x-) \leq f(x) \leq f(x+) \). It will also be convenient to extend \( f \) to a function on the extended reals \( \mathbb{R} \cup \{-\infty, +\infty\} \). Again by monotonicity the limits \( f(\pm\infty\mp) = \lim_{x \to \pm\infty} f(x) \) exist and we will set \( f(\pm\infty\mp) = \pm\infty \).

If we want to define an inverse, problems will occur at points where \( f \) jumps and on intervals where \( f \) is constant. Formally speaking, if \( f \) jumps, then the corresponding jump will be missing in the domain of the inverse and if \( f \) is constant, the inverse will be multivalued. For the first case there is a natural fix by choosing the inverse to be constant along the missing interval. In particular, observe that this natural choice is independent of the actual
value of $f$ at the jump and hence the inverse loses this information. The second case will result in a jump for the inverse function and here there is no natural choice for the value at the jump (except that it must be between the left and right limits such that the inverse is again a nondecreasing function).

To give a precise definition it will be convenient to look at relations instead of functions. Recall that a (binary) relation $R$ on $\mathbb{R}$ is a subset of $\mathbb{R}^2$.

To every nondecreasing function $f$ associate the relation

$$\Gamma(f) = \{(x, y) | y \in [f(x^-), f(x^+)]\}. \quad (7.68)$$

Note that $\Gamma(f)$ does not depend on the values of $f$ at a discontinuity and $f$ can be partially recovered from $\Gamma(f)$ using $f(x^-) = \inf \Gamma(f)(x)$ and $f(x^+) = \sup \Gamma(f)(x)$, where $\Gamma(f)(x) = \{y | (x, y) \in \Gamma(f)\} = [f(x^-), f(x^+)]$.

Moreover, the relation is nondecreasing in the sense that $x_1 < x_2$ implies $y_1 \leq y_2$ for $(x_1, y_1), (x_2, y_2) \in \Gamma(f)$. It is uniquely defined as the largest relation containing the graph of $f$ with this property.

The graph of any reasonable inverse should be a subset of the inverse relation

$$\Gamma(f)^{-1} = \{(y, x) | (x, y) \in \Gamma(f)\} \quad (7.69)$$

and we will call any function $f^{-1}$ whose graph is a subset of $\Gamma(f)^{-1}$ a generalized inverse of $f$. Note that any generalized inverse is again nondecreasing since a pair of points $(y_1, x_1), (y_2, x_2) \in \Gamma(f)^{-1}$ with $y_1 < y_2$ and $x_1 > x_2$ would contradict the fact that $\Gamma(f)$ is nondecreasing. Moreover, since $\Gamma(f)^{-1}$ and $\Gamma(f^{-1})$ are two nondecreasing relations containing the graph of $f^{-1}$, we conclude

$$\Gamma(f^{-1}) = \Gamma(f)^{-1} \quad (7.70)$$

since both are maximal. In particular, it follows that if $f^{-1}$ is a generalized inverse of $f$ then $f$ is a generalized inverse of $f^{-1}$.

There are two particular choices, namely the left continuous version $f^{-1}_-(y) = \inf \Gamma(f)^{-1}(y)$ and the right continuous version $f^{-1}_+(y) = \sup \Gamma(f)^{-1}(y)$. It is straightforward to verify that they can be equivalently defined via

$$f^{-1}_-(y) = \inf f^{-1}([y, \infty)) = \sup f^{-1}((\infty, y)), \quad f^{-1}_+(y) = \inf f^{-1}((y, \infty)) = \sup f^{-1}((\infty, y)). \quad (7.71)$$

For example, $\inf f^{-1}([y, \infty)) = \inf \{x | (x, \tilde{y}) \in \Gamma(f), \tilde{y} \geq y\} = \inf \Gamma(f)^{-1}(y)$. The first one is typically used in probability theory, where it corresponds to the quantile function of a distribution.

If $f$ is strictly increasing the generalized inverse coincides with the usual inverse and we have $f(f^{-1}(y)) = y$ for $y$ in the range of $f$. The purpose
of the next lemma is to investigate to what extent this remains valid for a
generalized inverse.

Note that for every \( y \) there is some \( x \) with \( y \in [f(x-), f(x+)] \). Moreover,
if we can find two values, say \( x_1 \) and \( x_2 \), with this property, then \( f(x) = y \)
is constant for \( x \in (x_1, x_2) \). Hence, the set of all such \( x \) is an interval which
is closed since at the left, right boundary point the left, right limit equals \( y \),
respectively.

We collect a few simple facts for later use.

**Lemma 7.34.** Let \( f \) be nondecreasing.

(i) \( f^{-1}(y) \leq x \) if and only if \( y \leq f(x+) \).

(ii) \( f^{-1}(y) \geq x \) if and only if \( y \geq f(x-) \).

(iii) \( f(f^{-1}(y)) \geq y \) if \( f \) is right continuous at \( f^{-1}(y) \) with equality if
\( y \in \text{Ran}(f) \).

(iv) \( f(f^{-1}(y)) \leq y \) if \( f \) is left continuous at \( f^{-1}(y) \) with equality if
\( y \in \text{Ran}(f) \).

**Proof.** Item (i) follows since both claims are equivalent to \( y \leq f(x) \) for all \( x > x \). Next, let \( f \) be right-continuous. If \( y \) is in the range of \( f \),
then \( y = f(f^{-1}(y)+) = f(f^{-1}(y)) \) if \( f^{-1}(y) \) contains more than one
point and \( y = f(f^{-1}(y)) \) else. If \( y \) is not in the range of \( m \) we must have
\( y \in [f(x-), f(x+)] \) for \( x = f^{-1}(y) \) which establishes item (ii). Similarly for
(iii) and (iv).

In particular, \( f(f^{-1}(y)) = y \) if \( f \) is continuous. We will also need the
set
\[
L(f) = \{ y | f^{-1}((y, \infty)) = (f^{-1}(y), \infty) \}. 
\]  
(7.72)
Note that \( y \notin L(f) \) if and only if there is some \( x \) such that \( y \in [f(x-), f(x)] \).

**Lemma 7.35.** Let \( m : \mathbb{R} \to \mathbb{R} \) be a nondecreasing function on \( \mathbb{R} \) and \( \mu \) its
associated measure via (7.5). Let \( f(x) \) be a nondecreasing function on \( \mathbb{R} \)
such that \( \mu((0, \infty)) < \infty \) if \( f \) is bounded above and \( \mu((-\infty, 0)) < \infty \) if \( f \) is
bounded below.

Then \( f_\# \mu \) is a Borel measure whose distribution function coincides up
to a constant with \( m_+ \circ f_+^{-1} \) at every point \( y \) which is in \( L(f) \) or satisfies
\( \mu(f_+^{-1}(y)) = 0 \). If \( y \in [f(x-), f(x)] \) and \( \mu(f_+^{-1}(y)) > 0 \), then \( m_+ \circ f_+^{-1} \)
jumps at \( f(x-) \) and \( (f_\# \mu)(y) \) jumps at \( f(x) \).

**Proof.** First of all note that the assumptions in case \( f \) is bounded from
above or below ensure that \( (f_\# \mu)(K) < \infty \) for any compact interval. Moreover,
we can assume \( m = m_+ \) without loss of generality. Now note that we
have $f^{-1}((y, \infty)) = (f^{-1}(y), \infty)$ for $y \in L(f)$ and $f^{-1}((y, \infty)) = [f^{-1}(y), \infty)$ else. Hence

$$(f_*\mu)((y_0, y_1]) = \mu(f^{-1}((y_0, y_1])) = \mu((f^{-1}(y_0), f^{-1}(y_1]))$$

$$= m(f_+(y_1)) - m(f_+(y_0)) = (m \circ f_+^{-1})(y_1) - (m \circ f_+^{-1})(y_0)$$

if $y_j$ is either in $L(f)$ or satisfies $\mu(\{f_+^{-1}(y_j)\}) = 0$. For the last claim observe that $f^{-1}((y, \infty))$ will jump from $(f_+^{-1}(y), \infty)$ to $[f_+^{-1}(y), \infty)$ at $y = f(x)$. \[ \square \]

Example. For example, consider $f(x) = \chi_{[0,\infty)}(x)$ and $\mu = \Theta$, the Dirac measure centered at 0 (note that $\Theta(x) = f(x)$). Then

$$f_+^{-1}(y) = \begin{cases} +\infty, & 1 \leq y, \\ 0, & 0 \leq y < 1, \\ -\infty, & y < 0, \end{cases}$$

and $L(f) = (-\infty, 0) \cup [1, \infty)$. Moreover, $\mu(f_+^{-1}(y)) = \chi_{[0,\infty)}(y)$ and $(f_*\mu)(y) = \chi_{[1,\infty)}(y)$. If we choose $g(x) = \chi_{[0,\infty)}(x)$, then $g_+^{-1}(y) = f_+^{-1}(y)$ and $L(g) = \mathbb{R}$. Hence $\mu(g_+^{-1}(y)) = \chi_{[0,\infty)}(y) = (g_*\mu)(y)$.

For later use it is worth while to single out the following consequence:

**Corollary 7.36.** Let $m$, $f$ be as in the previous lemma and denote by $\mu$, $\nu_\pm$ the measures associated with $m$, $m_+ \circ f^{-1}$, respectively. Then, $(f_+)\mu = \nu_\pm$ and hence

$$\int g \, d(m_+ \circ f^{-1}) = \int (g \circ f_+) \, dm. \quad (7.73)$$

In the special case where $\mu$ is Lebesgue measure this reduces to a way of expressing the Lebesgue–Stieltjes integral as a Lebesgue integral via

$$\int g \, dh = \int g(h^{-1}(y)) \, dy. \quad (7.74)$$

If we choose $f$ to be the distribution function of $\mu$ we get the following generalization of the integration by substitution rule. To formulate it we introduce

$$i_m(y) = m(m_-(y)). \quad (7.75)$$

Note that $i_m(y) = y$ if $m$ is continuous. By $hull(Ran(m))$ we denote the convex hull of the range of $m$.

**Corollary 7.37.** Suppose $m$, $m$ are two nondecreasing functions on $\mathbb{R}$ with $m$ right continuous. Then we have

$$\int_\mathbb{R} (g \circ m) \, d(n \circ m) = \int_{hull(Ran(m))} (g \circ i_m) \, dn \quad (7.76)$$
for any Borel function $g$ which is either nonnegative or for which one of the two integrals is finite. Similarly, if $n$ is left continuous and $i_m$ is replaced by $m(m^{-1}(y))$.

Hence the usual $\int_{\mathbb{R}} (g \circ m) \, d(n \circ m) = \int_{\text{Ran}(m)} g \, dn$ only holds if $m$ is continuous. In fact, the right-hand side loses all point masses of $\mu$. The above formula fixes this problem by rendering $g$ constant along a gap in the range of $m$ and includes the gap in the range of integration such that it makes up for the lost point mass. It should be compared with the previous example!

If one does not want to bother with $i_m$ one can at least get inequalities for monotone $g$.

**Corollary 7.38.** Suppose $m$, $n$ are nondecreasing functions on $\mathbb{R}$ and $g$ is monotone. Then we have

$$\int_{\mathbb{R}} (g \circ m) \, d(n \circ m) \leq \int_{\text{null}(\text{Ran}(m))} g \, dn$$

(7.77)

if $m$, $n$ are right continuous and $g$ nonincreasing or $m$, $n$ left continuous and $g$ nondecreasing. If $m$, $n$ are right continuous and $g$ nondecreasing or $m$, $n$ left continuous and $g$ nonincreasing the inequality has to be reversed.

**Proof.** Immediate from the previous corollary together with $i_m(y) \leq y$ if $y = f(x) = f(x^+)$ and $i_m(y) \geq y$ if $y = f(x) = f(x^-)$ according to Lemma 7.34. \qed

**Problem 7.30.** Show (7.71).

**Problem 7.31.** Show that $\Gamma(f) \circ \Gamma(f^{-1}) = \{(y, z) | y, z \in [f(x^-), f(x^+)] \text{ for some } x\}$.

**Problem 7.32.** Let $d\mu(\lambda) = \chi_{[0,1]}(\lambda)d\lambda$ and $f(\lambda) = \chi_{(t,\infty)}(\lambda)$, $t \in \mathbb{R}$. Compute $f_\ast \mu$.

### 7.9. Appendix: The connection with the Riemann integral

In this section we want to investigate the connection with the Riemann integral. We restrict our attention to compact intervals $[a, b]$ and bounded real-valued functions $f$. A **partition** of $[a, b]$ is a finite set $P = \{x_0, \ldots, x_n\}$ with

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b.$$  

(7.78)

The number

$$\|P\| = \max_{1 \leq j \leq n} x_j - x_{j-1}$$

(7.79)

...
7.9. Appendix: The connection with the Riemann integral

is called the norm of \( P \). Given a partition \( P \) and a bounded real-valued function \( f \) we can define

\[
s_{P,f,-}(x) = \sum_{j=1}^{n} m_{j}\chi_{[x_{j-1},x_{j})}(x), \quad m_{j} = \inf_{x \in [x_{j-1},x_{j})} f(x), \quad (7.80)
\]

\[
s_{P,f,+}(x) = \sum_{j=1}^{n} M_{j}\chi_{[x_{j-1},x_{j})}(x), \quad M_{j} = \sup_{x \in [x_{j-1},x_{j})} f(x), \quad (7.81)
\]

Hence \( s_{P,f,-}(x) \) is a step function approximating \( f \) from below and \( s_{P,f,+}(x) \) is a step function approximating \( f \) from above. In particular,

\[
m \leq s_{P,f,-}(x) \leq f(x) \leq s_{P,f,+}(x) \leq M, \quad m = \inf_{x \in [a,b]} f(x), \quad M = \sup_{x \in [a,b]} f(x). \quad (7.82)
\]

Moreover, we can define the upper and lower Riemann sum associated with \( P \) as

\[
L(P,f) = \sum_{j=1}^{n} m_{j}(x_{j} - x_{j-1}), \quad U(P,f) = \sum_{j=1}^{n} M_{j}(x_{j} - x_{j-1}). \quad (7.83)
\]

Of course, \( L(f,P) \) is just the Lebesgue integral of \( s_{P,f,-} \) and \( U(f,P) \) is the Lebesgue integral of \( s_{P,f,+} \). In particular, \( L(P,f) \) approximates the area under the graph of \( f \) from below and \( U(P,f) \) approximates this area from above.

By the above inequality

\[
m (b - a) \leq L(P,f) \leq U(P,f) \leq M (b - a). \quad (7.84)
\]

We say that the partition \( P_{2} \) is a refinement of \( P_{1} \) if \( P_{1} \subseteq P_{2} \) and it is not hard to check, that in this case

\[
s_{P_{1},f,-}(x) \leq s_{P_{2},f,-}(x) \leq f(x) \leq s_{P_{2},f,+}(x) \leq s_{P_{1},f,+}(x) \quad (7.85)
\]

as well as

\[
L(P_{1},f) \leq L(P_{2},f) \leq U(P_{2},f) \leq U(P_{1},f). \quad (7.86)
\]

Hence we define the **lower, upper Riemann integral** of \( f \) as

\[
\int f(x)dx = \sup_{P} L(P,f), \quad \int f(x)dx = \inf_{P} U(P,f), \quad (7.87)
\]

respectively. Since for arbitrary partitions \( P \) and \( Q \) we have

\[
L(P,f) \leq L(P \cup Q, f) \leq U(P \cup Q, f) \leq U(Q,f). \quad (7.88)
\]

we obtain

\[
m (b - a) \leq \int f(x)dx \leq \int f(x)dx \leq M (b - a). \quad (7.89)
\]
We will call \( f \) **Riemann integrable** if both values coincide and the common value will be called the **Riemann integral** of \( f \).

**Example.** Let \([a, b] = [0, 1]\) and \( f(x) = \chi_Q(x) \). Then \( s_{P,f,-}(x) = 0 \) and \( s_{P,f,+}(x) = 1 \). Hence \( \int f(x)dx = 0 \) and \( \bar{\int} f(x)dx = 1 \) and \( f \) is not Riemann integrable.

On the other hand, every continuous function \( f \in C[a,b] \) is Riemann integrable (Problem 7.33).

**Example.** Let \( f \) non-decreasing, then \( f \) is integrable. In fact, since \( m_j = f(x_{j-1}) \) and \( M_j = f(x_j) \) we obtain

\[
U(f, P) - L(f, P) \leq \|P\| \sum_{j=1}^{n} (f(x_j) - f(x_{j-1})) = \|P\|(f(b) - f(a))
\]

and the claim follows (cf. also the next lemma).

**Lemma 7.39.** A function \( f \) is Riemann integrable if and only if there exists a sequence of partitions \( P_j \) such that

\[
\lim_{n \to \infty} L(P_n, f) = \lim_{n \to \infty} U(P_n, f). \tag{7.90}
\]

In this case the above limits equal the Riemann integral of \( f \) and \( P_n \) can be chosen such that \( P_n \subseteq P_{n+1} \) and \( \|P_n\| \to 0 \).

**Proof.** If there is such a sequence of partitions then \( f \) is integrable by \( \lim_n L(P_n, f) \leq \sup_P L(P, f) \leq \inf_P U(P, f) \leq \lim_n U(P_n, f) \).

Conversely, given an integrable \( f \), there is a sequence of partitions \( P_{L,n} \) such that \( \int f(x)dx = \lim_n L(P_{L,n}, f) \) and a sequence \( P_{U,n} \) such that \( \int f(x)dx = \lim_n U(P_{U,n}, f) \). By (7.86) the common refinement \( P_n = P_{L,n} \cup P_{U,n} \) is the partition we are looking for. Since, again by (7.86), any refinement will also work, the last claim follows.

Note that when computing the Riemann integral as in the previous lemma one could choose instead of \( m_j \) or \( M_j \) any value in \([m_j, M_j]\) (e.g. \( f(x_{j-1}) \) or \( f(x_j) \)).

With the help of this lemma we can give a characterization of Riemann integrable functions and show that the Riemann integral coincides with the Lebesgue integral.

**Theorem 7.40** (Lebesgue). A bounded measurable function \( f : [a,b] \to \mathbb{R} \) is Riemann integrable if and only if the set of its discontinuities is of Lebesgue measure zero. Moreover, in this case the Riemann and the Lebesgue integral of \( f \) coincide.
7.9. Appendix: The connection with the Riemann integral

**Proof.** Suppose \( f \) is Riemann integrable and let \( P_j \) be a sequence of partitions as in Lemma 7.39. Then \( s_{f,P_j,-}(x) \) will be monotone and hence converge to some function \( s_{f,-}(x) \leq f(x) \). Similarly, \( s_{f,P_j,+}(x) \) will converge to some function \( s_{f,+}(x) \geq f(x) \). Moreover, by dominated convergence

\[
0 = \lim_{j} \int (s_{f,P_j,+}(x) - s_{f,P_j,-}(x)) \, dx = \int (s_{f,+}(x) - s_{f,-}(x)) \, dx
\]

and thus by Lemma 7.22 \( s_{f,+}(x) = s_{f,-}(x) \) almost everywhere. Moreover, \( f \) is continuous at every \( x \) at which equality holds and which is not in any of the partitions. Since the first as well as the second set have Lebesgue measure zero, \( f \) is continuous almost everywhere and

\[
\lim_{j} L(P_j, f) = \lim_{j} U(P_j, f) = \int s_{f,\pm}(x) \, dx = \int f(x) \, dx.
\]

Conversely, let \( f \) be continuous almost everywhere and choose some sequence of partitions \( P_j \) with \( \|P_j\| \to 0 \). Then at every \( x \) where \( f \) is continuous we have \( \lim_{j} s_{f,P_j,\pm}(x) = f(x) \) implying

\[
\lim_{j} L(P_j, f) = \int s_{f,-}(x) \, dx = \int f(x) \, dx = \int s_{f,+}(x) \, dx = \lim_{j} U(P_j, f)
\]

by the dominated convergence theorem. \( \square \)

Note that if \( f \) is not assumed to be measurable, the above proof still shows that \( f \) satisfies \( s_{f,-} \leq f \leq s_{f,+} \) for two measurable functions \( s_{f,\pm} \) which are equal almost everywhere. Hence if we replace the Lebesgue measure by its completion, we can drop this assumption.

**Problem 7.33.** Show that for any function \( f \in C[a,b] \) we have

\[
\lim_{\|P\| \to 0} L(P, f) = \lim_{\|P\| \to 0} U(P, f).
\]

In particular, \( f \) is Riemann integrable.

**Problem 7.34.** Prove that the Riemann integral is linear: If \( f, g \) are Riemann integrable and \( \alpha \in \mathbb{R} \), then \( \alpha f \) and \( f + g \) are Riemann integrable with

\[
\int (f + g) \, dx = \int f \, dx + \int g \, dx\quad \text{and}\quad \int \alpha f \, dx = \alpha \int f \, dx.
\]

**Problem 7.35.** Show that if \( f, g \) are Riemann integrable, so is \( fg \).
The Lebesgue spaces

$L^p$

8.1. Functions almost everywhere

We fix some measure space $(X, \Sigma, \mu)$ and define the $L^p$ norm by

$$\|f\|_p = \left( \int_X |f|^p \, d\mu \right)^{1/p}, \quad 1 \leq p,$$

and denote by $L^p(X, d\mu)$ the set of all complex-valued measurable functions for which $\|f\|_p$ is finite. First of all note that $L^p(X, d\mu)$ is a vector space, since $|f + g|^p \leq 2^p \max(|f|^p, |g|^p) = 2^p \max(|f|^p, |g|^p) \leq 2^p(|f|^p + |g|^p)$. Of course our hope is that $L^p(X, d\mu)$ is a Banach space. However, Lemma 7.22 implies that there is a small technical problem (recall that a property is said to hold almost everywhere if the set where it fails to hold is contained in a set of measure zero):

**Lemma 8.1.** Let $f$ be measurable. Then

$$\int_X |f|^p \, d\mu = 0$$

if and only if $f(x) = 0$ almost everywhere with respect to $\mu$.

Thus $\|f\|_p = 0$ only implies $f(x) = 0$ for almost every $x$, but not for all! Hence $\|\|_p$ is not a norm on $L^p(X, d\mu)$. The way out of this misery is to identify functions which are equal almost everywhere: Let

$$N(X, d\mu) = \{f | f(x) = 0 \, \mu\text{-almost everywhere} \}. \quad (8.3)$$
Then $\mathcal{N}(X, d\mu)$ is a linear subspace of $L^p(X, d\mu)$ and we can consider the quotient space

$$L^p(X, d\mu) = L^p(X, d\mu)/\mathcal{N}(X, d\mu).$$

(8.4)

If $d\mu$ is the Lebesgue measure on $X \subseteq \mathbb{R}^n$, we simply write $L^p(X)$. Observe that $\|f\|_p$ is well defined on $L^p(X, d\mu)$ and hence we have a normed space.

Even though the elements of $L^p(X, d\mu)$ are, strictly speaking, equivalence classes of functions, we will still treat them functions for notational convenience. However, if we do so it is important to ensure that every statement made does not depend on the representative in the equivalence classes. In particular, note that for $f \in L^p(X, d\mu)$ the value $f(x)$ is not well defined (unless there is a continuous representative and continuous functions with different values are in different equivalence classes, e.g., in the case of Lebesgue measure).

With this modification we are back in business since $L^p(X, d\mu)$ turns out to be a Banach space. We will show this in the following sections. Moreover, note that $L^2(X, d\mu)$ is a Hilbert space with scalar product given by

$$\langle f, g \rangle = \int_X f(x)^*g(x)d\mu(x).$$

(8.5)

But before that let us also define $L^\infty(X, d\mu)$. It should be the set of bounded measurable functions $B(X)$ together with the sup norm. The only problem is that if we want to identify functions equal almost everywhere, the supremum is no longer independent of the representative in the equivalence class. The solution is the essential supremum

$$\|f\|_\infty = \inf\{C | \mu(\{x| |f(x)| > C\}) = 0\}.\quad (8.6)$$

That is, $C$ is an essential bound if $|f(x)| \leq C$ almost everywhere and the essential supremum is the infimum over all essential bounds.

**Example.** If $\lambda$ is the Lebesgue measure, then the essential sup of $\chi_Q$ with respect to $\lambda$ is 0. If $\Theta$ is the Dirac measure centered at 0, then the essential sup of $\chi_Q$ with respect to $\Theta$ is 1 (since $\chi_Q(0) = 1$, and $x = 0$ is the only point which counts for $\Theta$).

As before we set

$$L^\infty(X, d\mu) = B(X)/\mathcal{N}(X, d\mu)$$

(8.7)

and observe that $\|f\|_\infty$ is independent of the representative from the equivalence class.

If you wonder where the $\infty$ comes from, have a look at Problem 8.2.

If $X$ is a locally compact Hausdorff space (together with the Borel sigma algebra), a function is called **locally integrable** if it is integrable when restricted to any compact subset $K \subseteq X$. The set of all (equivalence classes
of) locally integrable functions will be denoted by $L^1_{\text{loc}}(X,d\mu)$. Of course this definition extends to $L^p$ for any $1 \leq p \leq \infty$.

**Problem 8.1.** Let $\| \cdot \|$ be a seminorm on a vector space $X$. Show that $N = \{ x \in X | \| x \| = 0 \}$ is a vector space. Show that the quotient space $X/N$ is a normed space with norm $\| x + N \| = \| x \|$.

**Problem 8.2.** Suppose $\mu(X) < \infty$. Show that $L^\infty(X,d\mu) \subseteq L^p(X,d\mu)$ and

$$
\lim_{p \to \infty} \| f \|_p = \| f \|_\infty, \quad f \in L^\infty(X,d\mu).
$$

**Problem 8.3.** Construct a function $f \in L^p(0,1)$ which has a singularity at every rational number in $[0,1]$ (such that the essential supremum is infinite on every open subinterval). (Hint: Start with the function $f_0(x) = |x|^{-\alpha}$ which has a single singularity at 0, then $f_j(x) = f_0(x-x_j)$ has a singularity at $x_j$.)

### 8.2. Jensen $\leq$ Hölder $\leq$ Minkowski

As a preparation for proving that $L^p$ is a Banach space, we will need Hölder’s inequality, which plays a central role in the theory of $L^p$ spaces. In particular, it will imply Minkowski’s inequality, which is just the triangle inequality for $L^p$. Our proof is based on Jensen’s inequality and emphasizes the connection with convexity. In fact, the triangle inequality just states that a norm is convex:

$$
\| \lambda f + (1 - \lambda)g \| \leq \lambda \| f \| + (1 - \lambda)\| g \|, \quad \lambda \in (0,1).
$$

(8.8)

Recall that a real function $\varphi$ defined on an open interval $(a,b)$ is called **convex** if

$$
\varphi((1 - \lambda)x + \lambda y) \leq (1 - \lambda)\varphi(x) + \lambda \varphi(y), \quad \lambda \in (0,1),
$$

(8.9)

that is, on $(x,y)$ the graph of $\varphi(x)$ lies below or on the line connecting $(x,\varphi(x))$ and $(y,\varphi(y))$:

![Graph of a convex function](image)

If the inequality is strict, then $\varphi$ is called **strictly convex**. It is not hard to see (use $z = (1 - \lambda)x + \lambda y$) that the definition implies

$$
\frac{\varphi(z) - \varphi(x)}{z-x} \leq \frac{\varphi(y) - \varphi(x)}{y-x} \leq \frac{\varphi(y) - \varphi(z)}{y-z}, \quad x < z < y.
$$

(8.10)
where the inequalities are strict if \( \varphi \) is strictly convex.

**Lemma 8.2.** Let \( \varphi : (a, b) \to \mathbb{R} \) be convex. Then

1. \( \varphi \) is locally Lipschitz continuous.
2. The left/right derivatives \( \varphi'_{\pm}(x) = \lim_{\varepsilon \downarrow 0} \frac{\varphi(x \pm \varepsilon) - \varphi(x)}{\pm \varepsilon} \) exist and are monotone nondecreasing. Moreover, \( \varphi' \) exists except at a countable number of points.
3. For fixed \( x \) we have \( \varphi(y) \geq \varphi(x) + \alpha(y - x) \) for every \( \alpha \) with \( \varphi'(x) \leq \alpha \leq \varphi'(y) \). The inequality is strict for \( y \neq x \) if \( \varphi \) is strictly convex.

**Proof.** Abbreviate \( D(x, y) = D(y, x) = \frac{\varphi(y) - \varphi(x)}{y - x} \) and observe that (8.10) implies

\[
D(x, z) \leq D(y, z) \quad \text{for } x < z < y.
\]

Hence \( \varphi'_{\pm}(x) \) exist and we have \( \varphi'_-(x) \leq \varphi'_+(x) \leq \varphi'_-(y) \leq \varphi'_+(y) \) for \( x < y \). So (ii) follows after observing that a monotone function can have at most a countable number of jumps. Next

\[
\varphi'_+(x) \leq D(y, x) \leq \varphi'_-(y)
\]

shows \( \varphi(y) \geq \varphi(x) + \varphi'_+(x)(y - x) \) if \( \pm(y - x) > 0 \) and proves (iii). Moreover, \( \varphi'_+(z) \leq |D(y, x)| \leq \varphi'_-(z) \) for \( z < x, y < z \) proves (i).

**Remark:** It is not hard to see that \( \varphi \in C^1 \) is convex if and only if \( \varphi'(x) \) is monotone nondecreasing (e.g., \( \varphi'' \geq 0 \) if \( \varphi \in C^2 \)).

With these preparations out of the way we can show

**Theorem 8.3** (Jensen’s inequality). Let \( \varphi : (a, b) \to \mathbb{R} \) be convex (\( a = -\infty \) or \( b = \infty \) being allowed). Suppose \( \mu \) is a finite measure satisfying \( \mu(X) = 1 \) and \( f \in L^1(X, d\mu) \) with \( a < f(x) < b \). Then the negative part of \( \varphi \circ f \) is integrable and

\[
\varphi\left(\int_X f \, d\mu\right) \leq \int_X (\varphi \circ f) \, d\mu.
\]

(8.11)

For \( \varphi \geq 0 \) nondecreasing and \( f \geq 0 \) the requirement that \( f \) is integrable can be dropped if \( \varphi(b) \) is understood as \( \lim_{x \to b} \varphi(x) \).

**Proof.** By (iii) of the previous lemma we have

\[
\varphi(f(x)) \geq \varphi(I) + \alpha(f(x) - I), \quad I = \int_X f \, d\mu \in (a, b).
\]

This shows that the negative part of \( \varphi \circ f \) is integrable and integrating over \( X \) finishes the proof in the case \( f \in L^1 \). If \( f \geq 0 \) we note that for \( X_n = \{ x \in X | f(x) \leq n \} \) the first part implies

\[
\frac{1}{\mu(X_n)} \int_{X_n} f \, d\mu \leq \frac{1}{\mu(X_n)} \int_{X_n} \varphi(f) \, d\mu.
\]
Taking \( n \to \infty \) the claim follows from \( X_n \nearrow X \) and the monotone convergence theorem. \( \square \)

Observe that if \( \varphi \) is strictly convex, then equality can only occur if \( f \) is constant.

Now we are ready to prove

**Theorem 8.4** (Hölder’s inequality). Let \( p \) and \( q \) be dual indices; that is,
\[
\frac{1}{p} + \frac{1}{q} = 1 \tag{8.12}
\]
with \( 1 \leq p \leq \infty \). If \( f \in L^p(X, d\mu) \) and \( g \in L^q(X, d\mu) \), then \( fg \in L^1(X, d\mu) \) and
\[
\|fg\|_1 \leq \|f\|_p \|g\|_q. \tag{8.13}
\]

**Proof.** The case \( p = 1, q = \infty \) (respectively \( p = \infty, q = 1 \)) follows directly from the properties of the integral and hence it remains to consider \( 1 < p, q < \infty \).

First of all it is no restriction to assume \( \|g\|_q = 1 \). Let \( A = \{x | g(x) > 0\} \), then (note \( (1 - q)p = -q \))
\[
\|fg\|_1^p = \int_A |f| |g|^{1-q}|g|^qd\mu \leq \int_A (|f| |g|^{1-q})^p|g|^qd\mu = \int_A |f|^p d\mu \leq \|f\|_p^p,
\]
where we have used Jensen’s inequality with \( \varphi(x) = |x|^p \) applied to the function \( h = |f| |g|^{1-q} \) and measure \( d\nu = |g|^qd\mu \) (note \( \nu(X) = \int |g|^qd\mu = \|g\|_q^q = 1 \)). \( \square \)

Note that in the special case \( p = 2 \) we have \( q = 2 \) and Hölder’s inequality reduces to the Cauchy–Schwarz inequality. Moreover, in the case \( 1 < p < \infty \) the function \( x^p \) is strictly convex and equality will only occur if \( |f| \) is a multiple of \( |g|^{q-1} \) or \( g \) is trivial. In fact, this last observation gives us the following useful characterization of norms.

**Lemma 8.5.** Consider \( L^p(X, d\mu) \) and let \( q \) be the corresponding dual index, \( \frac{1}{p} + \frac{1}{q} = 1 \). Then for every measurable function \( f \)
\[
\|f\|_p = \sup_{s \text{ simple, } \|s\|_q = 1} \left| \int_X fs d\mu \right|
\]
if \( 1 \leq p < \infty \). If \( \mu \) is \( \sigma \)-finite the claim also holds for \( p = \infty \).

**Proof.** We begin with the case \( 1 \leq p < \infty \). Choosing \( h = |f|^{p-1} \text{sign}(f^*) \) (where \( |f|^0 = 1 \) in the case \( p = 1 \)) it follows that \( h \in L^q \) and \( g = h/\|h\|_q \) satisfies \( \int_X fg d\mu = \|f\|_p \). Now since the simple functions are dense in \( L^q \) there is a sequence of simple functions \( s_n \to g \) in \( L^q \) and Hölder’s inequality implies \( \int_X fs_n d\mu \to \int_X fg d\mu \).
Now let us turn to the case \( p = \infty \). For every \( \varepsilon > 0 \) the set \( A_\varepsilon = \{ x \mid |f(x)| \geq \|f\|_\infty - \varepsilon \} \) has positive measure. Moreover, considering \( X_n \uparrow X \) with \( \mu(X_n) < \infty \) there must be some \( n \) such that \( B_\varepsilon = A_\varepsilon \cap X_n \) satisfies \( 0 < \mu(B_\varepsilon) < \infty \). Then \( g_\varepsilon = \text{sign}(f^*)\chi_{B_\varepsilon}/\mu(B_\varepsilon) \) satisfies \( \int_X fg_\varepsilon \, d\mu \geq \|f\|_\infty - \varepsilon \).

Finally, choosing a sequence of simple functions \( s_n \to g_\varepsilon \) in \( L^1 \) finishes the proof. \( \square \)

Note that in the case \( p = \infty \) it suffices to assume that for every set \( A \) of positive measure, there is a subset \( B \subseteq A \) with \( 0 < \mu(B) < \infty \). Moreover, without this assumption the claim is clearly false since there will be no simple functions \( s \) supported on \( A \) with \( \|s\|_1 = 1 \).

As another consequence of Hölder’s inequality we also get

**Theorem 8.6** (Minkowski’s integral inequality). Suppose, \( \mu \) and \( \nu \) are two \( \sigma \)-finite measures and \( f \) is \( \mu \otimes \nu \)-measurable. Let \( 1 \leq p \leq \infty \). Then

\[
\left\| \int_Y f(.,y) d\nu(y) \right\|_p \leq \int_Y \|f(.,y)\|_p d\nu(y),
\]

where the \( p \)-norm is computed with respect to \( \mu \).

**Proof.** Let \( g \in L^q(X,d\mu) \) with \( \|g\|_q = 1 \). Then using Fubini

\[
\int_X g(x) \int_Y f(x,y) d\nu(y) d\mu(x) = \int_Y \int_X f(x,y) g(x) d\mu(x) d\nu(y) \\
\leq \int_Y \|f(.,y)\|_p d\nu(y)
\]

and the claim follows from Lemma 8.5. \( \square \)

In the special case where \( \nu \) is supported on two points this reduces to the triangle inequality (note that in this case the use of Fubini in the proof reduces to linearity of the integral and hence the assumption that \( \mu \) is \( \sigma \)-finite is not needed).

**Corollary 8.7** (Minkowski’s inequality). Let \( f, g \in L^p(X,d\mu) \), \( 1 \leq p \leq \infty \). Then

\[
\|f + g\|_p \leq \|f\|_p + \|g\|_p.
\]

This shows that \( L^p(X,d\mu) \) is a normed vector space.

Note that Fatou’s lemma implies that the norm is lower semi continuous \( \|f\|_p \leq \liminf_{n \to \infty} \|f_n\|_p \) with respect to pointwise convergence (a.e.). The next lemma sheds some light on the missing part.

**Lemma 8.8** (Brezis–Lieb). Let \( 1 \leq p < \infty \) and let \( f_n \in L^p(X,d\mu) \) be a sequence which converges pointwise a.e. to \( f \) such that \( \|f_n\|_p \leq C \). Then
\[ f \in L^p(X, d\mu) \text{ and} \]
\[ \lim_{n \to \infty} (\|f_n\|_p^p - \|f_n - f\|_p^p) = \|f\|_p^p. \tag{8.16} \]

**Proof.** As pointed out before \( \|f\|_p \leq \liminf_{n \to \infty} \|f_n\|_p \leq C \) which shows \( f \in L^p(X, d\mu) \). Moreover, one easy checks the elementary inequality
\[ |s-t|^p - |t|^p - |s|^p \leq \varepsilon |t|^p + C^p|s|^p \]
(note that by scaling it suffices to consider the case \( s = 1 \)). Setting \( t = f - f_n \) and \( s = f \), bringing everything to the right-hand-side and applying Fatou gives
\[ C^p|f|^p_\| \leq \liminf_{n \to \infty} \int_X (\varepsilon |f_n - f|^p + C|f|^p - |f_n|^p - |f - f_n|^p) \, d\mu \]
\[ \leq \varepsilon (2C)^p + C^p|f|^p_\| \limsup_{n \to \infty} \int_X |f_n|^p - |f - f_n|^p - |f|^p| \, d\mu. \]

Since \( \varepsilon > 0 \) is arbitrary the claim follows. \( \square \)

It might be more descriptive to write the conclusion of the lemma as
\[ \|f_n\|_p^p = \|f\|_p^p + \|f_n - f\|_p^p + o(1) \tag{8.17} \]
which shows an important consequence:

**Corollary 8.9.** Let \( 1 \leq p < \infty \) and let \( f_n \in L^p(X, d\mu) \) be a sequence which converges pointwise a.e. to \( f \) such that \( \|f_n\|_p \leq C \). Then \( \|f_n - f\|_p \to 0 \) if and only if \( \|f_n\|_p \to \|f\|_p \).

Note that it even suffices to show \( \limsup \|f_n\|_p \leq \|f\|_p \) since \( \|f\|_p \leq \liminf \|f_n\|_p \) comes for free from Fubini as pointed out before.

**Problem 8.4.** Prove
\[ \prod_{k=1}^n x_k^{\alpha_k} \leq \sum_{k=1}^n \alpha_k x_k, \quad \text{if} \sum_{k=1}^n \alpha_k = 1, \tag{8.18} \]
for \( \alpha_k > 0, x_k > 0 \). (Hint: Take a sum of Dirac-measures and use that the exponential function is convex.)

**Problem 8.5.** Show the generalized Hölder’s inequality:
\[ \|fg\|_r \leq \|f\|_p \|g\|_q, \quad \frac{1}{p} + \frac{1}{q} = \frac{1}{r}. \tag{8.19} \]

**Problem 8.6.** Show the iterated Hölder’s inequality:
\[ \|f_1 \cdots f_m\|_r \leq \prod_{j=1}^m \|f_j\|_{p_j}, \quad \frac{1}{p_1} + \cdots + \frac{1}{p_m} = \frac{1}{r}. \tag{8.20} \]
Problem 8.7. Suppose \( \mu \) is finite. Show that \( L^{p_0} \subseteq L^p \) and
\[
\|f\|_{p_0} \leq \mu(X)^{\frac{1}{p_0} - \frac{1}{p}} \|f\|_p, \quad 1 \leq p_0 \leq p.
\]
(Hint: Hölder’s inequality.)

8.3. Nothing missing in \( L^p \)

Finally it remains to show that \( L^p(X, d\mu) \) is complete.

**Theorem 8.10** (Riesz–Fischer). The space \( L^p(X, d\mu) \), \( 1 \leq p \leq \infty \), is a Banach space.

**Proof.** We begin with the case \( 1 \leq p < \infty \). Suppose \( f_n \) is a Cauchy sequence. It suffices to show that some subsequence converges (show this). Hence we can drop some terms such that
\[
\|f_{n+1} - f_n\|_p \leq \frac{1}{2^n}.
\]
Now consider \( g_n = f_n - f_{n-1} \) (set \( f_0 = 0 \)). Then
\[
G(x) = \sum_{k=1}^{\infty} |g_k(x)|
\]
is in \( L^p \). This follows from
\[
\left\| \sum_{k=1}^{n} |g_k| \right\|_p \leq \sum_{k=1}^{n} \|g_k\|_p \leq \|f_1\|_p + 1
\]
using the monotone convergence theorem. In particular, \( G(x) < \infty \) almost everywhere and the sum
\[
\sum_{n=1}^{\infty} g_n(x) = \lim_{n \to \infty} f_n(x)
\]
is absolutely convergent for those \( x \). Now let \( f(x) \) be this limit. Since \( |f(x) - f_n(x)|^p \) converges to zero almost everywhere and \( |f(x) - f_n(x)|^p \leq (2G(x))^p \in L^1 \), dominated convergence shows \( \|f - f_n\|_p \to 0 \).

In the case \( p = \infty \) note that the Cauchy sequence property \( |f_n(x) - f_m(x)| < \varepsilon \) for \( n, m > N \) holds except for sets \( A_{m,n} \) of measure zero. Since \( A = \bigcup_{n,m} A_{m,n} \) is again of measure zero, we see that \( f_n(x) \) is a Cauchy sequence for \( x \in X \setminus A \). The pointwise limit \( f(x) = \lim_{n \to \infty} f_n(x), \ x \in X \setminus A \), is the required limit in \( L^\infty(X, d\mu) \) (show this). \( \square \)

In particular, in the proof of the last theorem we have seen:

**Corollary 8.11.** If \( \|f_n - f\|_p \to 0 \), then there is a subsequence (of representatives) which converges pointwise almost everywhere.
Consequently, if \( f_n \in L^{p_0} \cap L^{p_1} \) converges in both \( L^{p_0} \) and \( L^{p_1} \), then the limits will be equal a.e. Be warned that the statement is not true in general without passing to a subsequence (Problem 8.8).

It even turns out that \( L^p \) is separable.

**Lemma 8.12.** Suppose \( X \) is a second countable topological space (i.e., it has a countable basis) and \( \mu \) is an outer regular Borel measure. Then \( L^p(X, d\mu) \), \( 1 \leq p < \infty \), is separable. In particular, for every countable base which is closed under finite unions the set of characteristic functions \( \chi_O(x) \) with \( O \) in this base is total.

**Proof.** The set of all characteristic functions \( \chi_A(x) \) with \( A \in \Sigma \) and \( \mu(A) < \infty \) is total by construction of the integral (Problem 8.10)). Now our strategy is as follows: Using outer regularity, we can restrict \( A \) to open sets and using the existence of a countable base, we can restrict \( A \) to open sets from this base.

Fix \( A \). By outer regularity, there is a decreasing sequence of open sets \( O_n \supseteq A \) such that \( \mu(O_n) \to \mu(A) \). Since \( \mu(A) < \infty \), it is no restriction to assume \( \mu(O_n) < \infty \), and thus \( \| \chi_A - \chi_{O_n} \|_p = \mu(O_n \setminus A) = \mu(O_n) - \mu(A) \to 0 \). Thus the set of all characteristic functions \( \chi_O(x) \) with \( O \) open and \( \mu(O) < \infty \) is total. Finally let \( \mathcal{B} \) be a countable base for the topology. Then, every open set \( O \) can be written as \( O = \bigcup_{j=1}^{\infty} \tilde{O}_j \) with \( \tilde{O}_j \in \mathcal{B} \). Moreover, by considering the set of all finite unions of elements from \( \mathcal{B} \), it is no restriction to assume \( \bigcup_{j=1}^{m} \tilde{O}_j \supset O \) with \( \tilde{O}_n \in \mathcal{B} \). By monotone convergence, \( \| \chi_O - \chi_{\tilde{O}_n} \|_p \to 0 \) and hence the set of all characteristic functions \( \chi_{\tilde{O}} \) with \( \tilde{O} \in \mathcal{B} \) is total. \( \square \)

**Problem 8.8.** Find a sequence \( f_n \) which converges to 0 in \( L^p(0, 1) \), \( 1 \leq p < \infty \), but for which \( f_n(x) \to 0 \) for a.e. \( x \in (0, 1) \) does not hold. (Hint: Every \( n \in \mathbb{N} \) can be uniquely written as \( n = 2^m + k \) with \( 0 \leq m \) and \( 0 \leq k < 2^m \). Now consider the characteristic functions of the intervals \( I_{m,k} = [k2^{-m}, (k + 1)2^{-m}] \).)

**Problem 8.9.** Let \( \mu_j \) be \( \sigma \)-finite regular Borel measures on some second countable topological spaces \( X_j \), \( j = 1, 2 \). Show that the set of characteristic functions \( \chi_{A_1 \times A_2} \) with \( A_j \) Borel sets is total in \( L^p(X_1 \times X_2, d(\mu_1 \otimes \mu_2)) \) for \( 1 \leq p < \infty \). (Hint: Problem 7.21 and Lemma 8.12.)

### 8.4. Approximation by nicer functions

Since measurable functions can be quite wild they are sometimes hard to work with. In fact, in many situations some properties are much easier to prove for a dense set of nice functions and the general case can then be reduced to the nice case by an approximation argument. But for such a
strategy to work one needs to identify suitable sets of nice functions which are dense in $L^p$.

**Theorem 8.13.** Let $X$ be a locally compact metric space and let $\mu$ be a regular Borel measure. Then the set $C_c(X)$ of continuous functions with compact support is dense in $L^p(X, d\mu)$, $1 \leq p < \infty$.

**Proof.** As in the proof of Lemma 8.12 the set of all characteristic functions $\chi_K(x)$ with $K$ compact is total (using inner regularity). Hence it suffices to show that $\chi_K(x)$ can be approximated by continuous functions. By outer regularity there is an open set $O \supset K$ such that $\mu(O \setminus K) \leq \varepsilon$. By Urysohn’s lemma (Lemma 1.21) there is a continuous function $f_\varepsilon : X \to [0, 1]$ with compact support which is 1 on $K$ and 0 outside $O$. Since

$$\int_X |\chi_K - f_\varepsilon|^p d\mu = \int_{O \setminus K} |f_\varepsilon|^p d\mu \leq \mu(O \setminus K) \leq \varepsilon,$$

we have $\|f_\varepsilon - \chi_K\|_p \to 0$ and we are done. \qed

Clearly this result has to fail in the case $p = \infty$ (in general) since the uniform limit of continuous functions is again continuous. In fact, the closure of $C_c(\mathbb{R}^n)$ in the infinity norm is the space $C_0(\mathbb{R}^n)$ of continuous functions vanishing at infinity (Problem 1.42).

If $X$ is some subset of $\mathbb{R}^n$, we can do even better and approximate integrable functions by smooth functions. The idea is to replace the value $f(x)$ by a suitable average computed from the values in a neighborhood. This is done by choosing a nonnegative bump function $\phi$, whose area is normalized to 1, and considering the **convolution**

$$(\phi \ast f)(x) = \int_{\mathbb{R}^n} \phi(x - y)f(y)dy = \int_{\mathbb{R}^n} \phi(y)f(x - y)dy. \quad (8.21)$$

For example, if we choose $\phi_r = |B_r(0)|^{-1} \chi_{B_r(0)}$ to be the characteristic function of a ball centered at 0, then $(\phi_r \ast f)(x)$ will be precisely the average of the values of $f$ in the ball $B_r(x)$. In the general case we can think of $(\phi \ast f)(x)$ as an weighted average. Moreover, if we choose $\phi$ differentiable, we can interchange differentiation and integration to conclude that $\phi \ast f$ will also be differentiable. Iterating this argument shows that $\phi \ast f$ will have as many derivatives as $\phi$. Finally, if the set over which the average is computed (i.e., the support of $\phi$) shrinks, we expect $(\phi \ast f)(x)$ to get closer and closer to $f(x)$.

To make these ideas precise we begin with a few properties of the convolution.

**Lemma 8.14.** The convolution has the following properties:
8.4. Approximation by nicer functions

(i) If \( f(x - .)g(.) \) is integrable if and only if \( f(.)g(x - .) \) is and
\[
(f * g)(x) = (g * f)(x)
\]
in this case.

(ii) Suppose \( \phi \in C^k_c(\mathbb{R}^n) \) and \( f \in L^1_{\text{loc}}(\mathbb{R}^n) \), then \( \phi \ast f \in C^k(\mathbb{R}^n) \) and
\[
\partial_\alpha(\phi \ast f) = (\partial_\alpha \phi) \ast f
\]
for any partial derivative of order at most \( k \).

(iii) Suppose \( \phi \in C^k_c(\mathbb{R}^n) \) and \( f \in L^1_c(\mathbb{R}^n) \), then \( \phi \ast f \in C^k_c(\mathbb{R}^n) \).

(iv) Suppose \( \phi \in L^1(\mathbb{R}^n) \) and \( f \in L^p(\mathbb{R}^n) \), \( 1 \leq p \leq \infty \), then their convolution is in \( L^p(\mathbb{R}^n) \) and satisfies Young’s inequality
\[
\|\phi \ast f\|_p \leq \|\phi\|_1 \|f\|_p.
\]

Proof. (i) is a simple affine change of coordinates. (ii) This follows by interchanging differentiation with the integral using Problems 7.19 and 7.20. (iii) By the previous item it suffices to observe that if the support of \( f \) and \( \phi \) are within a ball of radius \( R \), then the support of \( \phi \ast f \) will be within a ball of radius \( 2R \). (iv) The case \( p = \infty \) follows from Hölder’s inequality and we can assume \( 1 \leq p < \infty \). Without loss of generality let \( \|\phi\|_1 = 1 \). Then
\[
\|\phi \ast f\|_p \leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \left| f(y - x) |\phi(y)| d^n y \right|^p d^n x
\]
\[
\leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(y - x)|^p |\phi(y)| d^n y d^n x = \|f\|_p^p,
\]
where we have use Jensen’s inequality with \( \varphi(x) = |x|^p, d\mu = |\phi|d^n y \) in the first and Fubini in the second step. \( \square \)

Next we turn to approximation of \( f \). To this end we call a family of integrable functions \( \phi_\varepsilon, \varepsilon \in (0, 1] \), an approximate identity if it satisfies the following three requirements:

(i) \( \|\phi_\varepsilon\|_1 \leq C \) for all \( \varepsilon > 0 \).

(ii) \( \int_{\mathbb{R}^n} \phi_\varepsilon(x) d^n x = 1 \) for all \( \varepsilon > 0 \).

(iii) For every \( r > 0 \) we have \( \lim_{\varepsilon \downarrow 0} \int_{|x| \geq r} \phi_\varepsilon(x) d^n x = 0 \).

Moreover, a nonnegative function \( \phi \in C^\infty_c(\mathbb{R}^n) \) satisfying \( \|\phi\|_1 = 1 \) is called mollifier.

Example. The standard mollifier is \( \phi(x) = \exp(\frac{1}{|x|^2 - 1}) \) for \( |x| < 1 \) and \( \phi(x) = 0 \) otherwise. To show that this function is indeed smooth it suffices to show that all left derivatives of \( f(r) = \exp(\frac{1}{r - 1}) \) at \( r = 1 \) vanish, which can be done using l’Hôpital’s rule. \( \diamond \)
Example. Scaling a mollifier according to \( \phi_\varepsilon(x) = \varepsilon^{-n}\phi(\frac{x}{\varepsilon}) \) such that its mass is preserved (\( \|\phi_\varepsilon\|_1 = 1 \)) and it concentrates more and more around the origin as \( \varepsilon \downarrow 0 \) we obtain an approximate identity:

\[
\phi_\varepsilon
\]

In fact, (i), (ii) are obvious from \( \|\phi_\varepsilon\|_1 = 1 \) and the integral in (iii) will be identically zero for \( \varepsilon \geq \frac{r}{s} \), where \( s \) is chosen such that \( \text{supp} \phi \subseteq B_s(0) \).

Now we are ready to show that an approximate identity deserves its name.

Lemma 8.15. Let \( \phi_\varepsilon \) be an approximate identity. If \( f \in L^p(\mathbb{R}^n) \) with \( 1 \leq p < \infty \). Then

\[
\lim_{\varepsilon \downarrow 0} \phi_\varepsilon * f = f
\]

with the limit taken in \( L^p \). In the case \( p = \infty \) the claim holds for \( f \in C_0(\mathbb{R}^n) \).

Proof. We begin with the case where \( f \in C_c(\mathbb{R}^n) \). Fix some small \( \delta > 0 \). Since \( f \) is uniformly continuous we know \( |f(x - y) - f(x)| \to 0 \) as \( y \to 0 \) uniformly in \( x \). Since the support of \( f \) is compact, this remains true when taking the \( L^p \) norm and thus we can find some \( r \) such that

\[
\|f(\cdot - y) - f(\cdot)\|_p \leq \frac{\delta}{2C}, \quad |y| \leq r.
\]

(Here the \( C \) is the one for which \( \|\phi_\varepsilon\|_1 \leq C \) holds.) Now we use

\[
(\phi_\varepsilon * f)(x) - f(x) = \int_{\mathbb{R}^n} \phi_\varepsilon(y)(f(x - y) - f(x))d^ny
\]

Splitting the domain of integration according to \( \mathbb{R}^n = \{ y | |y| \leq r \} \cup \{ y | |y| > r \} \), we can estimate the \( L^p \) norms of the individual integrals using the Minkowski inequality as follows:

\[
\left\| \int_{|y| \leq r} \phi_\varepsilon(y)(f(x - y) - f(x))d^ny \right\|_p \leq \frac{\delta}{2}
\]

\[
\int_{|y| \leq r} |\phi_\varepsilon(y)||f(\cdot - y) - f(\cdot)||_p d^ny \leq \frac{\delta}{2}
\]
and
\[
\left\| \int_{|y|>r} \phi_\varepsilon(y)(f(x-y)-f(x))d^n y \right\|_p \leq 2\|f\|_p \int_{|y|>r} |\phi_\varepsilon(y)|d^n y \leq \frac{\delta}{2}
\]
provided \( \varepsilon \) is sufficiently small such that the integral in (iii) is less than \( \delta/2 \).

This establishes the claim for \( f \in C_c(\mathbb{R}^n) \). Since these functions are dense in \( L^p \) for \( 1 \leq p < \infty \) and in \( C_0(\mathbb{R}^n) \) for \( p = \infty \) the claim follows from Lemma 4.31 and Young’s inequality.

Note that in case of a mollifier with support in \( B_r(0) \) this result implies a corresponding local version since the value of \( (\phi_\varepsilon * f)(x) \) is only affected by the values of \( f \) on \( B_{\varepsilon r}(x) \). The question when the pointwise limit exists will be addressed in Problem 8.13.

Now we are ready to prove

**Theorem 8.16.** If \( X \subseteq \mathbb{R}^n \) is open and \( \mu \) is a regular Borel measure, then the set \( C_c^\infty(X) \) of all smooth functions with compact support is dense in \( L^p(X,d\mu) \), \( 1 \leq p < \infty \).

**Proof.** By Theorem 8.13 it suffices to show that every continuous function \( f(x) \) with compact support can be approximated by smooth ones. By setting \( f(x) = 0 \) for \( x \not\in X \), it is no restriction to assume \( X = \mathbb{R}^n \). Now choose a mollifier \( \phi \) and observe that \( \phi_\varepsilon * f \) has compact support (since \( f \) has). Moreover, \( \phi_\varepsilon * f \to f \) in \( L^p \) by the previous lemma.

**Lemma 8.17.** Suppose \( f \in L^1_{\text{loc}}(\mathbb{R}^n) \). Then
\[
\int_{\mathbb{R}^n} \varphi(x) f(x) d^n x = 0, \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n),
\]
if and only if \( f(x) = 0 \) (a.e.).

**Proof.** Let \( K \) be a compact set and \( g = \chi_K \text{sign}(f)^* \). Let \( \phi \) be a mollifier and consider \( g_\varepsilon = \phi_\varepsilon * g \). Since \( g_\varepsilon \to g \) in \( L^p \), there is a subsequence which converges pointwise. Moreover, \( \|g_\varepsilon\|_\infty \leq \|g\|_\infty = 1 \) by Young’s inequality and thus dominated convergence implies
\[
0 = \lim_{\varepsilon \downarrow 0} \int g_\varepsilon f d^n x = \int g f d^n x = \int_K |f| d^n x.
\]

**Problem 8.10.** Show that for any \( f \in L^p(X,d\mu) \), \( 1 \leq p < \infty \) there exists a sequence of integrable simple functions \( s_n \) such that \( |s_n| \leq |f| \) and \( s_n \to f \) in \( L^p(X,d\mu) \). Show that for \( p = \infty \) this only holds if the measure is finite.

(Hint: Split \( f \) into the sum of four nonnegative functions and use (7.26).)
Problem 8.11. Let $\phi$ be integrable and normalized such that $\int_{\mathbb{R}^n} \phi(x) d^nx = 1$. Show that $\phi_\varepsilon(x) = \varepsilon^{-n} \phi(\varepsilon x)$ is an approximate identity.

Problem 8.12. Show that the Poisson kernel

$$P_\varepsilon(x) = \frac{1}{\pi} \frac{\varepsilon}{x^2 + \varepsilon^2}$$

is an approximate identity on $\mathbb{R}$.

Show that the Cauchy transform

$$F(z) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{f(\lambda)}{\lambda - z} d\lambda$$

of a function $f \in L^p(\mathbb{R})$ is analytic in the upper half-plane with imaginary part given by

$$\text{Im}(F(x + iy)) = (P_y * f)(x).$$

In particular, by Young’s inequality $\|\text{Im}(F(\cdot + iy))\|_p \leq \|f\|_p$ and thus $\sup_{y>0} \|\text{Im}(F(\cdot + iy))\|_p = \|f\|_p$. Such analytic functions are said to be in the Hardy space $H^p(\mathbb{C}_+)$.  

(Hint: To see analyticity of $F$ use Problem 7.22 plus the estimate

$$\left| \frac{1}{\lambda - z} \right| \leq \frac{1}{1 + |\lambda| |\text{Im}(z)|}.)$$

Problem 8.13. Let $\phi$ be bounded with support in $\overline{B_1(0)}$ and normalized such that $\int_{\mathbb{R}^n} \phi(x) d^nx = 1$. Set $\phi_\varepsilon(x) = \varepsilon^{-n} \phi(\varepsilon x)$.

For $f$ locally integrable show

$$|(\phi_\varepsilon * f)(x) - f(x)| \leq \frac{V_n}{|B_\varepsilon(x)|} \int_{B_\varepsilon(x)} |f(y) - f(x)| d^ny.$$

Hence at every Lebesgue point (cf. Theorem 9.6) $x$ we have

$$\lim_{\varepsilon \downarrow 0} (\phi_\varepsilon * f)(x) = f(x). \quad (8.27)$$

If $f$ is uniformly continuous then the above limit will be uniform.

8.5. Integral operators

Using Hölder’s inequality, we can also identify a class of bounded operators from $L^p(Y, d\nu)$ to $L^p(X, d\mu)$.

Lemma 8.18 (Schur criterion). Let $\mu, \nu$ be $\sigma$-finite measures on $X, Y$, respectively, and let $\frac{1}{p} + \frac{1}{q} = 1$. Suppose that $K(x, y)$ is measurable and there are nonnegative measurable functions $K_1(x, y), K_2(x, y)$ such that $|K(x, y)| \leq K_1(x, y) K_2(x, y)$ and

$$\|K_1(x, \cdot)\|_{L^q(Y, d\nu)} \leq C_1, \quad \|K_2(\cdot, y)\|_{L^p(X, d\mu)} \leq C_2 \quad (8.28)$$
for \(\mu\)-almost every \(x\), respectively, for \(\nu\)-almost every \(y\). Then the operator
\[ K : L^p(Y, d\nu) \to L^p(X, d\mu), \]
defined by
\[ (Kf)(x) = \int_Y K(x, y)f(y)d\nu(y), \quad (8.29) \]
for \(\mu\)-almost every \(x\) is bounded with \(\|K\| \leq C_1C_2\).

**Proof.** We assume \(1 < p < \infty\) for simplicity and leave the cases \(p = 1, \infty\) to the reader. Choose \(f \in L^p(Y, d\nu)\). By Fubini’s theorem \(\int_Y |K(x, y)f(y)|d\nu(y)\) is measurable and by Hölder’s inequality we have
\[
\int_Y |K(x, y)f(y)|d\nu(y) \leq \int_Y K_1(x, y)K_2(x, y)|f(y)|d\nu(y)
\leq \left( \int_Y K_1(x, y)^q d\nu(y) \right)^{1/q} \left( \int_Y |K_2(x, y)f(y)|^p d\nu(y) \right)^{1/p}
\leq C_1 \left( \int_Y |K_2(x, y)f(y)|^p d\nu(y) \right)^{1/p}
\]
for \(\mu\) a.e. \(x\) (if \(K_2(x, .)f(.) \notin L^p(X, d\nu)\), the inequality is trivially true). Now take this inequality to the \(p\)’th power and integrate with respect to \(x\) using Fubini
\[
\int_X \left( \int_Y |K(x, y)f(y)|d\nu(y) \right)^p d\mu(x) \leq C_1^p \int_X \int_Y |K_2(x, y)f(y)|^p d\nu(y)d\mu(x)
\leq C_1^p \int_Y \int_X |K_2(x, y)f(y)|^p d\mu(x)d\nu(y) \leq C_1^p C_2^p \|f\|_p^p.
\]
Hence \(\int_Y |K(x, y)f(y)|d\nu(y) \in L^p(X, d\mu)\) and, in particular, it is finite for \(\mu\)-almost every \(x\). Thus \(K(x, .)f(.)\) is \(\nu\) integrable for \(\mu\)-almost every \(x\) and \(\int_Y K(x, y)f(y)d\nu(y)\) is measurable. \(\square\)

Note that the assumptions are, for example, satisfied if \(\|K(x, .)\|_{L^1(Y, d\nu)} \leq C\) and \(\|K(., y)\|_{L^1(X, d\mu)} \leq C\) which follows by choosing \(K_1(x, y) = |K(x, y)|^{1/q}\) and \(K_2(x, y) = |K(x, y)|^{1/p}\). For related results see also Problems 8.15 and 12.3.

Another case of special importance is the case of integral operators
\[ (Kf)(x) = \int_X K(x, y)f(y)d\mu(y), \quad f \in L^2(X, d\mu), \quad (8.30) \]
where \(K(x, y) \in L^2(X \times X, d\mu \otimes d\mu)\). Such an operator is called a Hilbert–Schmidt operator.

**Lemma 8.19.** Let \(K\) be a Hilbert–Schmidt operator in \(L^2(X, d\mu)\). Then
\[
\int_X \int_X |K(x, y)|^2 d\mu(x)d\nu(y) = \sum_{j \in J} \|Ku_j\|^2 \quad (8.31)
\]
for every orthonormal basis \( \{ u_j \}_{j \in J} \) in \( L^2(X, d\mu) \).

**Proof.** Since \( K(x, .) \in L^2(X, d\mu) \) for \( \mu \)-almost every \( x \) we infer from Parseval’s relation
\[
\sum_j \left| \int_X K(x, y)u_j(y)d\mu(y) \right|^2 = \int_X |K(x, y)|^2d\mu(y)
\]
for \( \mu \)-almost every \( x \) and thus
\[
\sum_j \|Ku_j\|^2 = \sum_j \int_X \left( \int_X K(x, y)u_j(y)d\mu(y) \right)^2 d\mu(x)
\]
\[
= \int_X \left( \sum_j \int_X K(x, y)u_j(y)d\mu(y) \right)^2 d\mu(x)
\]
\[
= \int_X \int_X |K(x, y)|^2d\mu(x)d\mu(y)
\]
as claimed. \( \square \)

Hence in combination with Lemma 5.7 this shows that our definition for
integral operators agrees with our previous definition from Section 5.2. In
particular, this gives us an easy to check test for compactness of an integral
operator.

**Example.** Let \([a, b]\) be some compact interval and suppose \( K(x, y) \) is
bounded. Then the corresponding integral operator in \( L^2(a, b) \) is Hilbert–
Schmidt and thus compact. This generalizes Lemma 3.5.

\[\diamond\]

**Problem 8.14.** Suppose \((Y, d\nu) = (X, d\mu) \) with \( X \subseteq \mathbb{R}^n \). Show that
the integral operator with kernel \( K(x, y) = k(x - y) \) is bounded in \( L^p(X, d\mu) \) if
\( k \in L^1(X, d\mu) \) with \( \|K\| \leq \|k\|_1 \).

**Problem 8.15** (Schur test). Let \( K(x, y) \) be given and suppose there are
positive measurable functions \( a(x) \) and \( b(y) \) such that
\[
\|K(\cdot)b(\cdot)\|_{L^1(Y, d\nu)} \leq C_1 a(x), \quad \|a(\cdot)K(\cdot, y)\|_{L^1(X, d\mu)} \leq C_2 b(y).
\]
Then the operator \( K : L^2(Y, d\nu) \to L^2(X, d\mu) \), defined by
\[
(Kf)(x) = \int_Y K(x, y)f(y)d\nu(y),
\]
for \( \mu \)-almost every \( x \) is bounded with \( \|K\| \leq \sqrt{C_1C_2} \). (Hint: Estimate
\( |(Kf)(x)|^2 = \int_Y K(x, y)b(y)f(y)b(y)^{-1}d\nu(y) \) using Cauchy–Schwarz and
integrate the result with respect to \( x \).)
More measure theory

9.1. Decomposition of measures

Let $\mu$, $\nu$ be two measures on a measurable space $(X, \Sigma)$. They are called mutually singular (in symbols $\mu \perp \nu$) if they are supported on disjoint sets. That is, there is a measurable set $N$ such that $\mu(N) = 0$ and $\nu(X \setminus N) = 0$.

Example. Let $\lambda$ be the Lebesgue measure and $\Theta$ the Dirac measure (centered at 0). Then $\lambda \perp \Theta$: Just take $N = \{0\}$; then $\lambda(\{0\}) = 0$ and $\Theta(\mathbb{R} \setminus \{0\}) = 0$.

On the other hand, $\nu$ is called absolutely continuous with respect to $\mu$ (in symbols $\nu \ll \mu$) if $\mu(A) = 0$ implies $\nu(A) = 0$.

Example. The prototypical example is the measure $d\nu = f \, d\mu$ (compare Lemma 7.19). Indeed by Lemma 7.22 $\mu(A) = 0$ implies

$$\nu(A) = \int_A f \, d\mu = 0$$

and shows that $\nu$ is absolutely continuous with respect to $\mu$. In fact, we will show below that every absolutely continuous measure is of this form.

The two main results will follow as simple consequence of the following result:

**Theorem 9.1.** Let $\mu$, $\nu$ be $\sigma$-finite measures. Then there exists a nonnegative function $f$ and a set $N$ of $\mu$ measure zero, such that

$$\nu(A) = \nu(A \cap N) + \int_A f \, d\mu.$$  

(9.2)
Proof. We first assume \( \mu, \nu \) to be finite measures. Let \( \alpha = \mu + \nu \) and consider the Hilbert space \( L^2(X, d\alpha) \). Then
\[
\ell(h) = \int_X h \, d\nu
\]
is a bounded linear functional on \( L^2(X, d\alpha) \) by Cauchy–Schwarz:
\[
|\ell(h)|^2 = \left| \int_X 1 \cdot h \, d\nu \right|^2 \leq \left( \int |1|^2 \, d\nu \right) \left( \int |h|^2 \, d\nu \right) \leq \nu(X) \left( \int |h|^2 \, d\alpha \right) = \nu(X)\|h\|^2.
\]
Hence by the Riesz lemma (Theorem 2.10) there exists a \( g \in L^2(X, d\alpha) \) such that
\[
\ell(h) = \int_X hg \, d\alpha.
\]
By construction
\[
\nu(A) = \int \chi_A \, d\nu = \int \chi_A g \, d\alpha = \int_A g \, d\alpha.
\]
(9.3)
In particular, \( g \) must be positive a.e. (take \( A \) the set where \( g \) is negative).
Moreover,
\[
\mu(A) = \alpha(A) - \nu(A) = \int_A (1 - g) \, d\alpha
\]
which shows that \( g \leq 1 \) a.e. Now chose \( N = \{x|g(x) = 1\} \) such that \( \mu(N) = 0 \) and set
\[
f = \frac{g}{1 - g} \chi_{N'}, \quad N' = X \setminus N.
\]
Then, since (9.3) implies \( d\nu = g \, d\alpha \), respectively, \( d\mu = (1 - g) \, d\alpha \), we have
\[
\int_A f \, d\mu = \int \chi_A \frac{g}{1 - g} \chi_{N'} \, d\mu = \int \chi_A \cap N' g \, d\alpha = \nu(A \cap N')
\]
as desired.
To see the \( \sigma \)-finite case, observe that \( Y_n \nrightarrow X, \mu(Y_n) < \infty \) and \( Z_n \nrightarrow X, \nu(Z_n) < \infty \) implies \( X_n = Y_n \cap Z_n \nrightarrow X \) and \( \alpha(X_n) < \infty \). Now we set \( \bar{X}_n = X_n \setminus X_{n-1} \) (where \( X_0 = \emptyset \)) and consider \( \mu_n(A) = \mu(A \cap \bar{X}_n) \) and \( \nu_n(A) = \nu(A \cap \bar{X}_n) \). Then there exist corresponding sets \( N_n \) and functions \( f_n \) such that
\[
\nu_n(A) = \nu_n(A \cap N_n) + \int_A f_n \, d\mu_n = \nu(A \cap N_n) + \int_A f_n \, d\mu,
\]
where for the last equality we have assumed \( N_n \subset \bar{X}_n \) and \( f_n(x) = 0 \) for \( x \in \bar{X}_n' \) without loss of generality. Now set \( N = \bigcup_n N_n \) as well as \( f = \sum_n f_n \),
9.1. Decomposition of measures

then \( \mu(N) = 0 \) and
\[
\nu(A) = \sum_n \nu_n(A) = \sum_n \nu(A \cap N_n) + \sum_n \int_A f_n \, d\mu = \nu(A \cap N) + \int_A f \, d\mu,
\]
which finishes the proof. \( \square \)

Now the anticipated results follow with no effort:

**Theorem 9.2** (Radon–Nikodym). Let \( \mu, \nu \) be two \( \sigma \)-finite measures on a measurable space \((X, \Sigma)\). Then \( \nu \) is absolutely continuous with respect to \( \mu \) if and only if there is a nonnegative measurable function \( f \) such that
\[
\nu(A) = \int_A f \, d\mu \tag{9.4}
\]
for every \( A \in \Sigma \). The function \( f \) is determined uniquely a.e. with respect to \( \mu \) and is called the **Radon–Nikodym derivative** \( \frac{d\nu}{d\mu} \) of \( \nu \) with respect to \( \mu \).

**Proof.** Just observe that in this case \( \nu(A \cap N) = 0 \) for every \( A \). Uniqueness will be shown in the next theorem. \( \square \)

**Example.** Take \( X = \mathbb{R} \). Let \( \mu \) be the counting measure and \( \nu \) Lebesgue measure. Then \( \nu \ll \mu \) but there is no \( f \) with \( \frac{d\nu}{d\mu} = f \). If there were such an \( f \), there must be a point \( x_0 \in \mathbb{R} \) with \( f(x_0) > 0 \) and we have \( 0 = \nu(\{x_0\}) = \int_{\{x_0\}} f \, d\mu = f(x_0) > 0 \), a contradiction. Hence the Radon–Nikodym theorem can fail if \( \mu \) is not \( \sigma \)-finite. \( \diamond \)

**Theorem 9.3** (Lebesgue decomposition). Let \( \mu, \nu \) be two \( \sigma \)-finite measures on a measurable space \((X, \Sigma)\). Then \( \nu \) can be uniquely decomposed as \( \nu = \nu_{ac} + \nu_{sing} \), where \( \mu \) and \( \nu_{sing} \) are mutually singular and \( \nu_{ac} \) is absolutely continuous with respect to \( \mu \).

**Proof.** Taking \( \nu_{sing}(A) = \nu(A \cap \tilde{N}) \) and \( d\nu_{ac} = f \, d\mu \) from the previous theorem, there is at least one such decomposition. To show uniqueness assume there is another one, \( \nu = \tilde{\nu}_{ac} + \tilde{\nu}_{sing} \), and let \( \tilde{N} \) be such that \( \mu(\tilde{N}) = 0 \) and \( \tilde{\nu}_{sing}(\tilde{N}) = 0 \). Then \( \nu_{sing}(A) - \tilde{\nu}_{sing}(A) = \int_A (\tilde{f} - f) \, d\mu \). In particular, \( \int_{A \cap N(\tilde{N})} (\tilde{f} - f) \, d\mu = 0 \) and hence, since \( A \) is arbitrary, \( \tilde{f} = f \) a.e. away from \( N \cup \tilde{N} \). Since \( \mu(N \cup \tilde{N}) = 0 \), we have \( \tilde{f} = f \) a.e. and hence \( \tilde{\nu}_{ac} = \nu_{ac} \) as well as \( \tilde{\nu}_{sing} = \nu - \tilde{\nu}_{ac} = \nu - \nu_{ac} = \nu_{sing} \). \( \square \)

**Problem 9.1.** Let \( \mu \) be a Borel measure on \( \mathcal{B} \) and suppose its distribution function \( \mu(x) \) is continuously differentiable. Show that the Radon–Nikodym derivative equals the ordinary derivative \( \mu'(x) \).

**Problem 9.2.** Suppose \( \mu \) is a Borel measure on \( \mathbb{R} \) and \( f : \mathbb{R} \to \mathbb{R} \) is continuous. Show that \( f_* \mu \) is absolutely continuous if \( \mu \) is. (Hint: Problem 7.13)
Problem 9.3. Suppose $\mu$ and $\nu$ are inner regular measures. Show that $\nu \ll \mu$ if and only if $\mu(C) = 0$ implies $\nu(C) = 0$ for every compact set.

Problem 9.4. Suppose $\nu(A) \leq C \mu(A)$ for all $A \in \Sigma$. Then $d\nu = f \, d\mu$ with $0 \leq f \leq C$ a.e.

Problem 9.5. Let $d\nu = f \, d\mu$. Suppose $f > 0$ a.e. with respect to $\mu$. Then $\mu \ll \nu$ and $d\mu = f^{-1} \, d\nu$.

Problem 9.6 (Chain rule). Show that $\nu \ll \mu$ is a transitive relation. In particular, if $\omega \ll \nu \ll \mu$, show that

$$\frac{d\omega}{d\mu} = \frac{d\omega}{d\nu} \cdot \frac{d\nu}{d\mu}. \quad (9.5)$$

Problem 9.7. Suppose $\nu \ll \mu$. Show that for every measure $\omega$ we have

$$\frac{d\omega}{d\mu} \, d\mu = \frac{d\omega}{d\nu} \, d\nu + d\zeta,$$

where $\zeta$ is a positive measure (depending on $\omega$) which is singular with respect to $\nu$. Show that $\zeta = 0$ if and only if $\mu \ll \nu$.

9.2. Derivatives of measures

If $\mu$ is a Borel measure on $\mathcal{B}$ and its distribution function $\mu(x)$ is continuously differentiable, then the Radon–Nikodym derivative is just the ordinary derivative $\mu'(x)$ (Problem 9.1). Our aim in this section is to generalize this result to arbitrary Borel measures on $\mathcal{B}^n$.

Let $\mu$ be a Borel measure on $\mathbb{R}^n$. We call

$$(D\mu)(x) = \lim_{\varepsilon \downarrow 0} \frac{\mu(B_{\varepsilon}(x))}{|B_{\varepsilon}(x)|} \quad (9.5)$$

the derivative of $\mu$ at $x \in \mathbb{R}^n$ provided the above limit exists. (Here $B_r(x) \subset \mathbb{R}^n$ is a ball of radius $r$ centered at $x \in \mathbb{R}^n$ and $|A|$ denotes the Lebesgue measure of $A \in \mathcal{B}^n$.)

**Example.** Consider a Borel measure on $\mathcal{B}$ and suppose its distribution $\mu(x)$ (as defined in (7.3)) is differentiable at $x$. Then

$$\frac{d\mu}{d\nu}(x) = \frac{\mu((x + \varepsilon, x - \varepsilon))}{2\varepsilon} = \frac{\mu(x + \varepsilon) - \mu(x - \varepsilon)}{2\varepsilon} = \mu'(x). \quad \diamond$$

To compute the derivative of $\mu$, we introduce the upper and lower derivative,

$$\overline{\frac{d\mu}{d\nu}}(x) = \limsup_{\varepsilon \downarrow 0} \frac{\mu(B_{\varepsilon}(x))}{|B_{\varepsilon}(x)|} \quad \text{and} \quad \underline{\frac{d\mu}{d\nu}}(x) = \liminf_{\varepsilon \downarrow 0} \frac{\mu(B_{\varepsilon}(x))}{|B_{\varepsilon}(x)|}. \quad (9.6)$$
9.2. Derivatives of measures

Clearly $\mu$ is differentiable at $x$ if $(D\mu)(x) = (\overline{D}\mu)(x) < \infty$. Next note that they are measurable: In fact, this follows from

$$(\overline{D}\mu)(x) = \lim_{n \to \infty} \sup_{0<\varepsilon<1/n} \frac{\mu(B_{\varepsilon}(x))}{|B_{\varepsilon}(x)|}$$

since the supremum on the right-hand side is lower semicontinuous with respect to $x$ (cf. Problem 7.10) as $x \mapsto \mu(B_{\varepsilon}(x))$ is lower semicontinuous (Problem 9.8). Similarly for $(D\mu)(x)$.

Next, the following geometric fact of $\mathbb{R}^n$ will be needed.

**Lemma 9.4 (Wiener covering lemma).** Given open balls $B_1 = B_{r_1}(x_1), \ldots, B_m = B_{r_m}(x_m)$ in $\mathbb{R}^n$, there is a subset of disjoint balls $B_{j_1}, \ldots, B_{j_k}$ such that

$$\bigcup_{j=1}^m B_j \subseteq \bigcup_{\ell=1}^k B_{3r_{j_\ell}}(x_{j_\ell})$$

**Proof.** Assume that the balls $B_j$ are ordered by decreasing radius. Start with $B_{j_1} = B_1$ and remove all balls from our list which intersect $B_{j_1}$. Observe that the removed balls are all contained in $B_{3r_1}(x_1)$. Proceeding like this, we obtain the required subset. \(\square\)

The upshot of this lemma is that we can select a disjoint subset of balls which still controls the Lebesgue volume of the original set up to a universal constant $3^n$ (recall $|B_{3r}(x)| = 3^n|B_r(x)|$).

Now we can show

**Lemma 9.5.** Let $\alpha > 0$. For every Borel set $A$ we have

$$\left| \{x \in A \mid (D\mu)(x) > \alpha \} \right| \leq 3^n \frac{\mu(A)}{\alpha}$$

and

$$\left| \{x \in A \mid (\overline{D}\mu)(x) > 0 \} \right| = 0, \text{ whenever } \mu(A) = 0.$$ (9.10)

**Proof.** Let $A_\alpha = \{x \in A \mid (D\mu)(x) > \alpha \}$. We will show

$$|K| \leq 3^n \frac{\mu(O)}{\alpha}$$

for every open set $O$ with $A \subseteq O$ and every compact set $K \subseteq A_\alpha$. The first claim then follows from outer regularity of $\mu$ and inner regularity of the Lebesgue measure.

Given fixed $K$, $O$, for every $x \in K$ there is some $r_x$ such that $B_{r_x}(x) \subseteq O$ and $|B_{r_x}(x)| < \alpha^{-1}\mu(B_{r_x}(x))$. Since $K$ is compact, we can choose a finite
subcover of $K$ from these balls. Moreover, by Lemma 9.4 we can refine our set of balls such that

$$|K| \leq 3^n \sum_{i=1}^{k} |B_{r_i}(x_i)| < \frac{3^n}{\alpha} \sum_{i=1}^{k} \mu(B_{r_i}(x_i)) \leq 3^n \frac{\mu(O)}{\alpha}.$$  

To see the second claim, observe that $A_0 = \bigcup_{j=1}^{\infty} A_{1/j}$ and by the first part $|A_{1/j}| = 0$ for every $j$ if $\mu(A) = 0$. □

**Theorem 9.6 (Lebesgue).** Let $f$ be (locally) integrable, then for a.e. $x \in \mathbb{R}^n$ we have

$$\lim_{r \downarrow 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)| d^n y = 0. \quad (9.11)$$

The points where (9.11) holds are called **Lebesgue points** of $f$.

**Proof.** Decompose $f$ as $f = g + h$, where $g$ is continuous and $\|h\|_1 < \varepsilon$ (Theorem 8.13) and abbreviate

$$D_r(f)(x) = \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)| d^n y.$$  

Then, since $\lim D_r(g)(x) = 0$ (for every $x$) and $D_r(f) \leq D_r(g) + D_r(h)$, we have

$$\limsup_{r \downarrow 0} D_r(f)(x) \leq \limsup_{r \downarrow 0} D_r(h)(x) \leq (\overline{D}\mu)(x) + |h(x)|,$$

where $d\mu = |h|d^n x$. This implies

$$\{x \mid \limsup_{r \downarrow 0} D_r(f)(x) \geq 2\alpha\} \subseteq \{x \mid (\overline{D}\mu)(x) \geq \alpha\} \cup \{x \mid |h(x)| \geq \alpha\}$$

and using the first part of Lemma 9.5 plus $|\{x \mid |h(x)| \geq \alpha\}| \leq \alpha^{-1} \|h\|_1$ (Problem 9.11), we see

$$|\{x \mid \limsup_{r \downarrow 0} D_r(f)(x) \geq 2\alpha\}| \leq (3^n + 1)\frac{\varepsilon}{\alpha}.$$  

Since $\varepsilon$ is arbitrary, the Lebesgue measure of this set must be zero for every $\alpha$. That is, the set where the lim sup is positive has Lebesgue measure zero. □

Note that the balls can be replaced by more general sets: A sequence of sets $A_j(x)$ is said to shrink to $x$ nicely if there are balls $B_{r_j}(x)$ with $r_j \to 0$ and a constant $\varepsilon > 0$ such that $A_j(x) \subseteq B_{r_j}(x)$ and $|A_j| \geq \varepsilon |B_{r_j}(x)|$. For example, $A_j(x)$ could be some balls or cubes (not necessarily containing $x$). However, the portion of $B_{r_j}(x)$ which they occupy must not go to zero! For example, the rectangles $(0, 1/j) \times (0, 2/j) \subset \mathbb{R}^2$ do shrink nicely to 0, but the rectangles $(0, 1/j) \times (0, 2/j^2)$ do not.
Lemma 9.7. Let \( f \) be (locally) integrable. Then at every Lebesgue point we have
\[
f(x) = \lim_{j \to \infty} \frac{1}{|A_j(x)|} \int_{A_j(x)} f(y) d^n y
\]
whenever \( A_j(x) \) shrinks to \( x \) nicely.

**Proof.** Let \( x \) be a Lebesgue point and choose some nicely shrinking sets \( A_j(x) \) with corresponding \( B_{r_j}(x) \) and \( \varepsilon \). Then
\[
\frac{1}{|A_j(x)|} \int_{A_j(x)} |f(y) - f(x)| d^n y \leq \frac{1}{\varepsilon |B_{r_j}(x)|} \int_{B_{r_j}(x)} |f(y) - f(x)| d^n y
\]
and the claim follows. \( \square \)

Corollary 9.8. Let \( \mu \) be a Borel measure on \( \mathbb{R} \) which is absolutely continuous with respect to Lebesgue measure. Then its distribution function is differentiable a.e. and \( d\mu(x) = \mu'(x) dx \).

**Proof.** By assumption \( d\mu(x) = f(x) dx \) for some locally integrable function \( f \). In particular, the distribution function \( \mu(x) = \int_0^x f(y) dy \) is continuous. Moreover, since the sets \( (x, x + r) \) shrink nicely to \( x \) as \( r \to 0 \), Lemma 9.7 implies
\[
\lim_{r \to 0} \frac{\mu((x, x + r))}{r} = \lim_{r \to 0} \frac{\mu(x + r) - \mu(x)}{r} = f(x)
\]
at every Lebesgue point of \( f \). Since the same is true for the sets \( (x - r, x) \), \( \mu(x) \) is differentiable at every Lebesgue point and \( \mu'(x) = f(x) \). \( \square \)

As another consequence we obtain

**Theorem 9.9.** Let \( \mu \) be a Borel measure on \( \mathbb{R}^n \). The derivative \( D\mu \) exists a.e. with respect to Lebesgue measure and equals the Radon–Nikodym derivative of the absolutely continuous part of \( \mu \) with respect to Lebesgue measure; that is,
\[
\mu_{ac}(A) = \int_A (D\mu)(x) d^n x. \tag{9.13}
\]

**Proof.** If \( d\mu = f \, d^n x \) is absolutely continuous with respect to Lebesgue measure, then \( (D\mu)(x) = f(x) \) at every Lebesgue point of \( f \) by Lemma 9.7 and the claim follows from Theorem 9.6. To see the general case, use the Lebesgue decomposition of \( \mu \) and let \( N \) be a support for the singular part with \( |N| = 0 \). Then \( (D\mu_{sing})(x) = 0 \) for a.e. \( x \in \mathbb{R}^n \setminus N \) by the second part of Lemma 9.5. \( \square \)

In particular, \( \mu \) is singular with respect to Lebesgue measure if and only if \( D\mu = 0 \) a.e. with respect to Lebesgue measure.
Using the upper and lower derivatives, we can also give supports for the absolutely and singularly continuous parts.

**Theorem 9.10.** The set \( \{ x \mid 0 < (D\mu)(x) < \infty \} \) is a support for the absolutely continuous and \( \{ x \mid (D\mu)(x) = \infty \} \) is a support for the singular part.

**Proof.** The first part is immediate from the previous theorem. For the second part first note that by \( (D\mu)(x) \geq (D\mu_{\text{sing}})(x) \) we can assume that \( \mu \) is purely singular. It suffices to show that the set \( A_k = \{ x \mid (D\mu)(x) < k \} \) satisfies \( \mu(A_k) = 0 \) for every \( k \in \mathbb{N} \).

Let \( K \subset A_k \) be compact, and let \( V_j \supset K \) be some open set such that \( |V_j \backslash K| \leq \frac{1}{j} \). For every \( x \in K \) there is some \( \varepsilon = \varepsilon(x) \) such that \( B_{\varepsilon}(x) \subseteq V_j \) and \( \mu(B_{3\varepsilon}(x)) \leq k |B_{3\varepsilon}(x)| \). By compactness, finitely many of these balls cover \( K \) and hence

\[
\mu(K) \leq \sum_i \mu(B_{\varepsilon_i}(x_i)).
\]

Selecting disjoint balls as in Lemma 9.4 further shows

\[
\mu(K) \leq \sum_{i} \mu(B_{3\varepsilon_i}(x_i)) \leq k 3^n \sum_{i} |B_{3\varepsilon_i}(x_i)| \leq k 3^n |V_j|.
\]

Letting \( j \to \infty \), we see \( \mu(K) \leq k 3^n |K| \) and by regularity we even have \( \mu(A) \leq k 3^n |A| \) for every \( A \subseteq A_k \). Hence \( \mu \) is absolutely continuous on \( A_k \) and since we assumed \( \mu \) to be singular, we must have \( \mu(A_k) = 0 \).

Finally, we note that these supports are minimal. Here a support \( M \) of some measure \( \mu \) is called a **minimal support** (it is sometimes also called an **essential support**) if every subset \( M_0 \subseteq M \) which does not support \( \mu \) (i.e., \( \mu(M_0) = 0 \)) has Lebesgue measure zero.

**Example.** Let \( X = \mathbb{R} \), \( \Sigma = \mathcal{B} \). If \( d\mu(x) = \sum_n \alpha_n d\theta(x - x_n) \) is a sum of Dirac measures, then the set \( \{ x_n \} \) is clearly a minimal support for \( \mu \). Moreover, it is clearly the smallest support as none of the \( x_n \) can be removed.

If we choose \( \{ x_n \} \) to be the rational numbers, then \( \text{supp}(\mu) = \mathbb{R} \), but \( \mathbb{R} \) is not a minimal support, as we can remove the irrational numbers.

On the other hand, if we consider the Lebesgue measure \( \lambda \), then \( \mathbb{R} \) is a minimal support. However, the same is true if we remove any set of measure zero, for example, the Cantor set. In particular, since we can remove any single point, we see that, just like supports, minimal supports are not unique.

**Lemma 9.11.** The set \( M_{\text{ac}} = \{ x \mid 0 < (D\mu)(x) < \infty \} \) is a minimal support for \( \mu_{\text{ac}} \).
Proof. Suppose $M_0 \subseteq M_{ac}$ and $\mu_{ac}(M_0) = 0$. Set $M_\varepsilon = \{x \in M_0|\varepsilon < (D\mu)(x)\}$ for $\varepsilon > 0$. Then $M_\varepsilon \not\supseteq M_0$ and
$$|M_\varepsilon| = \int_{M_\varepsilon} d^nx \leq \frac{1}{\varepsilon} \int_{M_\varepsilon} (D\mu)(x) dx = \frac{1}{\varepsilon} \mu_{ac}(M_\varepsilon) \leq \frac{1}{\varepsilon} \mu_{ac}(M_0) = 0$$
shows $|M_0| = \lim_{\varepsilon \downarrow 0} |M_\varepsilon| = 0$. □

Note that the set $M = \{x|0 < (D\mu)(x)\}$ is a minimal support of $\mu$.

Example. The Cantor function is constructed as follows: Take the sets $C_n$ used in the construction of the Cantor set: $C_n$ is the union of $2^n$ closed intervals with $2^n - 1$ open gaps in between. Set $f_n$ equal to $j/2^n$ on the $j$'th gap of $C_n$ and extend it to $[0, 1]$ by linear interpolation. Note that, since we are creating precisely one new gap between every old gap when going from $C_n$ to $C_{n+1}$, the value of $f_{n+1}$ is the same as the value of $f_n$ on the gaps of $C_n$. Explicitly, we have $f_0(x) = x$ and $f_{n+1} = K(f_n)$, where
$$K(f)(x) = \begin{cases} \frac{1}{2}f(3x), & 0 \leq x \leq \frac{1}{3}, \\ \frac{1}{2}f(3x), & \frac{1}{3} \leq x \leq \frac{2}{3}, \\ \frac{1}{2}(1 + f(3x - 2)), & \frac{2}{3} \leq x \leq 1. \end{cases}$$
Since $\|f_{n+1} - f_n\|_\infty \leq \frac{1}{2}\|f_{n+1} - f_n\|_\infty$ we can define the Cantor function as $f = \lim_{n \to \infty} f_n$. By construction $f$ is a continuous function which is constant on every subinterval of $[0, 1]\setminus C$. Since $C$ is of Lebesgue measure zero, this set is of full Lebesgue measure and hence $f' = 0$ a.e. in $[0, 1]$. In particular, the corresponding measure, the Cantor measure, is supported on $C$ and is purely singular with respect to Lebesgue measure.

Problem 9.8. Show that
$$\mu(B_\varepsilon(x)) \leq \liminf_{y \to x} \mu(B_\varepsilon(y)) \leq \limsup_{y \to x} \mu(B_\varepsilon(y)) \leq \mu(B_\varepsilon(x)).$$
In particular, conclude that $x \mapsto \mu(B_\varepsilon(x))$ is lower semicontinuous for $\varepsilon > 0$.

Problem 9.9. Show that $M = \{x|0 < (D\mu)(x)\}$ is a minimal support of $\mu$.

Problem 9.10. Suppose $\overline{D\mu} \leq \alpha$. Show that $d\mu = f d^n x$ with $\|f\|_\infty \leq \alpha$.

Problem 9.11 (Chebyshev inequality). For $f \in L^1(\mathbb{R}^n)$ show
$$\left|\{x \in A | f(x) > \alpha\}\right| \leq \frac{1}{\alpha} \int_A |f(x)| d^n x.$$

Problem 9.12. Show that the Cantor function is Hölder continuous $|f(x) - f(y)| \leq |x - y|^\alpha$ with exponent $\alpha = \log_3(2)$. (Hint: Show that if $g$ satisfies a Hölder estimate $|g(x) - g(y)| \leq M|x - y|^\alpha$, then so does $K(g)$: $|K(g)(x) - K(g)(y)| \leq \frac{3\alpha}{2} M|x - y|^\alpha$.)
9.3. Complex measures

Let \((X, \Sigma)\) be some measurable space. A map \(\nu : \Sigma \to \mathbb{C}\) is called a complex measure if
\[
\nu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \nu(A_n), \quad A_n \cap A_m = \emptyset, \quad n \neq m. \tag{9.14}
\]
Choosing \(A_n = \emptyset\) for all \(n\) in (9.14) shows \(\nu(\emptyset) = 0\).

Note that a positive measure is a complex measure only if it is finite (the value \(\infty\) is not allowed for complex measures). Moreover, the definition implies that the sum is independent of the order of the sets \(A_j\), that is, it converges unconditionally and thus absolutely by the Riemann series theorem.

**Example.** Let \(\mu\) be a positive measure. For every \(f \in L^1(X, d\mu)\) we have that \(fd\mu\) is a complex measure (compare the proof of Lemma 7.19 and use dominated in place of monotone convergence). In fact, we will show that every complex measure is of this form. \hfill \Box

**Example.** Let \(\nu_1, \nu_2\) be two complex measures and \(\alpha_1, \alpha_2\) two complex numbers. Then \(\alpha_1 \nu_1 + \alpha_2 \nu_2\) is again a complex measure. Clearly we can extend this to any finite linear combination of complex measures. \hfill \Box

Given a complex measure \(\nu\) it seems natural to consider the set function \(A \mapsto |\nu(A)|\). However, considering the simple example \(d\nu(x) = \text{sign}(x) dx\) on \(X = [-1, 1]\) one sees that this set function is not additive and this simple approach does not provide a positive measure associated with \(\nu\). However, using \(|\nu(A \cap [-1, 0])| + |\nu(A \cap [0, 1])|\) we do get a positive measure. Motivated by this we introduce the total variation of a measure defined as
\[
|\nu|(A) = \sup \left\{ \sum_{k=1}^{n} |\nu(A_k)| \middle| A_k \in \Sigma \text{ disjoint}, A = \bigcup_{k=1}^{n} A_k \right\}. \tag{9.15}
\]
Note that by construction we have
\[
|\nu(A)| \leq |\nu|(A). \tag{9.16}
\]
Moreover, the total variation is monotone \(|\nu|(A) \leq |\nu|(B)\) if \(A \subseteq B\) and for a positive measure \(\mu\) we have of course \(|\mu|(A) = \mu(A)\).

**Theorem 9.12.** The total variation \(|\nu|\) of a complex measure \(\nu\) is a finite positive measure.

**Proof.** We begin by showing that \(|\nu|\) is a positive measure. We need to show \(|\nu|(A) = \sum_{n=1}^{\infty} |\nu|(A_n)\) for any partition of \(A\) into disjoint sets \(A_n\). If \(|\nu|(A_n) = \infty\) for some \(n\) it is not hard to see that \(|\nu|(A) = \infty\) and hence we can assume \(|\nu|(A_n) < \infty\) for all \(n\).
Let \(\varepsilon > 0\) be fixed and for each \(A_n\) choose a disjoint partition \(B_{n,k}\) of \(A_n\) such that
\[
|\nu|(A_n) \leq \sum_{k=1}^{m} |\nu(B_{n,k})| + \frac{\varepsilon}{2^n}.
\]

Then
\[
\sum_{n=1}^{N} |\nu|(A_n) \leq \sum_{n=1}^{N} \sum_{k=1}^{m} |\nu(B_{n,k})| + \varepsilon \leq |\nu|\left(\bigcup_{n=1}^{N} A_n\right) + \varepsilon \leq |\nu|(A) + \varepsilon
\]
since \(\bigcup_{n=1}^{N} \bigcup_{k=1}^{m} B_{n,k} = \bigcup_{n=1}^{N} A_n\). Since \(\varepsilon\) was arbitrary this shows \(|\nu|(A) \geq \sum_{n=1}^{\infty} |\nu|(A_n)\).

Conversely, given a finite partition \(B_k\) of \(A\), then
\[
\sum_{k=1}^{m} |\nu(B_k)| = \sum_{k=1}^{m} \left| \sum_{n=1}^{\infty} \nu(B_k \cap A_n) \right| \leq \sum_{k=1}^{m} \sum_{n=1}^{\infty} |\nu(B_k \cap A_n)| \leq \sum_{n=1}^{\infty} \sum_{k=1}^{m} |\nu(B_k \cap A_n)| \leq \sum_{n=1}^{\infty} |\nu|(A_n).
\]

Taking the supremum over all partitions \(B_k\) shows \(|\nu|(A) \leq \sum_{n=1}^{\infty} |\nu|(A_n)\).

Hence \(|\nu|\) is a positive measure and it remains to show that it is finite. Splitting \(\nu\) into its real and imaginary part, it is no restriction to assume that \(\nu\) is real-valued since \(|\nu|(A) \leq |\text{Re}(\nu)|(A) + |\text{Im}(\nu)|(A)|\).

The idea is as follows: Suppose we can split any given set \(A\) with \(|\nu|(A) = \infty\) into two subsets \(B\) and \(A \setminus B\) such that \(|\nu(B)| \geq 1\) and \(|\nu|(A \setminus B) = \infty\). Then we can construct a sequence \(B_n\) of disjoint sets with \(|\nu(B_n)| \geq 1\) for which
\[
\sum_{n=1}^{\infty} \nu(B_n)
\]
diverges (the terms of a convergent series must converge to zero). But \(\sigma\)-additivity requires that the sum converges to \(\nu(\bigcup_n B_n)\), a contradiction.

It remains to show existence of this splitting. Let \(A\) with \(|\nu|(A) = \infty\) be given. Then there are disjoint sets \(A_j\) such that
\[
\sum_{j=1}^{n} |\nu(A_j)| \geq 2 + |\nu(A)|.
\]

Now let \(A_+ = \bigcup\{A_j | \nu(A_j) \geq 0\}\) and \(A_- = A \setminus A_+ = \bigcup\{A_j | \nu(A_j) < 0\}\). Then the above inequality reads \(\nu(A_+) + |\nu(A_-)| \geq 2 + |\nu(A_+) - |\nu(A_-)||\) implying (show this) that for both of them we have \(|\nu(A_+)\) \(\geq 1\) and by \(|\nu|(A) = |\nu|(A_+) + |\nu|(A_-)\) either \(A_+\) or \(A_-\) must have infinite \(|\nu|\) measure.
Note that this implies that every complex measure $\nu$ can be written as a linear combination of four positive measures. In fact, first we can split $\nu$ into its real and imaginary part
\[ \nu = \nu_r + i\nu_i, \quad \nu_r(A) = \text{Re}(\nu(A)), \quad \nu_i(A) = \text{Im}(\nu(A)). \tag{9.17} \]
Second we can split every real (also called signed) measure according to
\[ \nu = \nu_+ - \nu_- \quad \nu_{\pm}(A) = \frac{|\nu|(A) \pm \nu(A)}{2}. \tag{9.18} \]
By (9.16) both $\nu_-$ and $\nu_+$ are positive measures. This splitting is also known as Jordan decomposition of a signed measure.

Of course such a decomposition of a signed measure is not unique (we can always add a positive measure to both parts), however, the Jordan decomposition is unique in the sense that it is the smallest possible decomposition.

**Lemma 9.13.** Let $\nu$ be a complex measure and $\mu$ a positive measure satisfying $|\nu(A)| \leq \mu(A)$ for all measurable sets $A$. Then $\mu \geq |\nu|$. (Here $|\nu| \leq \mu$ has to be understood as $|\nu|(A) \leq \mu(A)$ for every measurable set $A$.)

Furthermore, let $\nu$ be a signed measure and $\nu = \tilde{\nu}_+ - \tilde{\nu}_-$ a decomposition into positive measures. Then $\tilde{\nu}_\pm \geq \nu_\pm$, where $\nu_\pm$ is the Jordan decomposition.

**Proof.** It suffices to prove the first part since the second is a special case. But for every measurable set $A$ and a corresponding finite partition $A_k$ we have $\sum_k |\nu(A_k)| \leq \sum \mu(A_k) = \mu(A)$ implying $|\nu|(A) \leq \mu(A)$.

Moreover, we also have:

**Theorem 9.14.** The set of all complex measures $M(X)$ together with the norm $\|\nu\| = |\nu|(X)$ is a Banach space.

**Proof.** Clearly $M(X)$ is a vector space and it is straightforward to check that $|\nu|(X)$ is a norm. Hence it remains to show that every Cauchy sequence $\nu_k$ has a limit.

First of all, by $|\nu_k(A) - \nu_j(A)| = |(\nu_k - \nu_j)(A)| \leq |\nu_k - \nu_j|(A) \leq \|\nu_k - \nu_j\|$, we see that $\nu_k(A)$ is a Cauchy sequence in $\mathbb{C}$ for every $A \in \Sigma$ and we can define
\[ \nu(A) = \lim_{k \to \infty} \nu_k(A). \]
Moreover, $C_j = \sup_{k \geq j} \|\nu_k - \nu_j\| \to 0$ as $j \to \infty$ and we have
\[ |\nu_j(A) - \nu(A)| \leq C_j. \]
Next we show that $\nu$ satisfies (9.14). Let $A_m$ be given disjoint sets and set $\tilde{A}_n = \bigcup_{m=1}^n A_m$, $A = \bigcup_{m=1}^\infty A_m$. Since we can interchange limits with
finite sums, (9.14) holds for finitely many sets. Hence it remains to show 
\( \nu(\tilde{A}_n) \to \nu(A) \). This follows from
\[
|\nu(\tilde{A}_n) - \nu(A)| \leq |\nu(\tilde{A}_n) - \nu_k(\tilde{A}_n)| + |\nu_k(\tilde{A}_n) - \nu_k(A)| + |\nu_k(A) - \nu(A)|
\]
\[
\leq 2C_k + |\nu_k(\tilde{A}_n) - \nu_k(A)|.
\]
Finally, \( \nu_k \to \nu \) since \( |\nu_k(A) - \nu(A)| \leq C_k \) implies \( \|\nu_k - \nu\| \leq 4C_k \) (Problem 9.17).

□

If \( \mu \) is a positive and \( \nu \) a complex measure we say that \( \nu \) is absolutely continuous with respect to \( \mu \) if \( \mu(A) = 0 \) implies \( \nu(A) = 0 \).

Lemma 9.15. If \( \mu \) is a positive and \( \nu \) a complex measure then \( \nu \ll \mu \) if and only if \( |\nu| \ll \mu \).

Proof. If \( \nu \ll \mu \), then \( \mu(A) = 0 \) implies \( \mu(B) = 0 \) for every \( B \subseteq A \) and hence \( |\nu|(A) = 0 \). Conversely, if \( |\nu| \ll \mu \), then \( \mu(A) = 0 \) implies \( |\nu(A)| \leq |\nu|(A) = 0 \). □

Now we can prove the complex version of the Radon–Nikodym theorem:

Theorem 9.16 (Complex Radon–Nikodym). Let \((X, \Sigma)\) be a measurable space, \( \mu \) a positive \( \sigma \)-finite measure and \( \nu \) a complex measure which is absolutely continuous with respect to \( \mu \). Then there is a unique \( f \in L^1(X, d\mu) \)

\[
\nu(A) = \int_A f \, d\mu. \tag{9.19}
\]

Proof. By treating the real and imaginary part separately it is no restriction to assume that \( \nu \) is real-valued. Let \( \nu = \nu_+ - \nu_- \) be its Jordan decomposition. Then both \( \nu_+ \) and \( \nu_- \) are absolutely continuous with respect to \( \mu \) and by the Radon–Nikodym theorem there are nonnegative functions \( f_\pm \) such that \( d\nu_\pm = f_\pm d\mu \). By construction
\[
\int_X f_\pm d\mu = \nu_\pm(X) \leq |\nu|(X) < \infty,
\]
which shows \( f = f_+ - f_- \in L^1(X, d\mu) \). Moreover, \( d\nu = d\nu_+ - d\nu_- = f \, d\mu \) as required. □

In this case the total variation of \( d\nu = f \, d\mu \) is just \( d|\nu| = |f|d\mu \):

Lemma 9.17. Suppose \( d\nu = f \, d\mu \), where \( \mu \) is a positive measure and \( f \in L^1(X, d\mu) \). Then
\[
|\nu|(A) = \int_A |f|d\mu. \tag{9.20}
\]
Proof. If $A_n$ are disjoint sets and $A = \bigcup_n A_n$ we have
\[
\sum_n |\nu(A_n)| = \sum_n \left| \int_{A_n} f \, d\mu \right| \leq \sum_n \int_{A_n} |f| \, d\mu = \int_A |f| \, d\mu.
\]
Hence $|\nu|(A) \leq \int_A |f| \, d\mu$. To show the converse define
\[
A^+_k = \{ x \in A | \frac{k-1}{n} < \frac{\text{arg}(f(x)) + \pi}{2\pi} \leq \frac{k}{n}, \quad 1 \leq k \leq n \}.
\]
Then the simple functions
\[
s_n(x) = \sum_{k=1}^n e^{-2\pi i \frac{k-1}{n} \chi_{A^+_k}(x)}
\]
converge to $\text{sign}(f(x^*))$ for every $x \in A$ and hence
\[
\lim_{n \to \infty} \int_A s_n f \, d\mu = \int_A |f| \, d\mu
\]
by dominated convergence. Moreover,
\[
\left| \int_A s_n f \, d\mu \right| \leq \sum_{k=1}^n \left| \int_{A^+_k} s_n f \, d\mu \right| = \sum_{k=1}^n |\nu(A^+_k)| \leq |\nu|(A)
\]
shows $\int_A |f| \, d\mu \leq |\nu|(A)$. \qed

As a consequence we obtain (Problem 9.13):

**Corollary 9.18.** If $\nu$ is a complex measure, then $d\nu = h \, d|\nu|$, where $|h| = 1$.

If $\nu$ is a signed measure, then $h$ is real-valued and we obtain:

**Corollary 9.19.** If $\nu$ is a signed measure, then $d\nu = h \, d|\nu|$, where $h^2 = 1$. In particular, $d\nu_\pm = \chi_{A_\pm} \, d|\nu|$, where $A_\pm = h^{-1}(\{\pm 1\})$.

The decomposition $X = A_+ \cup A_-$ from the previous corollary is known as **Hahn decomposition** and it is characterized by the property that $\pm \nu(A) \geq 0$ if $A \subseteq A_\pm$. This decomposition is not unique since we can shift sets of $|\nu|$ measure zero from one to the other.

Clearly we can use Corollary 9.18 to define the integral of a bounded function $f$ with respect to a complex measure $d\nu = h \, d|\nu|$ as
\[
\int f \, d\nu = \int f h \, d|\nu|.
\]
In fact, it suffices to assume that $f$ is integrable with respect to $d|\nu|$ and we obtain
\[
\left| \int f \, d\nu \right| \leq \int |f| \, d|\nu|.
\]
For bounded functions this implies
\[ \left| \int_A f \, d\nu \right| \leq \|f\|_\infty |\nu|(A). \] (9.23)

Finally, there is an interesting equivalent definition of absolute continuity:

**Lemma 9.20.** If \( \mu \) is a positive and \( \nu \) a complex measure then \( \nu \ll \mu \) if and only if for every \( \varepsilon > 0 \) there is a corresponding \( \delta > 0 \) such that
\[ \mu(A) < \delta \implies |\nu(A)| < \varepsilon, \quad \forall A \in \Sigma. \] (9.24)

**Proof.** Suppose \( \nu \ll \mu \) implying that it is of the form (9.19). Let \( X_n = \{x \in X | |f(x)| \leq n\} \) and note that \( |\nu|(X \setminus X_n) \to 0 \) since \( X_n \nearrow X \) and \( |\nu|(X) < \infty \). Given \( \varepsilon > 0 \) we can choose \( n \) such that \( |\nu|(X \setminus X_n) \leq \frac{\varepsilon}{2} \) and \( \delta = \frac{\varepsilon}{2n} \). Then, if \( \mu(A) < \delta \) we have
\[ |\nu(A)| \leq |\nu|(A \cap X_n) + |\nu|(X \setminus X_n) \leq n \mu(A) + \frac{\varepsilon}{2} < \varepsilon. \]
The converse direction is obvious. \( \square \)

It is important to emphasize that the fact that \( |\nu|(X) < \infty \) is crucial for the above lemma to hold. In fact, it can fail for positive measures as the simple counterexample \( d\nu(\lambda) = \lambda^2 d\lambda \) on \( \mathbb{R} \) shows.

**Problem 9.13.** Prove Corollary 9.18. (Hint: Use the complex Radon–Nikodym theorem to get existence of \( h \). Then show that \( 1 - |h| \) vanishes a.e.)

**Problem 9.14** (Chebyshev inequality). Let \( \nu \) be a complex and \( \mu \) a positive measure. If \( f \) denotes the Radon–Nikodym derivative of \( \nu \) with respect to \( \mu \), then show that
\[ \mu(\{x \in A | |f(x)| \geq \alpha\}) \leq \frac{|\nu|(A)}{\alpha}. \]

**Problem 9.15.** Let \( \nu \) be a complex and \( \mu \) a positive measure and suppose \( |\nu(A)| \leq C \mu(A) \) for all \( A \in \Sigma \). Then \( d\nu = f \, d\mu \) with \( \|f\|_\infty \leq C \). (Hint: First show \( |\nu|(A) \leq C \mu(A) \) and then use Problem 9.4.)

**Problem 9.16.** Let \( \nu \) be a signed measure and \( \nu_\pm \) its Jordan decomposition. Show
\[ \nu_+(A) = \max_{B \in \Sigma, B \subseteq A} \nu(B), \quad \nu_-(A) = -\min_{B \in \Sigma, B \subseteq A} \nu(B). \]

**Problem 9.17.** Let \( \nu \) be a complex measure and let
\[ \nu = \nu_{r,+} - \nu_{r,-} + i(\nu_{i,+} - \nu_{i,-}) \]
be its decomposition into positive measures. Show the estimate
\[ \frac{1}{\sqrt{2}} \nu_s(A) \leq |\nu|(A) \leq \nu_s(A), \quad \nu_s = \nu_{r,+} + \nu_{r,-} + \nu_{i,+} + \nu_{i,-}. \]
Conclude that \(|\nu(A)| \leq C\) for all measurable sets \(A\) implies \(\|\nu\| \leq 4C\).

**Problem 9.18.** Define the convolution of two complex Borel measures \(\mu\) and \(\nu\) on \(\mathbb{R}^n\) via

\[ (\mu \ast \nu)(A) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \chi_A(x+y)\,d\mu(x)\,d\nu(y). \]

Note \(|\mu \ast \nu|(\mathbb{R}^n) \leq |\mu|(\mathbb{R}^n)|\nu|(\mathbb{R}^n)\). Show that this implies

\[ \int_{\mathbb{R}^n} h(x)\,d(\mu \ast \nu)(x) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} h(x+y)\,d\mu(x)\,d\nu(y) \]

for any bounded measurable function \(h\). Conclude that it coincides with our previous definition in case \(\mu\) and \(\nu\) are absolutely continuous with respect to Lebesgue measure.

### 9.4. Appendix: Functions of bounded variation and absolutely continuous functions

Let \([a, b] \subseteq \mathbb{R}\) be some compact interval and \(f : [a, b] \to \mathbb{C}\). Given a partition \(P = \{a = x_0, \ldots, x_n = b\}\) of \([a, b]\) we define the variation of \(f\) with respect to the partition \(P\) by

\[ V(P, f) = \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|. \] (9.25)

The supremum over all partitions

\[ V^b_a(f) = \sup_{\text{partitions } P \text{ of } [a, b]} V(P, f) \] (9.26)

is called the total variation of \(f\) over \([a, b]\). If the total variation is finite, \(f\) is called of bounded variation. Since we clearly have

\[ V^b_a(\alpha f) = |\alpha|V^b_a(f), \quad V^b_a(f+g) \leq V^b_a(f) + V^b_a(g) \] (9.27)

the space \(BV[a, b]\) of all functions of finite total variation is a vector space. However, the total variation is not a norm since (consider the partition \(P = \{a, x, b\}\))

\[ V^b_a(f) = 0 \iff f(x) = c. \] (9.28)

Moreover, any function of bounded variation is in particular bounded (consider again the partition \(P = \{a, x, b\}\))

\[ \sup_{x \in [a,b]} |f(x)| \leq |f(a)| + V^b_a(f). \] (9.29)

Furthermore, observe \(V^a_a(f) = 0\) as well as (Problem 9.20)

\[ V^b_a(f) = V^c_a(f) + V^b_c(f), \quad c \in [a, b], \] (9.30)

and it will be convenient to set

\[ V^b_a(f) = -V^a_a(f). \] (9.31)
Example. Every Lipschitz continuous function is of bounded variation. In fact, if \(|f(x) - f(y)| \leq L|x - y|\) for \(x, y \in [a, b]\), then \(V^b_a(f) \leq L(b - a)\). However, (Hölder) continuity is not sufficient (cf. Problems 9.21 and 9.21).

Example. Any real-valued nondecreasing function \(f\) is of bounded variation with variation given by \(V^b_a(f) = f(b) - f(a)\). Similarly, every real-valued nonincreasing function \(g\) is of bounded variation with variation given by \(V^a_b(g) = g(a) - g(b)\). In particular, the sum \(f + g\) is of bounded variation with variation given by \(V^b_a(f + g) \leq V^b_a(f) + V^b_a(g)\). The following theorem shows that the converse is also true.

Theorem 9.21 (Jordan). Let \(f : [a, b] \to \mathbb{R}\) be of bounded variation, then \(f\) can be decomposed as

\[
f(x) = f_+(x) - f_-(x), \quad f_\pm(x) = \frac{1}{2} (V^x_a(f) \pm f(x)),
\]

where \(f_\pm\) are nondecreasing functions. Moreover, \(V^b_a(f_\pm) \leq V^b_a(f)\).

Proof. From

\[
f(y) - f(x) \leq |f(y) - f(x)| \leq V^y_x(f) = V^y_a(f) - V^x_a(f)
\]

for \(x \leq y\) we infer \(V^x_a(f) - f(x) \leq V^y_a(f) - f(y)\), that is, \(f_+\) is nondecreasing. Moreover, replacing \(f\) by \(-f\) shows that \(f_-\) is nondecreasing and the claim follows.

In particular, we see that functions of bounded variation have at most countably many discontinuities and at every discontinuity the limits from the left and right exist.

For functions \(f : (a, b) \to \mathbb{C}\) (including the case where \((a, b)\) is unbounded) we will set

\[
V^b_a(f) = \lim_{c \downarrow a, d \uparrow b} V^d_c(f).
\]

In this respect the following lemma is of interest (whose proof is left as an exercise):

Lemma 9.22. Suppose \(f \in BV[a, b]\). We have \(\lim_{c \downarrow b} V^c_a(f) = V^b_a(f)\) if and only if \(f(b) = f(b-)\) and \(\lim_{c \uparrow a} V^b_c(f) = V^b_a(f)\) if and only if \(f(a) = f(a+)\). In particular, \(V^x_a(f)\) is left, right continuous if and only \(f\) is.

If \(f : \mathbb{R} \to \mathbb{C}\) is of bounded variation, then we can write it as a linear combination of four nondecreasing functions and hence associate a complex measure \(df\) with \(f\) via Theorem 7.3 (since all four functions are bounded, so are the associated measures).
Theorem 9.23. There is a one-to-one correspondence between functions in $f \in BV(\mathbb{R})$ which are right continuous and normalized by $f(0) = 0$ and complex Borel measures $\nu$ on $\mathbb{R}$ such that $f$ is the distribution function of $\nu$ as defined in (7.3). Moreover, in this case the distribution function of the total variation of $\nu$ is $|\nu|(x) = V^x_0(f)$.

**Proof.** We have already seen how to associate a complex measure $df$ with a function of bounded variation. If $f$ is right continuous and normalized, it will be equal to the distribution function of $df$ by construction. Conversely, let $d\nu$ be a complex measure with distribution function $\nu$. Then for every $a < b$ we have

$$V^b_a(\nu) = \sup_{P = \{a = x_0, \ldots, x_n = b\}} V(P, \nu) = \sup_{P = \{a = x_0, \ldots, x_n = b\}} \sum_{k=1}^n |\nu((x_{k-1}, x_k])| \leq |\nu|((a, b])$$

and thus the distribution function is of bounded variation. Furthermore, consider the measure $\mu$ whose distribution function is $\mu(x) = V^x_0(f)$. Then we see $|\nu((a, b])| = |\nu(b) - \nu(a)| \leq V^b_a(\nu) = \mu((a, b]) \leq |\nu|((a, b])$. Hence we obtain $|\nu(A)| \leq \mu(A) \leq |\nu|(A)$ for all intervals $A$, thus for all open sets (by Problem 1.7), and thus for all Borel sets by outer regularity. Hence Lemma 9.13 implies $\mu = |\nu|$ and hence $|\nu|(x) = V^x_0(f)$. □

We will call a function $f : [a, b] \to \mathbb{C}$ absolutely continuous if for every $\varepsilon > 0$ there is a corresponding $\delta > 0$ such that

$$\sum_k |y_k - x_k| < \delta \Rightarrow \sum_k |f(y_k) - f(x_k)| < \varepsilon \quad (9.34)$$

for every countable collection of pairwise disjoint intervals $(x_k, y_k) \subset [a, b]$. The set of all absolutely continuous functions on $[a, b]$ will be denoted by $AC[a, b]$. The special choice of just one interval shows that every absolutely continuous function is (uniformly) continuous, $AC[a, b] \subset C[a, b]$.

**Example.** Every Lipschitz continuous function is absolutely continuous. In fact, if $|f(x) - f(y)| \leq L|x - y|$ for $x, y \in [a, b]$, then we can choose $\delta = \frac{\varepsilon}{L}$. In particular, $C^1[a, b] \subset AC[a, b]$. Note that Hölder continuity is not sufficient (cf. Problem 9.22 and Theorem 9.25 below). □

**Theorem 9.24.** A complex Borel measure $\nu$ on $\mathbb{R}$ is absolutely continuous with respect to Lebesgue measure if and only if its distribution function is locally absolutely continuous (i.e., absolutely continuous on every compact sub-interval). Moreover, in this case the distribution function $\nu(x)$ is
differentiable almost everywhere and
\[ \nu(x) = \nu(0) + \int_0^x \nu'(y)dy \tag{9.35} \]
with \( \nu' \) integrable, \( \int_{\mathbb{R}} |\nu'(y)|dy = |\nu|_{(\mathbb{R})} \).

**Proof.** Suppose the measure \( \nu \) is absolutely continuous. Since we can write \( \nu \) as a sum of four positive measures, we can suppose \( \nu \) is positive. Now (9.34) follows from (9.24) in the special case where \( A \) is a union of pairwise disjoint intervals.

Conversely, suppose \( \nu(x) \) is absolutely continuous on \([a, b]\). We will verify (9.24). To this end fix \( \varepsilon \) and choose \( \delta \) such that \( \nu(x) \) satisfies (9.34). By outer regularity it suffices to consider the case where \( A \) is open. Moreover, by Problem 1.7, every open set \( O \subset (a, b) \) can be written as a countable union of disjoint intervals \( I_k = (x_k, y_k) \) and thus \( |O| = \sum_k |y_k - x_k| \leq \delta \) implies
\[ \nu(O) = \sum_k (\nu(y_k) - \nu(x_k)) \leq \sum_k |\nu(y_k) - \nu(x_k)| \leq \varepsilon \]
as required.

The rest follows from Corollary 9.8. \( \square \)

As a simple consequence of this result we can give an equivalent definition of absolutely continuous functions as precisely the functions for which the **fundamental theorem of calculus** holds.

**Theorem 9.25.** A function \( f : [a, b] \to \mathbb{C} \) is absolutely continuous if and only if it is of the form
\[ f(x) = f(a) + \int_a^x g(y)dy \tag{9.36} \]
for some integrable function \( g \). Moreover, in this case \( f \) is differentiable \( a.e \) with respect to Lebesgue measure and \( f'(x) = g(x) \). In addition, \( f \) is of bounded variation and
\[ V_a^x(f) = \int_a^x |g(y)|dy. \tag{9.37} \]

**Proof.** This is just a reformulation of the previous result. To see the last claim combine the last part of Theorem 9.23 with Lemma 9.17. \( \square \)

In particular, since the fundamental theorem of calculus fails for the Cantor function, this function is an example of a continuous which is not absolutely continuous. Note that even if \( f \) is differentiable everywhere it might fail the fundamental theorem of calculus (Problem 9.29).
Finally, we note that in this case the integration by parts formula continues to hold.

**Lemma 9.26.** Let \( f, g \in BV[a, b] \), then

\[
\int_{[a,b]} f(x-)dg(x) = f(b-)g(b-) - f(a-)g(a-) - \int_{[a,b]} g(x+)df(x) \quad (9.38)
\]

as well as

\[
\int_{[a,b]} f(x+)dg(x) = f(b+)g(b+) - f(a+)g(a+) - \int_{[a,b]} g(x-)df(x). \quad (9.39)
\]

**Proof.** Since the formula is linear in \( f \) and holds if \( f \) is constant, we can assume \( f(a-) = 0 \) without loss of generality. Similarly, we can assume \( g(b-) = 0 \). Plugging \( f(x-) = \int_{[a,x)} df(y) \) into the left-hand side of the first formula we obtain from Fubini

\[
\int_{[a,b)} f(x-)dg(x) = \int_{[a,b)} \int_{[a,x)} df(y)dg(x) = \int_{[a,b)} \int_{[a,b)} \chi_{\{x,y \mid y < x\}}(x,y)df(y)dg(x)
\]

\[
= \int_{[a,b]} \int_{[a,b)} \chi_{\{x,y \mid y < x\}}(x,y)dg(x)df(y)
\]

\[
= \int_{[a,b]} \int_{(y,b]} dg(x)df(y) = - \int_{[a,b]} g(y+)df(y).
\]

The second formula is shown analogously. \( \square \)

If both \( f, g \in AC[a, b] \) this takes the usual form

\[
\int_a^b f(x)g'(x)dx = f(b)g(b) - f(a)g(a) - \int_a^b g(x)f'(x)dx. \quad (9.40)
\]

**Problem 9.19.** Compute \( V_b^a(f) \) for \( f(x) = \text{sign}(x) \) on \([a, b] = [-1, 1]\).

**Problem 9.20.** Show (9.30).

**Problem 9.21.** Consider \( f_j(x) = x^j \cos(\pi/x) \) for \( j \in \mathbb{N} \). Show that \( f_j \in C[0, 1] \) if we set \( f_j(0) = 0 \). Show that \( f_j \) is of bounded variation for \( j \geq 2 \) but not for \( j = 1 \).

**Problem 9.22.** Let \( \alpha \in (0, 1) \) and \( \beta > 1 \) with \( \alpha \beta < 1 \). Set \( M = \sum_{k=1}^{\infty} k^{-\beta} \) and \( x_n = M^{-1} \sum_{k=1}^{n} k^{-\beta} \). Then we can define a function on \([0, 1]\) via

\[
f(x) = \left| x - \frac{x_{n-1} + x_n}{2} \right|^\alpha, \quad x \in [x_{n-1}, x_n]
\]

and \( f(1) = 0 \). Show that \( f \) is Hölder continuous of exponent \( \alpha \) but not of bounded variation. (Hint: What is the variation on each subinterval?) To
show Hölder continuity consider the cases when both points are in the same interval and in different intervals. For the latter case note that one point can be replaced by one in the same interval as the other without changing the value of $f$.)

**Problem 9.23.** Show that if $f \in BV[a,b]$ then so is $f^*$, $|f|$ and

$$V_a^b(f^*) = V_a^b(f), \quad V_a^b(|f|) \leq V_a^b(f).$$

Moreover, show

$$V_a^b(\Re(f)) \leq V_a^b(f), \quad V_a^b(\Im(f)) \leq V_a^b(f).$$

**Problem 9.24.** Show that if $f, g \in BV[a,b]$ then so is $fg$ and

$$V_a^b(fg) \leq V_a^b(f) \sup |g| + V_a^b(g) \sup |f|.$$

**Problem 9.25.** Show that $BV[a,b]$ together with the norm

$$|f(a)| + V_a^b(f)$$

is a Banach space.

**Problem 9.26** (Product rule for absolutely continuous functions). Show that if $f, g \in AC[a,b]$ then so is $fg$ and $(fg)' = f'g + fg'$. (Hint: Integration by parts.)

**Problem 9.27.** Let $X \subseteq \mathbb{R}$ be an interval, $Y$ some measure space, and $f : X \times Y \to \mathbb{C}$ some measurable function. Suppose $x \mapsto f(x,y)$ is absolutely continuous for a.e. $y$ such that

$$\int_a^b \int_A \left| \frac{\partial}{\partial x} f(x,y) \right| d\mu(y) dx < \infty \quad (9.41)$$

for every compact interval $[a,b] \subseteq X$ and $\int_A |f(c,y)| d\mu(y) < \infty$ for one $c \in X$.

Show that

$$F(x) = \int_A f(x,y) d\mu(y) \quad (9.42)$$

is absolutely continuous and

$$F'(x) = \int_A \frac{\partial}{\partial x} f(x,y) d\mu(y) \quad (9.43)$$

in this case. (Hint: Fubini.)

**Problem 9.28.** Show that if $f \in AC(a,b)$ and $f' \in L^p(a,b)$, then $f$ is Hölder continuous:

$$|f(x) - f(y)| \leq \|f'\|_p |x - y|^{1 - \frac{1}{p}}.$$

Show that the function $f(x) = \log(x)^{-1}$ is absolutely continuous but not Hölder continuous on $[0, \frac{1}{2}]$. 
Problem 9.29. Consider \( f(x) = x^2 \sin(\frac{\pi}{x^2}) \) on \([0, 1]\) (here \( f(0) = 0 \)). Show that \( f \) is differentiable everywhere and compute its derivative. Show that its derivative is not integrable. In particular, this function is not absolutely continuous and the fundamental theorem of calculus does not hold for this function.

Problem 9.30. Show that the function from the previous problem is Hölder continuous of exponent \( \frac{1}{2} \). (Hint: Consider \( 0 < x < y \). There is an \( x' < y \) with \( f(x') = f(x) \) and \((x')^{-2} - y^{-2} \leq 2\pi\). Hence \((x')^{-1} - y^{-1} \leq \sqrt{2\pi}\). Now use the Cauchy–Schwarz inequality to estimate \(|f(y) - f(x)| = |f(y) - f(x')| = \int_{x'}^{y} 1 \cdot f'(t)dt|\).)
The dual of $L^p$

10.1. The dual of $L^p$, $p < \infty$

By the Hölder inequality every $g \in L^q(X, d\mu)$ gives rise to a linear functional on $L^p(X, d\mu)$ and this clearly raises the question if every linear functional is of this form. For $1 \leq p < \infty$ this is indeed the case:

**Theorem 10.1.** Consider $L^p(X, d\mu)$ with some $\sigma$-finite measure $\mu$ and let $q$ be the corresponding dual index, $\frac{1}{p} + \frac{1}{q} = 1$. Then the map $g \in L^q \mapsto \ell_g \in (L^p)^*$ given by

$$\ell_g(f) = \int_X g f \, d\mu$$

(10.1)

is an isometric isomorphism for $1 \leq p < \infty$. If $p = \infty$ it is at least isometric.

**Proof.** Given $g \in L^q$ it follows from Hölder’s inequality that $\ell_g$ is a bounded linear functional with $\|\ell_g\| \leq \|g\|_q$. Moreover, $\|\ell_g\| = \|g\|_q$ follows from Lemma 8.5.

To show that this map is surjective if $1 \leq p < \infty$, first suppose $\mu(X) < \infty$ and choose some $\ell \in (L^p)^*$. Since $\|\chi_A\|_p = \mu(A)^{1/p}$, we have $\chi_A \in L^p$ for every $A \in \Sigma$ and we can define

$$\nu(A) = \ell(\chi_A).$$

Suppose $A = \bigcup_{j=1}^\infty A_j$, where the $A_j$’s are disjoint. Then, by dominated convergence, $\|\sum_{j=1}^n \chi_{A_j} - \chi_A\|_p \to 0$ (this is false for $p = \infty$!) and hence

$$\nu(A) = \ell(\sum_{j=1}^\infty \chi_{A_j}) = \sum_{j=1}^\infty \ell(\chi_{A_j}) = \sum_{j=1}^\infty \nu(A_j).$$
Thus $\nu$ is a complex measure. Moreover, $\mu(A) = 0$ implies $\chi_A = 0$ in $L^p$ and hence $\nu(A) = \ell(\chi_A) = 0$. Thus $\nu$ is absolutely continuous with respect to $\mu$ and by the complex Radon–Nikodym theorem $d\nu = gd\mu$ for some $g \in L^1(X,d\mu)$. In particular, we have

$$\ell(f) = \int_X fg\,d\mu$$

for every simple function $f$. Next let $A_n = \{x||g| < n\}$, then $g_n = g\chi_{A_n} \in L^q$ and by Lemma 8.5 we conclude $\|g_n\|_q \leq \|\ell\|$. Letting $n \to \infty$ shows $g \in L^q$ and finishes the proof for finite $\mu$.

If $\mu$ is $\sigma$-finite, let $X_n \uparrow X$ with $\mu(X_n) < \infty$. Then for every $n$ there is some $g_n$ on $X_n$ and by uniqueness of $g_n$ we must have $g_n = g_m$ on $X_n \cap X_m$. Hence there is some $g$ and by $\|g_n\| \leq \|\ell\|$ independent of $n$, we have $g \in L^q$.

**Corollary 10.2.** Let $\mu$ be some $\sigma$-finite measure. Then $L^p(X,d\mu)$ is reflexive for $1 < p < \infty$.

**Proof.** Identify $L^p(X,d\mu)^*$ with $L^q(X,d\mu)$ and choose $h \in L^p(X,d\mu)^*$. Then there is some $f \in L^p(X,d\mu)$ such that

$$h(g) = \int g(x)f(x)d\mu(x), \quad g \in L^q(X,d\mu) \equiv L^p(X,d\mu)^*.$$ 

But this implies $h(g) = g(f)$, that is, $h = J(f)$, and thus $J$ is surjective. $\square$

Note that in the case $0 < p < 1$, where $L^p$ fails to be a Banach space, the dual might even be empty (see Problem 10.1)!

**Problem 10.1.** Show that $L^p(0,1)$ is a quasinormed space if $0 < p < 1$ (cf. Problem 1.20). Moreover, show that $L^p(0,1)^* = \{0\}$ in this case. (Hint: Suppose there were a nontrivial $\ell \in L^p(0,1)^*$. Start with $f_0 \in L^p$ such that $|\ell(f_0)| \geq 1$. Set $g_0 = \chi_{(0,1]}f$ and $h_0 = \chi_{(s,1]}f$, where $s \in (0,1)$ is chosen such that $\|g_0\|_p = \|h_0\|_p = 2^{-1/p}\|f_0\|_p$. Then $|\ell(g_0)| \geq \frac{1}{2}$ or $|\ell(h_0)| \geq \frac{1}{2}$ and we set $f_1 = 2g_0$ in the first case and $f_1 = 2h_0$ else. Iterating this procedure gives a sequence $f_n$ with $|\ell(f_n)| \geq 1$ and $\|f_n\|_p = 2^{n-1/p}\|f_0\|_p$.)

10.2. The dual of $L^\infty$ and the Riesz representation theorem

In the last section we have computed the dual space of $L^p$ for $p < \infty$. Now we want to investigate the case $p = \infty$. Recall that we already know that the dual of $L^\infty$ is much larger than $L^1$ since it cannot be separable in general.

**Example.** Let $\nu$ be a complex measure. Then

$$\ell_\nu(f) = \int_X f\,d\nu \quad (10.2)$$
10.2. The dual of $L^\infty$ and the Riesz representation theorem

is a bounded linear functional on $B(X)$ (the Banach space of bounded measurable functions) with norm

$$\|\ell_\nu\| = |\nu|(X)$$

by (9.23) and Corollary 9.18. If $\nu$ is absolutely continuous with respect to $\mu$, then it will even be a bounded linear functional on $L^\infty(X, d\mu)$ since the integral will be independent of the representative in this case.

So the dual of $B(X)$ contains all complex measures. However, this is still not all of $B(X)^*$. In fact, it turns out that it suffices to require only finite additivity for $\nu$.

Let $(X, \Sigma)$ be a measurable space. A complex content $\nu$ is a map $\nu : \Sigma \to \mathbb{C}$ such that (finite additivity)

$$\nu\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \nu(A_k), \quad A_j \cap A_k = \emptyset, \ j \neq k.$$  

(10.4)

Given a content $\nu$ we can define the corresponding integral for simple functions $s(x) = \sum_{k=1}^{n} \alpha_k \chi_{A_k}$ as usual

$$\int_A s \, d\nu = \sum_{k=1}^{n} \alpha_k \nu(A_k \cap A).$$  

(10.5)

As in the proof of Lemma 7.17 one shows that the integral is linear. Moreover,

$$|\int_A s \, d\nu| \leq |\nu|(A) \|s\|_\infty,$$  

(10.6)

where $|\nu|(A)$ is defined as in (9.15) and the same proof as in Theorem 9.12 shows that $|\nu|$ is a content. However, since we do not require $\sigma$-additivity, it is not clear that $|\nu|(X)$ is finite. Hence we will call $\nu$ finite if $|\nu|(X) < \infty$. Moreover, for a finite content this integral can be extended to all of $B(X)$ such that

$$|\int_X f \, d\nu| \leq |\nu|(X) \|f\|_\infty$$  

(10.7)

by Theorem 1.35 (compare Problem 7.15). However, note that our convergence theorems (monotone convergence, dominated convergence) will no longer hold in this case (unless $\nu$ happens to be a measure).

In particular, every complex content gives rise to a bounded linear functional on $B(X)$ and the converse also holds:

**Theorem 10.3.** Every bounded linear functional $\ell \in B(X)^*$ is of the form

$$\ell(f) = \int_X f \, d\nu$$  

(10.8)

for some unique finite complex content $\nu$ and $\|\ell\| = |\nu|(X)$.
Proof. Let $\ell \in B(X)^*$ be given. If there is a content $\nu$ at all it is uniquely determined by $\nu(A) = \ell(\chi_A)$. Using this as definition for $\nu$, we see that finite additivity follows from linearity of $\ell$. Moreover, (10.8) holds for characteristic functions and by

$$
\ell\left(\sum_{k=1}^{n} \alpha_k \chi_{A_k}\right) = \sum_{k=1}^{n} \alpha_k \nu(A_k) = \sum_{k=1}^{n} |\nu(A_k)|, \quad \alpha_k = \text{sign}(\nu(A_k)),
$$

we see $|\nu|(X) \leq \|\ell\|$.

Since the characteristic functions are total, (10.8) holds everywhere by continuity and (10.7) shows $\|\ell\| = |\nu|(X)$. □

Remark: To obtain the dual of $L^\infty(X, d\mu)$ from this you just need to restrict to those linear functionals which vanish on $\mathcal{N}(X, d\mu)$ (cf. Problem 10.2), that is, those whose content is absolutely continuous with respect to $\mu$ (note that the Radon–Nikodym theorem does not hold unless the content is a measure).

Example. Consider $B(\mathbb{R})$ and define

$$
\ell(f) = \lim_{\varepsilon \downarrow 0} \lambda f(-\varepsilon) + (1 - \lambda)f(\varepsilon), \quad \lambda \in [0, 1],
$$

for $f$ in the subspace of bounded measurable functions which have left and right limits at 0. Since $\|\ell\| = 1$ we can extend it to all of $B(\mathbb{R})$ using the Hahn–Banach theorem. Then the corresponding content $\nu$ is no measure:

$$
\lambda = \nu([-1, 0)) = \nu\left(\bigcup_{n=1}^{\infty} \left[\frac{-1}{n}, \frac{-1}{n+1}\right)\right) \neq \sum_{n=1}^{\infty} \nu\left(\left[\frac{-1}{n}, \frac{-1}{n+1}\right)\right) = 0. \quad (10.10)
$$

Observe that the corresponding distribution function (defined as in (7.3)) is nondecreasing but not right continuous! If we render $\nu$ right continuous, we get the distribution function of the Dirac measure (centered at 0). In addition, the Dirac measure has the same integral at least for continuous functions!

Theorem 10.4 (Riesz representation). Let $I = [a, b] \subseteq \mathbb{R}$ be a compact interval. Every bounded linear functional $\ell \in C(I)^*$ is of the form

$$
\ell(f) = \int_{I} f \, d\nu \quad (10.11)
$$

for some unique complex Borel measure $\nu$ and $\|\ell\| = |\nu|(I)$.

Moreover, in the case $I = \mathbb{R}$ every bounded linear functional $\ell \in C_0(\mathbb{R})^*$ is of the above form.

Proof. Extending $\ell$ to a bounded linear functional $\tilde{\ell} \in B(I)^*$ we have a corresponding content $\tilde{\nu}$. Splitting this content into real and imaginary part
we see that it is no restriction to assume that $\tilde{\nu}$ is real. Moreover, splitting $\tilde{\nu}$ into $\tilde{\nu}_\pm = (|\tilde{\nu}| \pm \tilde{\nu})/2$ it is no restriction to assume $\tilde{\nu}$ is positive.

Now the idea is as follows: Define a distribution function for $\tilde{\nu}$ as in (7.3). By finite additivity of $\tilde{\nu}$ it will be nondecreasing and we can use Theorem 7.3 to obtain an associated measure $\nu$ whose distribution function coincides with $\tilde{\nu}$ except possibly at points where $\nu$ is discontinuous. It remains to show that the corresponding integral coincides with $\ell$ for continuous functions.

Let $f \in C(I)$ be given. Fix points $a < x_0^n < x_1^n < \ldots < x_n^n < b$ such that $x_0^n \to a$, $x_n^n \to b$, and $\sup_k |x_k^n - x_{k-1}^n| \to 0$ as $n \to \infty$. Then the sequence of simple functions

$$f_n(x) = f(x_0^n)\chi_{[x_0^n,x_1^n]} + f(x_1^n)\chi_{[x_1^n,x_2^n]} + \cdots + f(x_{n-1}^n)\chi_{[x_{n-1}^n,x_n^n]}.$$ 

converges uniformly to $f$ by continuity of $f$ (and the fact that $f$ vanishes as $x \to \pm \infty$ in the case $I = \mathbb{R}$). Moreover,

$$\int_I f d\nu = \lim_{n \to \infty} \int_I f_n d\nu = \lim_{n \to \infty} \sum_{k=1}^n f(x_{k-1}^n)(\nu(x_k^n) - \nu(x_{k-1}^n))$$

$$= \lim_{n \to \infty} \sum_{k=1}^n f(x_{k-1}^n)(\tilde{\nu}(x_k^n) - \tilde{\nu}(x_{k-1}^n)) = \lim_{n \to \infty} \int_I f_n d\tilde{\nu}$$

$$= \int_I f d\tilde{\nu} = \ell(f)$$

provided the points $x_k^n$ are chosen to stay away from all discontinuities of $\nu(x)$ (recall that there are at most countably many).

To see $\|\ell\| = |\nu|(I)$ recall $d\nu = hd|\nu|$ where $|h| = 1$ (Corollary 9.18). Now choose continuous functions $h_n(x) \to h(x)$ pointwise a.e. Using $h_n = \max(1,|h_n|)\text{sign}(h_n)$ we even get such a sequence with $|h_n| \leq 1$. Hence

$$\ell(h_n) = \int h_n^* h d|\nu| \to \int |h|^2 d|\nu| = |\nu|(I)$$

implying $\|\ell\| \geq |\nu|(I)$. The converse follows from (10.7). \hfill $\square$

Note that $\nu$ will be a positive measure if $\ell$ is a positive functional, that is, $\ell(f) \geq 0$ whenever $f \geq 0$.

**Problem 10.2.** Let $M$ be a closed subspace of a Banach space $X$. Show that $(X/M)^* \cong \{\ell \in X^*|M \subseteq \text{Ker}(\ell)\}$. (cf. Theorem 4.22).

**Problem 10.3** (Vague convergence of measures). Let $I$ be a compact interval. A sequence of measures $\nu_n$ is said to converge vaguely to a measure $\nu$ if

$$\int_I f d\nu_n \to \int_I f d\nu, \quad f \in C(I).$$

Show that every bounded sequence of measures has a vaguely convergent subsequence. Show that the limit $\nu$ is a positive measure if all $\nu_n$ are.
(Hint: Compare this definition to the definition of weak-∗ convergence in Section 4.5.)
Chapter 11

The Fourier transform

11.1. The Fourier transform on $L^1$ and $L^2$

For $f \in L^1(\mathbb{R}^n)$ we define its Fourier transform via

$$\mathcal{F}(f)(p) \equiv \hat{f}(p) = (2\pi)^{n/2} \int_{\mathbb{R}^n} e^{-i px} f(x) d^nx. \quad (11.1)$$

Here $px = p_1 x_1 + \cdots + p_n x_n$ is the usual scalar product in $\mathbb{R}^n$ and we will use $|x| = \sqrt{x_1^2 + \cdots + x_n^2}$ for the Euclidean norm.

Lemma 11.1. The Fourier transform is a bounded map from $L^1(\mathbb{R}^n)$ into $C_b(\mathbb{R}^n)$ satisfying

$$\|\hat{f}\|_{\infty} \leq (2\pi)^{-n/2} \|f\|_1. \quad (11.2)$$

Proof. Since $|e^{-ipx}| = 1$ the estimate (11.2) is immediate from

$$|\hat{f}(p)| \leq \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} |e^{-ipx} f(x)| d^nx = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} |f(x)| d^nx.$$

Moreover, a straightforward application of the dominated convergence theorem shows that $\hat{f}$ is continuous. \hfill \Box

Note that if $f$ is nonnegative we have equality: $\|\hat{f}\|_{\infty} = (2\pi)^{-n/2} \|f\|_1 = \hat{f}(0)$.

The following simple properties are left as an exercise.
Lemma 11.2. Let \( f \in L^1(\mathbb{R}^n) \). Then
\[
(f(x + a))^\wedge (p) = e^{iap} \hat{f}(p), \quad a \in \mathbb{R}^n, \\
(e^{ia} f(x))^\wedge (p) = \hat{f}(p - a), \quad a \in \mathbb{R}^n, \\
(f(\lambda x))^\wedge (p) = \frac{1}{\lambda^n} \hat{f}(\frac{p}{\lambda}), \quad \lambda > 0, \\
(f(-x))^\wedge (p) = (f)^\wedge (-p).
\]

Next we look at the connection with differentiation.

Lemma 11.3. Suppose \( f \in C^1(\mathbb{R}^n) \) such that \( \lim_{|x| \to \infty} f(x) = 0 \) and \( f, \partial_j f \in L^1(\mathbb{R}^n) \) for some \( 1 \leq j \leq n \). Then
\[
(\partial_j f)^\wedge (p) = ip_j \hat{f}(p).
\]
Similarly, if \( f(x), x_j f(x) \in L^1(\mathbb{R}^n) \) for some \( 1 \leq j \leq n \), then \( \hat{f}(p) \) is differentiable with respect to \( p_j \) and
\[
(x_j f(x))^\wedge (p) = i\partial_j \hat{f}(p).
\]

Proof. First of all, by integration by parts, we see
\[
(\partial_j f)^\wedge (p) = \frac{1}{(2\pi)^n/2} \int_{\mathbb{R}^n} e^{-ipx} \frac{\partial}{\partial x_j} f(x) d^n x \\
= \frac{1}{(2\pi)^n/2} \int_{\mathbb{R}^n} e^{-ipx} \frac{\partial}{\partial x_j} \hat{f}(x) d^n x \\
= \frac{1}{(2\pi)^n/2} \int_{\mathbb{R}^n} ip_j e^{-ipx} f(x) d^n x = ip_j \hat{f}(p).
\]
Similarly, the second formula follows from
\[
(x_j f(x))^\wedge (p) = \frac{1}{(2\pi)^n/2} \int_{\mathbb{R}^n} x_j e^{-ipx} f(x) d^n x \\
= \frac{1}{(2\pi)^n/2} \int_{\mathbb{R}^n} \left( i \frac{\partial}{\partial p_j} e^{-ipx} \right) f(x) d^n x = i \frac{\partial}{\partial p_j} \hat{f}(p),
\]
where interchanging the derivative and integral is permissible by Problem 7.20. In particular, \( \hat{f}(p) \) is differentiable.

This result immediately extends to higher derivatives. To this end let \( C^\infty(\mathbb{R}^n) \) be the set of all complex-valued functions which have partial derivatives of arbitrary order. For \( f \in C^\infty(\mathbb{R}^n) \) and \( \alpha \in \mathbb{N}_0^n \) we set
\[
\partial_\alpha f = \frac{\partial^{\alpha_1} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}, \quad x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \quad |\alpha| = \alpha_1 + \cdots + \alpha_n.
\]

\[11. The Fourier transform\]
An element $\alpha \in \mathbb{N}_0^n$ is called a multi-index and $|\alpha|$ is called its order. We will also set $(\lambda x)^\alpha = \lambda^{|\alpha|} x^\alpha$ for $\lambda \in \mathbb{R}$. Recall the Schwartz space

$$\mathcal{S}(\mathbb{R}^n) = \{ f \in C^\infty(\mathbb{R}^n) | \sup_x |x^\alpha (\partial_\beta f)(x)| < \infty, \forall \alpha, \beta \in \mathbb{N}_0^n \}$$  \hspace{1cm} (11.10)

which is a subspace of $L^p(\mathbb{R}^n)$ and which is dense for $1 \leq p < \infty$ (since $C_\infty^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$). Together with the seminorms $\|x^\alpha (\partial_\beta f)(x)\|_\infty$ it is a Fréchet space. Note that if $f \in \mathcal{S}(\mathbb{R}^n)$, then the same is true for $x^\alpha f(x)$ and $(\partial_\alpha f)(x)$ for every multi-index $\alpha$.

**Lemma 11.4.** The Fourier transform satisfies $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$. Furthermore, for every multi-index $\alpha \in \mathbb{N}_0^n$ and every $f \in \mathcal{S}(\mathbb{R}^n)$ we have

$$(\partial_\alpha f)^\wedge(p) = (ip)^\alpha \hat{f}(p), \quad (x^\alpha f(x))^\wedge(p) = i^{|\alpha|} \partial_\alpha \hat{f}(p).$$  \hspace{1cm} (11.11)

**Proof.** The formulas are immediate from the previous lemma. To see that $\hat{f} \in \mathcal{S}(\mathbb{R}^n)$ if $f \in \mathcal{S}(\mathbb{R}^n)$, we begin with the observation that $\hat{f}$ is bounded by (11.2). But then $p^\alpha (\partial_\beta \hat{f})(p) = i^{-|\alpha|}|\beta| (\partial_\alpha x^\beta f(x))^\wedge(p)$ is bounded since $\partial_\alpha x^\beta f(x) \in \mathcal{S}(\mathbb{R}^n)$ if $f \in \mathcal{S}(\mathbb{R}^n)$.

Hence we will sometimes write $pf(x)$ for $-i\partial f(x)$, where $\partial = (\partial_1, \ldots, \partial_n)$ is the gradient. Roughly speaking this lemma shows that the decay of a functions is related to the smoothness of its Fourier transform and the smoothness of a functions is related to the decay of its Fourier transform.

In particular, this allows us to conclude that the Fourier transform of an integrable function will vanish at $\infty$. Recall that we denote the space of all continuous functions $f : \mathbb{R}^n \rightarrow \mathbb{C}$ which vanish at $\infty$ by $C_0(\mathbb{R}^n)$.

**Corollary 11.5** (Riemann-Lebesgue). The Fourier transform maps $L^1(\mathbb{R}^n)$ into $C_0(\mathbb{R}^n)$.

**Proof.** First of all recall that $C_0(\mathbb{R}^n)$ equipped with the sup norm is a Banach space and that $\mathcal{S}(\mathbb{R}^n)$ is dense (Problem 1.42). By the previous lemma we have $\hat{f} \in C_0(\mathbb{R}^n)$ if $f \in \mathcal{S}(\mathbb{R}^n)$. Moreover, since $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$, the estimate (11.2) shows that the Fourier transform extends to a continuous map from $L^1(\mathbb{R}^n)$ into $C_0(\mathbb{R}^n)$.

Next we will turn to the inversion of the Fourier transform. As a preparation we will need the Fourier transform of a Gaussian.

**Lemma 11.6.** We have $e^{-z|x|^2/2} \in \mathcal{S}(\mathbb{R}^n)$ for $\text{Re}(z) > 0$ and

$$\mathcal{F}(e^{-z|x|^2/2})(p) = \frac{1}{z^{n/2}} e^{-|p|^2/(2z)}.$$  \hspace{1cm} (11.12)

Here $z^{n/2}$ is the standard branch with branch cut along the negative real axis.
Proof. Due to the product structure of the exponential, one can treat each coordinate separately, reducing the problem to the case \( n = 1 \) (Problem 11.3).

Let \( \phi_z(x) = \exp(-zx^2/2) \). Then \( \phi_z'(x) + zx\phi_z(x) = 0 \) and hence \( i(p\hat{\phi}_z(p) + z\hat{\phi}_z'(p)) = 0 \). Thus \( \hat{\phi}_z(p) = c\phi_{1/z}(p) \) and (Problem 7.26)

\[
c = \hat{\phi}_z(0) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp(-zx^2/2) dx = \frac{1}{\sqrt{z}}
\]

at least for \( z > 0 \). However, since the integral is holomorphic for \( \text{Re}(z) > 0 \) by Problem 7.22, this holds for all \( z \) with \( \text{Re}(z) > 0 \) if we choose the branch cut of the root along the negative real axis. \( \square \)

Now we can show

**Theorem 11.7.** The Fourier transform is a bounded injective map from \( L^1(\mathbb{R}^n) \) into \( C_0(\mathbb{R}^n) \). Its inverse is given by

\[
f(x) = \lim_{\varepsilon \downarrow 0} \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ipx-\varepsilon|p|^2/2} \hat{f}(p) d^n p,
\]

where the limit has to be understood in \( L^1 \).

**Proof.** Abbreviate \( \phi_\varepsilon(x) = (2\pi)^{-n/2} \exp(-\varepsilon|x|^2/2) \). Then the right-hand side is given by

\[
\int_{\mathbb{R}^n} \phi_\varepsilon(p) e^{ipx} \hat{f}(p) d^n p = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \phi_\varepsilon(p) e^{ipx} f(y) e^{-ipy} d^n y d^n p
\]

and, invoking Fubini and Lemma 11.2, we further see that this is equal to

\[
= \int_{\mathbb{R}^n} (\phi_\varepsilon(p)e^{-ipx})^* f(y) d^n y = \int_{\mathbb{R}^n} \frac{1}{\varepsilon^{n/2}} \phi_{1/\varepsilon}(y-x) f(y) d^n y.
\]

But the last integral converges to \( f \) in \( L^1(\mathbb{R}^n) \) by Lemma 8.15. \( \square \)

Of course when \( \hat{f} \in L^1(\mathbb{R}^n) \), the limit is superfluous and we obtain

**Corollary 11.8.** Suppose \( f, \hat{f} \in L^1(\mathbb{R}^n) \). Then

\[
(\hat{f})^\vee = f,
\]

where

\[
\hat{f}(p) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ipx} f(x) d^n x = \hat{f}(-p).
\]

In particular, \( F : F^1(\mathbb{R}^n) \to F^1(\mathbb{R}^n) \) is a bijection, where \( F^1(\mathbb{R}^n) = \{ f \in L^1(\mathbb{R}^n) | \hat{f} \in L^1(\mathbb{R}^n) \} \). Moreover, \( \mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n) \) is a bijection.
11.1. The Fourier transform on $L^1$ and $L^2$

However, note that $\mathcal{F} : L^1(\mathbb{R}^n) \to C_0(\mathbb{R}^n)$ is not onto (cf. Problem 11.6). Nevertheless the inverse Fourier transform $\mathcal{F}^{-1}$ is a closed map from $\text{Ran}(\mathcal{F}) \to L^1(\mathbb{R}^n)$ by Lemma 4.7. As $\mathcal{F}^{-1}$ and $\mathcal{F}$ differ only by a reflection in the argument, the same is true for $\mathcal{F}$.

Lemma 11.9. Suppose $f \in F^1(\mathbb{R}^n)$. Then $f, \hat{f} \in L^2(\mathbb{R}^n)$ and

$$\|f\|_2^2 = \|\hat{f}\|_2^2 \leq (2\pi)^{-n/2} \|f\|_1 \|\hat{f}\|_1$$

(11.16) holds.

Proof. This follows from Fubini’s theorem since

$$\int_{\mathbb{R}^n} |\hat{f}(p)|^2 d^n p = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x)^* \hat{f}(p) e^{ipx} d^n p d^n x$$

$$= \int_{\mathbb{R}^n} |f(x)|^2 d^n x$$

for $f, \hat{f} \in L^1(\mathbb{R}^n)$. □

In fact, we have $F^1(\mathbb{R}^n) \subset L^1(\mathbb{R}^n) \cap C_0(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ for any $p \in [1, \infty]$ and the inequality in (11.16) can be generalized to arbitrary $p$ (cf. Problem 11.2).

The identity $\|f\|_2 = \|\hat{f}\|_2$ is known as the Plancherel identity. Thus, by Theorem 1.35, we can extend $\mathcal{F}$ to all of $L^2(\mathbb{R}^n)$ by setting $\mathcal{F}(f) = \lim_{m \to \infty} \mathcal{F}(f_m)$, where $f_m$ is an arbitrary sequence from, say, $\mathcal{S}(\mathbb{R}^n)$ converging to $f$ in the $L^2$ norm.

Theorem 11.10 (Plancherel). The Fourier transform $\mathcal{F}$ extends to a unitary operator $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$.

Proof. As already noted, $\mathcal{F}$ extends uniquely to a bounded operator on $L^2(\mathbb{R}^n)$. Since Plancherel’s identity remains valid by continuity of the norm and since its range is dense, this extension is a unitary operator. □

We also note that this extension is still given by (11.1) whenever the right-hand side is integrable.

Lemma 11.11. Let $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then (11.1) continues to hold, where $\mathcal{F}$ now denotes the extension of the Fourier transform from $\mathcal{S}(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n)$.

Proof. If $f$ has compact support, then by Lemma 8.14 its mollification $\phi_\varepsilon * f \in C^\infty_c(\mathbb{R}^n)$ converges to $f$ both in $L^1$ and $L^2$. Hence the claim holds for every $f$ with compact support. Finally, for general $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ consider $f_m = f_{\chi_{B_m(0)}}$. Then $f_m \to f$ in both $L^1(\mathbb{R}^n)$ and $L^2(\mathbb{R}^n)$ and the claimed follows. □
In particular,

\[
\hat{f}(p) = \lim_{m \to \infty} \frac{1}{(2\pi)^{n/2}} \int_{|x| \leq m} e^{-ipx} f(x) d^n x,
\]

(11.17)

where the limit has to be understood in \(L^2(\mathbb{R}^n)\) and can be omitted if \(f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)\).

Another useful property is the convolution formula.

**Lemma 11.12.** The convolution

\[
(f * g)(x) = \int_{\mathbb{R}^n} f(y) g(x-y) d^n y = \int_{\mathbb{R}^n} f(x-y) g(y) d^n y
\]

(11.18)

of two functions \(f, g \in L^1(\mathbb{R}^n)\) is again in \(L^1(\mathbb{R}^n)\) and we have **Young’s inequality**

\[
\|f \ast g\|_1 \leq \|f\|_1 \|g\|_1.
\]

(11.19)

Moreover, its Fourier transform is given by

\[
(f \ast g)^\wedge(p) = (2\pi)^{n/2} \hat{f}(p) \hat{g}(p).
\]

(11.20)

**Proof.** The fact that \(f \ast g\) is in \(L^1\) together with Young’s inequality follows by applying Fubini’s theorem to \(h(x,y) = f(x-y)g(y)\) (in fact we have shown a more general version in Lemma 8.14). For the last claim we compute

\[
(f \ast g)^\wedge(p) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ipx} \int_{\mathbb{R}^n} f(y) g(x-y) d^n y d^n x
\]

\[
= \int_{\mathbb{R}^n} e^{-ipy} f(y) \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ip(x-y)} g(x-y) d^n x d^n y
\]

\[
= \int_{\mathbb{R}^n} e^{-ipy} f(y) \hat{g}(p) d^n y = (2\pi)^{n/2} \hat{f}(p) \hat{g}(p),
\]

where we have again used Fubini’s theorem. \(\square\)

In other words, \(L^1(\mathbb{R}^n)\) together with convolution as a product is a Banach algebra (without identity). As a consequence we can also deal with the case of convolution on \(S(\mathbb{R}^n)\) as well as on \(L^2(\mathbb{R}^n)\).

**Corollary 11.13.** The convolution of two \(S(\mathbb{R}^n)\) functions as well as their product is in \(S(\mathbb{R}^n)\) and

\[
(f \ast g)^\wedge = (2\pi)^{n/2} \hat{f} \hat{g}, \quad (fg)^\wedge = (2\pi)^{-n/2} \hat{f} \ast \hat{g}
\]

(11.21)

in this case.

**Proof.** Clearly the product of two functions in \(S(\mathbb{R}^n)\) is again in \(S(\mathbb{R}^n)\) (show this!). Since \(S(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)\) the previous lemma implies \((f \ast g)^\wedge = (2\pi)^{n/2} \hat{f} \hat{g} \in S(\mathbb{R}^n)\). Moreover, since the Fourier transform is injective on \(L^1(\mathbb{R}^n)\) we conclude \(f \ast g = (2\pi)^{n/2}(\hat{f} \hat{g})^\vee \in S(\mathbb{R}^n)\). Replacing \(f, g\) by \(\hat{f}, \hat{g}\)
in the last formula finally shows \( \hat{f} \ast \hat{g} = (2\pi)^{n/2}(fg)^\vee \) and the claim follows by a simple change of variables using \( \hat{f}(p) = \hat{f}(-p) \).

\[\text{Corollary 11.14.} \quad \text{The convolution of two } L^2(\mathbb{R}^n) \text{ functions is in } \text{Ran}(\mathcal{F}) \subset C_0(\mathbb{R}^n) \text{ and we have } \|f \ast g\|_\infty \leq \|f\|_2\|g\|_2 \text{ as well as}
\]
\[(fg)\wedge = (2\pi)^{-n/2}\hat{f} \ast \hat{g}, \quad (f \ast g)\wedge = (2\pi)^{n/2}\hat{f} \hat{g} \quad (11.22)\]
in this case.

\[\text{Proof.} \quad \text{The inequality } \|f \ast g\|_\infty \leq \|f\|_2\|g\|_2 \text{ is immediate from Cauchy–Schwarz and shows that the convolution is a continuous bilinear form from } L^2(\mathbb{R}^n) \text{ to } L^\infty(\mathbb{R}^n). \text{ Now take sequences } f_m, g_m \in \mathcal{S}(\mathbb{R}^n) \text{ converging to } f, g \in L^2(\mathbb{R}^n). \text{ Then using the previous corollary together with continuity of the Fourier transform from } L^1(\mathbb{R}^n) \text{ to } C_0(\mathbb{R}^n) \text{ and on } L^2(\mathbb{R}^n) \text{ we obtain}
\]
\[(fg)\wedge = \lim_{m \to \infty} (f_m g_m)\wedge = (2\pi)^{-n/2}\lim_{m \to \infty} \hat{f}_m \ast \hat{g}_m = (2\pi)^{-n/2}\hat{f} \ast \hat{g}.
\]
Similarly,
\[(f \ast g)\wedge = \lim_{m \to \infty} (f_m \ast g_m)\wedge = (2\pi)^{n/2}\lim_{m \to \infty} \hat{f}_m \hat{g}_m = (2\pi)^{n/2}\hat{f} \hat{g}
\]
from which that last claim follows since \( \mathcal{F} : \text{Ran}(\mathcal{F}) \to L^1(\mathbb{R}^n) \) is closed by Lemma 4.7.

Finally, note that by looking at the Gaussian’s \( \phi_\lambda(x) = \exp(-\lambda x^2/2) \) one observes that a well centered peak transforms into a broadly spread peak and vice versa. This turns out to be a general property of the Fourier transform known as \textbf{uncertainty principle}. One quantitative way of measuring this fact is to look at
\[
\|(x_j - x^0)f(x)\|_2^2 = \int_{\mathbb{R}^n} (x_j - x^0)^2 |f(x)|^2 d^n x \quad (11.23)
\]
which will be small if \( f \) is well concentrated around \( x^0 \) in the \( j \)'th coordinate direction.

\[\text{Theorem 11.15 (Heisenberg uncertainty principle).} \quad \text{Suppose } f \in \mathcal{S}(\mathbb{R}^n). \text{ Then for any } x^0, p^0 \in \mathbb{R} \text{ we have}
\]
\[
\|(x_j - x^0)f(x)\|_2\|(p_j - p^0)\hat{f}(p)\|_2 \geq \frac{\|f\|_2^2}{2}. \quad (11.24)
\]

\[\text{Proof.} \quad \text{Replacing } f(x) \text{ by } e^{ix_jp^0}f(x + x^0e_j) \text{ (where } e_j \text{ is the unit vector into the } j \text{'th coordinate direction)} \text{ we can assume } x^0 = p^0 = 0 \text{ by Lemma 11.2. Using integration by parts we have}
\]
\[
\|f\|_2^2 = \int_{\mathbb{R}^n} |f(x)|^2 d^n x = - \int_{\mathbb{R}^n} x_j \partial_j |f(x)|^2 d^n x = -2\mathrm{Re} \int_{\mathbb{R}^n} x_j f(x)^* \partial_j f(x) d^n x.
\]
Hence, by Cauchy–Schwarz,
\[ \|f\|^2 \leq 2 \|x_j f(x)\| \|\partial_j f(x)\| = 2 \|x_j f(x)\| \|p_j \hat{f}(p)\| \]
the claim follows. \( \square \)

The name stems from quantum mechanics, where \(|f(x)|^2\) is interpreted as the probability distribution for the position of a particle and \(|\hat{f}(x)|^2\) is interpreted as the probability distribution for its momentum. Equation (11.24) says that the variance of both distributions cannot both be small and thus one cannot simultaneously measure position and momentum of a particle with arbitrary precision.

Another version states that \(f\) and \(\hat{f}\) cannot both have compact support.

**Theorem 11.16.** Suppose \(f \in L^2(\mathbb{R}^n)\). If both \(f\) and \(\hat{f}\) have compact support, then \(f = 0\).

**Proof.** Let \(A, B \subseteq \mathbb{R}^n\) be two compact sets and consider the subspace of all functions with \(\text{supp}(f) \subseteq A\) and \(\text{supp}(\hat{f}) \subseteq B\). Then
\[ f(x) = \int_{\mathbb{R}^n} K(x, y) f(y) d^n y, \]
where
\[ K(x, y) = \frac{1}{(2\pi)^n} \int_B e^{i(x-y)p} \chi_A(y) d^n p = \frac{1}{(2\pi)^n} \hat{\chi}_B(y - x) \chi_A(y). \]
Since \(K \in L^2(\mathbb{R}^n \times \mathbb{R}^n)\) the corresponding integral operator is Hilbert–Schmidt, and thus its eigenspace corresponding to the eigenvalue 1 can be at most finite dimensional.

Now if there is a nonzero \(f\), we can find a sequence of vectors \(x^n \to 0\) such the functions \(f_n(x) = f(x - x^n)\) are linearly independent (look at their supports) and satisfy \(\text{supp}(f_n) \subseteq 2A\), \(\text{supp}(\hat{f}_n) \subseteq B\). But this a contradiction by the first part applied to the sets \(2A\) and \(B\). \( \square \)

**Problem 11.1.** Show that \(S(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)\). (Hint: If \(f \in S(\mathbb{R}^n)\), then \(|f(x)| \leq C_m \prod_{j=1}^n (1 + x_j^2)^{-m}\) for every \(m\).)

**Problem 11.2.** Show that \(F^1(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)\) with
\[ \|f\|_p \leq (2\pi)^{\frac{n}{2}(1 - \frac{1}{p})} \|f\|_1^{\frac{1}{2}} \|f\|_1^{1 - \frac{1}{p}}. \]
Moreover, show that \(S(\mathbb{R}^n) \subset F^1(\mathbb{R}^n)\) and conclude that \(F^1(\mathbb{R}^n)\) is dense in \(L^p(\mathbb{R}^n)\) for \(p \in [1, \infty)\). (Hint: Use \(x_p \leq x\) for \(0 \leq x \leq 1\) to show \(\|f\|_p \leq \|f\|_1^{\frac{1}{p}}\|f\|_1^{1 - \frac{1}{p}}\).

**Problem 11.3.** Suppose \(f_j \in L^1(\mathbb{R}), j = 1, \ldots, n\) and set \(f(x) = \prod_{j=1}^n f_j(x_j)\). Show that \(f \in L^1(\mathbb{R}^n)\) with \(\|f\|_1 = \prod_{j=1}^n |f_j|_1\) and \(\hat{f}(p) = \prod_{j=1}^n \hat{f}_j(p_j)\).
Problem 11.4. Compute the Fourier transform of the following functions $f : \mathbb{R} \to \mathbb{C}$:

(i) $f(x) = \chi_{(-1,1)}(x)$. 
(ii) $f(p) = \frac{1}{p^2 + k^2}$, $\text{Re}(k) > 0$.

Problem 11.5. A function $f : \mathbb{R}^n \to \mathbb{C}$ is called spherically symmetric if it is invariant under rotations; that is, $f(Ox) = f(x)$ for all $O \in SO(\mathbb{R}^n)$ (equivalently, $f$ depends only on the distance to the origin $|x|$). Show that the Fourier transform of a spherically symmetric function is again spherically symmetric.

Problem 11.6. Show that $\mathcal{F} : L^1(\mathbb{R}^n) \to C_0(\mathbb{R}^n)$ is not onto as follows:

(i) The range of $\mathcal{F}$ is dense.
(ii) $\mathcal{F}$ is onto if and only if it has a bounded inverse.
(iii) $\mathcal{F}$ has no bounded inverse.

(Hint for (iii) in the case $n = 1$: Investigate $f_m = \chi_{(-1,1)} * \chi_{(-m,m)}$.)

Problem 11.7 (Wiener). Suppose $f \in L^2(\mathbb{R}^n)$. Then the set $\{f(x+a) | a \in \mathbb{R}^n\}$ is total in $L^2(\mathbb{R}^n)$ if and only if $\hat{f}(p) \neq 0$ a.e. (Hint: Use Lemma 11.2 and the fact that a subspace is total if and only if its orthogonal complement is zero.)

Problem 11.8. Suppose $f(x)e^{k|x|} \in L^1(\mathbb{R})$ for some $k > 0$. Then $\hat{f}(p)$ has an analytic extension to the strip $|\text{Im}(p)| < k$.

11.2. Applications to linear partial differential equations

By virtue of Lemma 11.4 the Fourier transform can be used to map linear partial differential equations with constant coefficients to algebraic equations, thereby providing a mean of solving them. To illustrate this procedure we look at the famous Poisson equation, that is, given a function $g$, find a function $f$ satisfying

$$-\Delta f = g.$$  \hfill (11.25)

For simplicity, let us start by investigating this problem in the space of Schwartz functions $S(\mathbb{R}^n)$. Assuming there is a solution we can take the Fourier transform on both sides to obtain

$$|p|^2 \hat{f}(p) = \hat{g}(p) \quad \Rightarrow \quad \hat{f}(p) = |p|^{-2} \hat{g}(p).$$  \hfill (11.26)

Since the right-hand side is integrable for $n \geq 3$ we obtain that our solution is necessarily given by

$$f(x) = (|p|^{-2} \hat{g}(p))^\vee(x).$$  \hfill (11.27)

In fact, this formula still works provided $g(x), |p|^{-2} \hat{g}(p) \in L^1(\mathbb{R}^n)$. Moreover, if we additionally assume $\hat{g} \in L^1(\mathbb{R}^n)$, then $|p|^2 \hat{f}(p) = \hat{g}(p) \in L^1(\mathbb{R}^n)$ and Lemma 11.3 implies that $f \in C^2(\mathbb{R}^n)$ as well as that it is indeed a solution.
Note that if $n \geq 3$, then $|p|^{-2} \hat{g}(p) \in L^1(\mathbb{R}^n)$ follows automatically from $g, \hat{g} \in L^1(\mathbb{R}^n)$ (show this!).

Moreover, we clearly expect that $f$ should be given by a convolution. However, since $|p|^{-2}$ is not in $L^p(\mathbb{R}^n)$ for any $p$, the formulas derived so far do not apply.

**Lemma 11.17.** Let $0 < \alpha < n$ and suppose $g \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ as well as $|p|^{-\alpha} \hat{g}(p) \in L^1(\mathbb{R}^n)$. Then

$$
(|p|^{-\alpha} \hat{g}(p))^\vee(x) = \int_{\mathbb{R}^n} I_\alpha(|x-y|)g(y)d^n y,
$$

(11.28)

where the Riesz potential is given by

$$
I_\alpha(r) = \frac{\Gamma\left(\frac{n-\alpha}{2}\right)}{2^{\alpha} \pi^{n/2} \Gamma\left(\frac{\alpha}{2}\right)} \frac{1}{r^{\alpha-n}}.
$$

(11.29)

**Proof.** Note that, while $|.|^{-\alpha}$ is not in $L^p(\mathbb{R}^n)$ for any $p$, our assumption $0 < \alpha < n$ ensures that the singularity at zero is integrable.

We set $\phi_t(p) = \exp(-t|p|^2/2)$ and begin with the elementary formula

$$
|p|^{-\alpha} = c_\alpha \int_0^\infty \phi_t(p)t^{\alpha/2-1}dt,
$$

where the Riesz potential is given by

$$
I_\alpha(r) = \frac{\Gamma\left(\frac{n-\alpha}{2}\right)}{2^{\alpha} \pi^{n/2} \Gamma\left(\frac{\alpha}{2}\right)} \frac{1}{r^{\alpha-n}}.
$$

(11.29)

which follows from the definition of the gamma function (Problem 7.27) after a simple scaling. Since $|p|^{-\alpha} \hat{g}(p)$ is integrable we can use Fubini and Lemma 11.6 to obtain

$$
(|p|^{-\alpha} \hat{g}(p))^\vee(x) = \frac{c_\alpha}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \exp\left(\int_0^\infty \phi_t(p)t^{\alpha/2-1}dt\right) \hat{g}(p)d^n p.
$$

Since $\phi, g \in L^1$ we know by Lemma 11.12 that $\hat{\phi} \hat{g} = (2\pi)^{-n/2} (\phi * g)^\vee$ Moreover, since $\hat{\phi} \hat{g} \in L^1$ Theorem 11.7 gives us $(\hat{\phi} \hat{g})^\vee = (2\pi)^{-n/2} \phi * g$. Thus, we can make a change of variables and use Fubini once again (since $g \in L^\infty$)

$$
(|p|^{-\alpha} \hat{g}(p))^\vee(x) = \frac{c_\alpha}{(2\pi)^{n/2}} \int_0^\infty \left(\int_{\mathbb{R}^n} \phi_{t} (x-y)g(y)d^n y\right) t^{(\alpha-n)/2-1}dt
$$

$$
= \frac{c_\alpha}{(2\pi)^{n/2}} \int_0^\infty \left(\int_{\mathbb{R}^n} \phi_{t} (x-y)g(y)d^n y\right) t^{(\alpha-n)/2-1}dt
$$

$$
= \frac{c_\alpha}{(2\pi)^{n/2}} \int_0^\infty \left(\int_{\mathbb{R}^n} \phi_{t} (x-y)g(y)d^n y\right) t^{(\alpha-n)/2-1}dt
$$

$$
= \frac{c_\alpha/\Gamma(n-\alpha)}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} g(y) |x-y|^{n-\alpha}d^n y
$$

(11.28)

to obtain the desired result. \(\square\)
Note that the conditions of the above theorem are, for example, satisfied if \( g, \hat{g} \in L^1(\mathbb{R}^n) \) which holds, for example, if \( g \in \mathcal{S}(\mathbb{R}^n) \). In summary, if \( g \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n) \), \( |p|^{-2} \hat{g}(p) \in L^1(\mathbb{R}^n) \) and \( n \geq 3 \), then

\[
f = \Phi \ast g
g\tag{11.30}\]

is a classical solution of the Poisson equation, where

\[
\Phi(x) = \frac{\Gamma\left(\frac{n}{2} - 1\right)}{4\pi^{n/2}} \frac{1}{|x|^{n-2}}, \quad n \geq 3,
\tag{11.31}\]

is known as the fundamental solution of the Laplace equation.

A few words about this formula are in order. First of all, our original formula in Fourier space shows that the multiplication with \( |p|^{-2} \) improves the decay of \( \hat{g} \) and hence, by virtue of Lemma 11.4, \( f \) should have, roughly speaking, two derivatives more than \( g \). However, unless \( \hat{g}(0) \) vanishes, multiplication with \( |p|^{-2} \) will create a singularity at 0 and hence, again by Lemma 11.4, \( f \) will not inherit any decay properties from \( g \). In fact, evaluating the above formula with \( g = \chi_{B_1(0)} \) (Problem 11.9) shows that \( f \) might not decay better than \( \Phi \) even for \( g \) with compact support.

Moreover, our conditions on \( g \) might not be easy to check as it will not be possible to compute \( \hat{g} \) explicitly in general. So if one wants to deduce \( \hat{g} \in L^1(\mathbb{R}^n) \) from properties of \( g \), one could use Lemma 11.4 together with the Riemann–Lebesgue lemma to show that this condition holds if \( g \in C^k(\mathbb{R}^n) \), \( k > n-2 \), such that all derivatives are integrable and all derivatives of order less than \( k \) vanish at \( \infty \) (Problem 11.10). This seems a rather strong requirement since our solution formula will already make sense under the sole assumption \( g \in L^1(\mathbb{R}^n) \). However, as the example \( g = \chi_{B_1(0)} \) shows, this solution might not be \( C^2 \) and hence one needs to weaken the notion of a solution if one wants to include such situations. This will lead us to the concepts of weak derivatives and Sobolev spaces. As a preparation we will develop some further tools which will allow us to investigate continuity properties of the operator \( \mathcal{L}_\alpha f = I_\alpha \ast f \) in the next section.

Before that, let us summarize the procedure in the general case. Suppose we have the following linear partial differential equations with constant coefficients:

\[
P(i\partial) f = g, \quad P(i\partial) = \sum_{\alpha \leq k} c_\alpha |\alpha| \partial_\alpha.
\tag{11.32}\]

Then the solution can be found via the procedure
The Fourier transform

\[ \hat{g} \xrightarrow{\mathcal{F}} P^{-1}\hat{g} \]

and is formally given by

\[ f(x) = (P(p)^{-1}\hat{g}(p))^\vee(x). \quad (11.33) \]

It remains to investigate the properties of the solution operator. In general, given a measurable function \( m \) one might try to define a corresponding operator via

\[ A_m f = (m\hat{g})^\vee, \quad (11.34) \]

in which case \( m \) is known as a Fourier multiplier. It is said to be an \( L^p \)-multiplier if \( A_m \) can be extended to a bounded operator in \( L^p(\mathbb{R}^n) \). For example, it will be an \( L^2 \) multiplier if \( m \) is bounded (in fact the converse is also true — Problem 11.11). As we have seen, in some cases \( A_m \) can be expressed as a convolution, but this is not always the case as the trivial example \( m = 1 \) (corresponding to the identity operator) shows.

Another famous example which can be solved in this way is the **Helmholtz equation**

\[ -\Delta f + f = g. \quad (11.35) \]

As before we find that if \( g(x), (1 + |p|^2)^{-1}\hat{g}(p) \in L^1(\mathbb{R}^n) \) then the solution is given by

\[ f(x) = ((1 + |p|^2)^{-1}\hat{g}(p))^\vee(x). \quad (11.36) \]

**Lemma 11.18.** Let \( \alpha > 0 \). Suppose \( g \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n) \) as well as \( (1 + |p|^2)^{-\alpha/2}\hat{g}(p) \in L^1(\mathbb{R}^n) \). Then

\[ ((1 + |p|^2)^{-\alpha/2}\hat{g}(p))^\vee(x) = \int_{\mathbb{R}^n} B_{(n-\alpha)/2}(|x-y|)g(y)dy, \quad (11.37) \]

where the Bessel potential is given by

\[ J_\alpha(r) = \frac{2}{(4\pi)^{n/2}\Gamma(n/2)} \left( \frac{r^2}{2} \right)^{-(n-\alpha)/2} K_{(n-\alpha)/2}(r^2), \quad r > 0, \quad (11.38) \]

with

\[ K_\nu(r) = K_{-\nu}(r) = \frac{1}{2} \left( \frac{r}{2} \right)^\nu \int_0^\infty e^{-t-r/t^{\nu+1}} dt, \quad r > 0, \nu \in \mathbb{R}, \quad (11.39) \]

the modified Bessel function of the second kind of order \( \nu \) ([20, (10.32.10)]).
Proof. We proceed as in the previous lemma. We set \( \phi_t(p) = \exp(-t|p|^2/2) \) and begin with the elementary formula

\[
\frac{\Gamma\left(\frac{\alpha}{2}\right)}{(1 + |p|^2)^{\alpha/2}} = \int_0^\infty t^{\alpha/2 - 1}e^{-t(1+|p|^2)}dt.
\]

Since \( \hat{g}(p) \) are integrable we can use Fubini and Lemma 11.6 to obtain

\[
\left(\frac{\hat{g}(p)}{1 + |p|^2}\right)^\alpha(x) = \frac{\Gamma\left(\frac{\alpha}{2}\right)^{-1}}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{i\phi} \left(\int_0^\infty t^{\alpha/2 - 1}e^{-t(1+|p|^2)}dt\right) \hat{g}(p)d^n p
\]

\[
= \frac{\Gamma\left(\frac{\alpha}{2}\right)^{-1}}{(4\pi)^{n/2}} \int_0^\infty \left(\int_{\mathbb{R}^n} e^{i\phi} \phi_1/2t(\hat{g}(p))d^n p\right) e^{-t(\alpha-n)/2-1}dt
\]

\[
= \frac{\Gamma\left(\frac{\alpha}{2}\right)^{-1}}{(4\pi)^{n/2}} \int_0^\infty \left(\int_{\mathbb{R}^n} \phi_1/2t(x-y)(g(y))d^n y\right) e^{-t(\alpha-n)/2-1}dt
\]

\[
= \frac{\Gamma\left(\frac{\alpha}{2}\right)^{-1}}{(4\pi)^{n/2}} \int_{\mathbb{R}^n} \left(\int_0^\infty \phi_1/2t(x-y)e^{-t(\alpha-n)/2-1}dt\right) g(y)d^n y
\]

to obtain the desired result. Using Fubini in the last step is allowed since \( g \) is bounded and \( J_\alpha(|x|) \in L^1(\mathbb{R}^n) \) (Problem 11.12).

Note that since the first condition \( g \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n) \) implies \( g \in L^2(\mathbb{R}^n) \) and thus the second condition \( (1 + |p|^2)^{-\alpha/2} \hat{g}(p) \in L^1(\mathbb{R}^n) \) will be satisfied if \( \frac{n}{2} < \alpha \).

In particular, if \( g, \hat{g} \in L^1(\mathbb{R}^n) \), then

\[
f = J_1 * g
\]

is a solution of Helmholtz equation. Note that since our multiplier \( (1 + |p|^2)^{-1} \) does not have a singularity near zero, the solution \( f \) will preserve (some) decay properties of \( g \). For example, it will map Schwartz functions to Schwartz functions and thus for every \( g \in S(\mathbb{R}^n) \) there is a unique solution of the Helmholtz equation \( f \in S(\mathbb{R}^n) \). This is also reflected by the fact that the Bessel potential decays much faster than the Riesz potential. Indeed, one can show that [20, (10.25.3)]

\[
K_\nu(r) = \left(\frac{\pi}{2r}\right)^{-\nu}e^{-r(1 + O(r^{-1}))}
\]

as \( r \to \infty \). The singularity near zero is of the same type as for \( I_\alpha \) since (see [20, (10.30.2) and (10.30.3)])

\[
K_\nu(r) = \begin{cases} 
\frac{\Gamma(\nu)}{2} \left(\frac{r}{2}\right)^{-\nu} + O(r^{-\nu+2}), & \nu > 0, \\
-\log\left(\frac{r}{2}\right) + O(1), & \nu = 0,
\end{cases}
\]

for \( r \to 0 \).
Problem 11.9. Show that for \( n \geq 3 \) we have
\[
(\Phi * \chi_{B_1(0)})(x) = \begin{cases} 
\frac{1}{2^n|x|^{n-2}}, & |x| \geq 1, \\
\frac{2^n}{|x|^2}, & |x| \leq 1.
\end{cases}
\]
(Hint: Observe that the result depends only on \(|x|\). Then choose \( x = (0, \ldots, 0, R) \) and evaluate the integral using spherical coordinates.)

Problem 11.10. Suppose \( g \in C^k(\mathbb{R}^n) \) and \( \partial^l_j g \in L^1(\mathbb{R}^n) \) for \( j = 1, \ldots, n \) and \( 0 \leq l \leq k \) as well as \( \lim_{|x| \to \infty} \partial^l_j g(x) = 0 \) for \( j = 1, \ldots, n \) and \( 0 \leq l < k \). Then
\[
|\hat{g}(p)| \leq C \frac{1}{(1 + |p|^2)^{k/2}}.
\]

Problem 11.11. Show that \( m \) is an \( L^2 \) multiplier if and only if \( m \in L^\infty(\mathbb{R}^n) \).

Problem 11.12. Show
\[
\int_0^\infty J_\alpha(r)r^{n-1}dr = \frac{\Gamma(n/2)}{2\pi^{n/2}}, \quad \alpha > 0.
\]
Conclude that
\[
\|J_\alpha * g\|_p \leq \|g\|_p.
\]
(Hint: Fubini.)

11.3. Sobolev spaces

We begin by introducing the Sobolev space
\[
H^r(\mathbb{R}^n) = \{ f \in L^2(\mathbb{R}^n)|||p|^r \hat{f}(p) \in L^2(\mathbb{R}^n)\}. \tag{11.43}
\]
The most important case is when \( r \) is an integer, however our definition makes sense for any \( r \geq 0 \). Moreover, note that \( H^r(\mathbb{R}^n) \) becomes a Hilbert space if we introduce the scalar product
\[
\langle f, g \rangle = \int_{\mathbb{R}^n} \hat{f}(p)^* \hat{g}(p)(1 + |p|^2)^r d^n p. \tag{11.44}
\]
In particular, note that by construction \( \mathcal{F} \) maps \( H^r(\mathbb{R}^n) \) unitarily onto \( L^2(\mathbb{R}^n, (1 + |p|^2)^r d^n p) \). Clearly \( H^{r+1}(\mathbb{R}^n) \subset H^r(\mathbb{R}^n) \) with the embedding being continuous. Moreover, \( \mathcal{S}(\mathbb{R}^n) \subset H^r(\mathbb{R}^n) \) and this subset is dense (since \( \mathcal{S}(\mathbb{R}^n) \) is dense in \( L^2(\mathbb{R}^n, (1 + |p|^2)^r d^n p) \)).

The motivation for the definition (11.43) stems from Lemma 11.4 which allows us to extend differentiation to a larger class. In fact, every function in \( H^r(\mathbb{R}^n) \) has partial derivatives up to order \(|r|\), which are defined via
\[
\partial_\alpha f = ((ip)^\alpha \hat{f}(p))^\vee, \quad f \in H^r(\mathbb{R}^n), \ |\alpha| \leq r. \tag{11.45}
\]
By Lemma 11.4 this definition coincides with the usual one for every \( f \in \mathcal{S}(\mathbb{R}^n) \).
Example. Consider $f(x) = (1 - |x|)\chi_{[-1,1]}(x)$. Then $\hat{f}(p) = \sqrt{\frac{2}{\pi}} \frac{\cos(p) - 1}{p^3}$ and $f \in H^1(\mathbb{R})$. The weak derivative is $f'(x) = -\operatorname{sign}(x)\chi_{[-1,1]}(x)$. \hfill \Diamond

We also have

$$
\int_{\mathbb{R}^n} g(x)(\partial_\alpha f)(x)\,d^n x = \langle g^*, (\partial_\alpha f) \rangle = \langle \hat{g}(p)^*, (ip)^\alpha \hat{f}(p) \rangle
= (-1)^{|\alpha|}(ip)^\alpha \langle \hat{g}(p)^*, \hat{f}(p) \rangle = (-1)^{|\alpha|}\langle \partial_\alpha g^*, f \rangle
= (-1)^{|\alpha|} \int_{\mathbb{R}^n} (\partial_\alpha g)(x)f(x)\,d^n x,
$$

(11.46)

for $f, g \in H^r(\mathbb{R}^n)$. Furthermore, recall that a function $h \in L^1_{\text{loc}}(\mathbb{R}^n)$ satisfying

$$
\int_{\mathbb{R}^n} \varphi(x)h(x)\,d^n x = (-1)^{|\alpha|} \int_{\mathbb{R}^n} (\partial_\alpha \varphi)(x)f(x)\,d^n x, \quad \forall \varphi \in C_0^\infty(\mathbb{R}^n),
$$

(11.47)

is also called the weak derivative or the derivative in the sense of distributions of $f$ (by Lemma 8.17 such a function is unique if it exists). Hence, choosing $g = \varphi$ in (11.46), we see that $H^r(\mathbb{R}^n)$ is the set of all functions having partial derivatives (in the sense of distributions) up to order $r$, which are in $L^2(\mathbb{R}^n)$.

In this connection the following norm for $H^m(\mathbb{R}^n)$ with $m \in \mathbb{N}_0$ is more common:

$$
\|f\|_{2,m}^2 = \sum_{|\alpha| \leq m} \|\partial_\alpha f\|_2^2.
$$

(11.48)

By $|p^\alpha| \leq |p|^{|\alpha|} \leq (1 + |p|^2)^{m/2}$ it follows that this norm is equivalent to (11.44).

Example. This definition of a weak derivative is tailored for the method of solving linear constant coefficient partial differential equations as outlined in Section 11.2. While Lemma 11.3 only gives us a sufficient condition on $\hat{f}$ for $f$ to be differentiable, the weak derivatives gives us necessary and sufficient conditions. For example, we see that the Poisson equation (11.25) will have a (unique) solution $f \in H^2(\mathbb{R}^n)$ if and if $|p|^{-2}\hat{g} \in L^2(\mathbb{R}^n)$. That this is not true for all $g \in L^2(\mathbb{R}^n)$ is connected with the fact that $|p|^{-2}$ is unbounded and hence no $L^2$ multiplier (cf. Problem 11.11). Consequently the range of $\Delta$ when defined on $H^2(\mathbb{R}^n)$ will not be all of $L^2(\mathbb{R}^n)$ and hence the Poisson equation is not solvable within the class $H^2(\mathbb{R}^n)$ for all $g \in L^2(\mathbb{R}^n)$. Nevertheless, we get a unique weak solution under some conditions. Under which conditions this weak solution is also a classical solution can then be investigated separately.

Note that the situation is even simpler for the Helmholtz equation (11.35) since the corresponding multiplier $(1 + |p|^2)^{-1}$ does map $L^2$ to $L^2$. Hence we
get that the Helmholtz equation has a unique solution \( f \in H^2(\mathbb{R}^n) \) if and only if \( g \in L^2(\mathbb{R}^n) \). Moreover, \( f \in H^{r+2}(\mathbb{R}^n) \) if and only if \( g \in H^{r}(\mathbb{R}^n) \).

Of course a natural question to ask is when the weak derivatives are in fact classical derivatives. To this end observe that the Riemann–Lebesgue lemma implies that \( \partial_\alpha f(x) \in C_0(\mathbb{R}^n) \) provided \( p^\alpha f(p) \in L^1(\mathbb{R}^n) \). Moreover, in this situation the derivatives will exist as classical derivatives:

**Lemma 11.19.** Suppose \( f \in L^1(\mathbb{R}^n) \) or \( f \in L^2(\mathbb{R}^n) \) with \((1 + |p|^k) \hat{f}(p) \in L^1(\mathbb{R}^n)\) for some \( k \in \mathbb{N}_0 \). Then \( f \in C_0^k(\mathbb{R}^n) \), the set of functions with continuous partial derivatives of order \( k \) all of which vanish at \( \infty \). Moreover,
\[
(\partial_\alpha f)\hat{}(p) = (ip)^\alpha \hat{f}(p), \quad |\alpha| \leq k, \quad (11.49)
\]
in this case.

**Proof.** We begin by observing that by Theorem 11.7
\[
f(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ipx} \hat{f}(p) d^n p.
\]
Now the claim follows as in the proof of Lemma 11.4 by differentiating the integral using Problem 7.20. \( \square \)

Now we are able to prove the following embedding theorem.

**Theorem 11.20** (Sobolev embedding). Suppose \( r > k + \frac{n}{2} \) for some \( k \in \mathbb{N}_0 \). Then \( H^r(\mathbb{R}^n) \) is continuously embedded into \( C_\infty^k(\mathbb{R}^n) \) with
\[
\|\partial_\alpha f\|_\infty \leq C_{n,r}\|f\|_{2,r}, \quad |\alpha| \leq k. \quad (11.50)
\]

**Proof.** Abbreviate \( \langle p \rangle = (1 + |p|^2)^{1/2} \). Now use \( |(ip)^\alpha \hat{f}(p)| \leq \langle p \rangle^{\alpha} |\hat{f}(p)| = \langle p \rangle^{-s} \cdot \langle p \rangle^{\alpha+s} |\hat{f}(p)| \). Now \( \langle p \rangle^{-s} \in L^2(\mathbb{R}^n) \) if \( s > \frac{n}{2} \) (use polar coordinates to compute the norm) and \( \langle p \rangle^{\alpha+s} |\hat{f}(p)| \in L^2(\mathbb{R}^n) \) if \( s + |\alpha| \leq r \). Hence \( \langle p \rangle^{\alpha} |\hat{f}(p)| \in L^1(\mathbb{R}^n) \) and the claim follows from the previous lemma. \( \square \)

In fact, we can even do a bit better.

**Lemma 11.21** (Morrey inequality). Suppose \( f \in H^{n/2+\gamma}(\mathbb{R}^n) \) for some \( \gamma \in (0,1) \). Then \( f \in C_0^{\alpha,\gamma}(\mathbb{R}^n) \), the set of functions which are Hölder continuous with exponent \( \gamma \) and vanish at \( \infty \). Moreover,
\[
|f(x) - f(y)| \leq C_{n,\gamma}\|\hat{f}(p)\|_{2,n/2+\gamma}|x-y|^\gamma \quad (11.51)
\]
in this case.

**Proof.** We begin with
\[
f(x + y) - f(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ipx} (e^{ipy} - 1) \hat{f}(p) d^n p
\]
11.3. Sobolev spaces

implying

\[ |f(x + y) - f(x)| \leq \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \frac{|e^{ipy} - 1|}{\langle p \rangle^{n/2 + \gamma}} |\hat{f}(p)| dp, \]

where again \( \langle p \rangle = (1 + |p|^2)^{1/2} \). Hence, after applying Cauchy–Schwarz, it remains to estimate (recall (7.62))

\[
\int_{\mathbb{R}^n} \frac{|e^{ipy} - 1|^2}{\langle p \rangle^{n+2\gamma}} d^n p \leq \frac{S_n}{2(1 - \gamma)} |y|^{2\gamma} + \frac{S_n}{2\gamma} |y|^{2\gamma} = \frac{S_n}{2\gamma(1 - \gamma)} |y|^{2\gamma},
\]

where \( S_n = nV_n \) is the surface area of the unit sphere in \( \mathbb{R}^n \).

Using this lemma we immediately obtain:

**Corollary 11.22.** Suppose \( r \geq k + \gamma + \frac{n}{2} \) for some \( k \in \mathbb{N}_0 \) and \( \gamma \in (0, 1) \). Then \( H^r(\mathbb{R}^n) \) is continuously embedded into \( C^{k,\gamma}_\infty(\mathbb{R}^n) \), the set of functions in \( C^k_\infty(\mathbb{R}^n) \) whose highest derivatives are Hölder continuous of exponent \( \gamma \).

**Example.** The function \( f(x) = \log(|x|) \) is in \( H^1(\mathbb{R}^n) \) for \( n \geq 3 \). In fact, the weak derivatives are given by

\[ \partial_j f(x) = \frac{x_j}{|x|^2}. \quad (11.52) \]

However, observe that \( f \) is not continuous.

The last example shows that in the case \( r < \frac{n}{2} \) functions in \( H^r \) are no longer necessarily continuous. In this case we at least get an embedding into some better \( L^p \) space:

**Theorem 11.23** (Sobolev inequality). Suppose \( 0 < r < \frac{n}{2} \). Then \( H^r(\mathbb{R}^n) \) is continuously embedded into \( L^p(\mathbb{R}^n) \) with \( p = \frac{2n}{n - 2r} \), that is,

\[ \|f\|_p \leq \tilde{C}_{n,r} \|p|^r \hat{f}(p)\|_2 \leq C_{n,r} \|f\|_2. \quad (11.53) \]

**Proof.** We will give a prove based on the Hardy–Littlewood–Sobolev inequality to be proven in Theorem 12.10 below.

It suffices to prove the first inequality. Set \( |p|^r \hat{f}(p) = \hat{g}(p) \in L^2 \). Moreover, choose some sequence \( f_m \in S \to f \in H^r \). Then, by Lemma 11.17 \( f_m = I_r g_m \), and since the Hardy–Littlewood–Sobolev inequality implies that the map \( I_r : L^2 \to L^p \) is continuous, we have \( \|f_m\|_p = \|I_r g_m\|_p \leq \tilde{C} \|g_m\|_2 = \tilde{C} \|\hat{g}_m\|_2 = \tilde{C} \|p|^r \hat{f}_m(p)\|_2 \) and the claim follows after taking limits. □
Problem 11.13. Use dilations \( f(x) \to f(\lambda x), \lambda > 0 \), to show that \( p = \frac{2n}{n-2r} \) is the only index for which the Sobolev inequality \( \|f\|_p \leq \tilde{C}_{n,r}\|p^r \hat{f}(p)\|_2 \) can hold.

Problem 11.14. Suppose \( f \in L^2(\mathbb{R}^n) \) show that \( \varepsilon^{-1}(f(x + e_j \varepsilon) - f(x)) \to g_j(x) \) in \( L^2 \) if and only if \( p_j \hat{f}(p) \in L^2 \), where \( e_j \) is the unit vector into the \( j \)th coordinate direction. Moreover, show \( g_j = \partial_j f \) if \( f \in H^1(\mathbb{R}^n) \).

Problem 11.15. Show that \( u \) is weakly differentiable in the interval \((0, 1)\) if and only if \( u \) is absolutely continuous and \( u' = v \) in this case. (Hint: You will need that \( \int_0^1 u(t)\varphi'(t)dt = 0 \) for all \( \varphi \in C_c^\infty(0, 1) \) if and only if \( u \) is constant.

To see this, choose some \( \varphi_0 \in C_c^\infty(0, 1) \) with \( I(\varphi_0) = \int_0^1 \varphi_0(t)dt = 1 \). Then invoke Lemma 8.17 and use that every \( \varphi \in C_c^\infty(0, 1) \) can be written as \( \varphi(t) = \Phi(t) + I(\varphi)\varphi_0(t) \) with \( \Phi(t) = \int_0^t \varphi(s)ds - I(\varphi)\int_0^t \varphi_0(s)ds \).

11.4. Applications to evolution equations

In this section we want to show how to apply these considerations to evolution equations. As a prototypical example we start with the Cauchy problem for the heat equation

\[ u_t - \Delta u = 0, \quad u(0) = g. \]  

(11.54)

It turns out useful to view \( u(t, x) \) as a function of \( t \) with values in a Banach space \( X \). To this end we let \( I \subseteq \mathbb{R} \) be some interval and denote by \( C(I, X) \) the set of continuous functions from \( I \) to \( X \). Given \( t \in I \) we call \( u : I \to X \) differentiable at \( t \) if the limit

\[ \hat{u}(t) = \lim_{\varepsilon \to 0} \frac{u(t + \varepsilon) - u(t)}{\varepsilon} \]  

(11.55)

exists. The set of functions \( u : I \to X \) which are differentiable at all \( t \in I \) and for which \( \hat{u} \in C(I, X) \) is denoted by \( C^1(I, X) \). As usual we set \( C^{k+1}(I, X) = \{ u \in C^1(I, x) | \hat{u} \in C^k(I, X) \} \). Note that if \( U \in \mathfrak{L}(X,Y) \) and \( u \in C^1(I, X) \), then \( Uu \in C^1(I, Y) \) and \( \frac{d}{dt}Uu = U\hat{u} \).

A strongly continuous operator semigroup (also \( C_0 \) semigroup) is a family of operators \( T(t) \in \mathfrak{L}(X), t \geq 0 \), such that

(i) \( T(t)g \in C([0, \infty), X) \) for every \( g \in X \) (strong continuity) and

(ii) \( T(0) = \mathbb{I}, \ T(t + s) = T(t)T(s) \) for every \( t, s \geq 0 \) (semigroup property).

We first note that \( \|T(t)\| \) is uniformly bounded on compact time intervals.

Lemma 11.24. Let \( T(t) \) be a \( C_0 \)-semigroup. Then there are constants \( M, \omega \geq 0 \) such that

\[ \|T(t)\| \leq Me^{\omega t}. \]  

(11.56)
Proof. Since \( \|T(\cdot)f\| \in C([0, 1]) \) for every \( f \in X \) we have \( \sup_{t \in [0,1]} \|T(t)f\| \leq M_f \). Hence by the uniform boundedness principle \( \sup_{t \in [0,1]} \|T(t)\| \leq M \) for some \( M \geq 1 \). Setting \( \omega = \log(M) \) the claim follows by induction using the semigroup property.

Given a strongly continuous semigroup we can define its generator \( A \) as the linear operator

\[
Af = \lim_{t \downarrow 0} \frac{1}{t}(T(t)f - f) \tag{11.57}
\]

where the domain \( \mathcal{D}(A) \) is precisely the set of all \( f \in X \) for which the above limit exists.

Lemma 11.25. Let \( T(t) \) be a \( C_0 \) semigroup with generator \( A \). If \( g \in \mathcal{D}(A) \) then \( T(t)g \in \mathcal{D}(A) \) and \( AT(t)g = T(t)Ag \). Moreover, suppose \( g \in X \) with \( u(t) = T(t)g \in \mathcal{D}(A) \) for \( t > 0 \). Then \( u(t) \in C^1((0, \infty), X) \cap C([0, \infty), X) \) and \( u(t) \) is the unique solution of the abstract Cauchy problem

\[
\dot{u}(t) = Au(t), \quad u(0) = g. \tag{11.58}
\]

This is, for example, the case if \( g \in \mathcal{D}(A) \) in which case we even have \( u(t) \in C^1([0, \infty), X) \).

Proof. Let \( g \in \mathcal{D}(A) \), then

\[
\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} (u(t+\varepsilon) - u(t)) = \lim_{\varepsilon \downarrow 0} T(t)\frac{1}{\varepsilon} (T(\varepsilon)f - f) = T(t)Ag.
\]

This shows the first part. To show that \( u(t) \) is differentiable for \( t > 0 \) it remains to compute

\[
\lim_{\varepsilon \downarrow 0} \frac{1}{-\varepsilon} (u(t-\varepsilon) - u(t)) = \lim_{\varepsilon \downarrow 0} T(t-\varepsilon)\frac{1}{\varepsilon} (T(\varepsilon)g - g) = \lim_{\varepsilon \downarrow 0} T(t-\varepsilon)(Ag + o(1)) = T(t)Ag
\]

since \( \|T(t)\| \) is bounded on compact \( t \) intervals by the previous lemma. Hence \( u(t) \in C^1([0, \infty), X) \) solves (11.58). In the general case \( f = T(t_0)g \in \mathcal{D}(A) \) and \( u(t) = T(t-t_0)f \) solves our differential equation for every \( t > t_0 \). Since \( t_0 > 0 \) is arbitrary it follows that \( u(t) \) solves (11.58) by the first part. To see that it is the only solution, let \( v(t) \) be a solution corresponding to the initial condition \( v(0) = 0 \). For \( 0 \leq s \leq t \) we have

\[
\frac{d}{ds}T(t-s)v(s) = T(t-s)Av(s) - T(t-s)Av(s) = 0
\]

Whence, by Problem 11.17, \( v(t) = T(t-s)v(s) = T(t)v(0) = 0 \). □
After these preparations we are ready to return to our original problem. Let \( g \in L^2(\mathbb{R}^n) \) and let \( u \in C^1((0, \infty), L^2(\mathbb{R}^n)) \) be a solution such that \( u(t) \in H^2(\mathbb{R}^n) \) for \( t > 0 \). Then we can take the Fourier transform to obtain

\[
\hat{u}_t + |p|^2 \hat{u} = 0, \quad \hat{u}(0) = \hat{g}. \tag{11.59}
\]

Next, one verifies (Problem 11.16) that the solution (in the sense defined above) of this differential equation is given by

\[
\hat{u}(t)(p) = \hat{g}(p)e^{-|p|^2 t}. \tag{11.60}
\]

Accordingly, the solution of our original problem is

\[
u(t) = T_H(t)g, \quad T_H(t) = \mathcal{F}^{-1}e^{-|p|^2 t} \mathcal{F}. \tag{11.61}
\]

Note that \( T_H(t) : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n) \) is a bounded linear operator with \( \|T_H(t)\| \leq 1 \) (since \( |e^{-|p|^2 t}| \leq 1 \)). In fact, for \( t > 0 \) we even have \( T_H(t)g \in H^r(\mathbb{R}^n) \) for any \( r \geq 0 \) showing that \( u(t) \) is smooth even for rough initial functions \( g \). In summary,

**Theorem 11.26.** The family \( T_H(t) \) is a \( C_0 \)-semigroup whose generator is \( \Delta, \mathcal{D}(\Delta) = H^2(\mathbb{R}^n) \).

**Proof.** That \( H^2(\mathbb{R}^n) \subseteq \mathcal{D}(A) \) follows from Problem 11.16. Conversely, let \( g \not\in H^2(\mathbb{R}^n) \). Then \( t^{-1}(e^{-|p|^2 t} - 1) \to -|p|^2 \) uniformly on every compact subset \( K \subset \mathbb{R}^n \). Hence \( \int_K |p|^2 |\hat{g}(p)|^2 \, dp = \int_K |Ag(x)|^2 \, dx \) which gives a contradiction as \( K \) increases. \( \square \)

Next we want to derive a more explicit formula for our solution. To this end we assume \( g \in L^1(\mathbb{R}^n) \) and introduce

\[
\Phi_t(x) = \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/4t}, \tag{11.62}
\]

known as the **fundamental solution** of the heat equation, such that

\[
\hat{u}(t) = (2\pi)^{n/2} \hat{g} \Phi_t = (\Phi_t \ast g)^\wedge \tag{11.63}
\]

by Lemma 11.6 and Lemma 11.12. Finally, by injectivity of the Fourier transform (Theorem 11.7) we conclude

\[
u(t) = \Phi_t \ast g. \tag{11.64}
\]

Moreover, one can check directly that (11.64) defines a solution for arbitrary \( g \in L^p(\mathbb{R}^n) \).

**Theorem 11.27.** Suppose \( g \in L^p(\mathbb{R}^n), 1 \leq p \leq \infty \). Then (11.64) defines a solution for the heat equation which satisfies \( u \in C^\infty((0, \infty) \times \mathbb{R}^n) \). The solutions has the following properties:

(i) If \( 1 \leq p < \infty \), then \( \lim_{t \downarrow 0} u(t) = g \) in \( L^p \). If \( p = \infty \) this holds for \( g \in C_0(\mathbb{R}^n) \).
(ii) If $p = \infty$, then
\[ \|u(t)\|_{\infty} \leq \|g\|_{\infty}. \] (11.65)

If $g$ is real-valued then so is $u$ and
\[ \inf g \leq u(t) \leq \sup g. \] (11.66)

(iii) (Mass conservation) If $p = 1$, then
\[ \int_{\mathbb{R}^n} u(t, x) d^n x = \int_{\mathbb{R}^n} g(x) d^n x \] (11.67)
and
\[ \|u(t)\|_{\infty} \leq \frac{1}{(4\pi t)^{n/2}} \|g\|_1. \] (11.68)

**Proof.** That $u \in C^\infty$ follows since $\Phi \in C^\infty$ from Problem 7.20. To see the remaining claims we begin by noting (by Problem 7.26)
\[ \int_{\mathbb{R}^n} \Phi_t(x) d^n x = 1. \] (11.69)

Now (i) follows from Lemma 8.15, (ii) is immediate, and (iii) follows from Fubini. \qed

Note that using Hölder’s inequality we even have
\[ \|u(t)\|_{\infty} \leq \|\Phi_t\|_q \|g\|_p = \frac{1}{q \frac{1}{q} (4\pi t)^{n/2}} \|g\|_p, \quad \frac{1}{p} + \frac{1}{q} = 1. \] (11.70)

Another closely related equation is the **Schrödinger equation**
\[ -i u_t - \Delta u = 0, \quad u(0) = g. \] (11.71)

As before we obtain that the solution for $g \in H^2(\mathbb{R}^n)$ is given by
\[ u(t) = T_S(t)g, \quad T_S(t) = \mathcal{F}^{-1}e^{-|\cdot|^2t} \mathcal{F}. \] (11.72)

Note that $T_S(t) : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is a unitary operator (since $|e^{-|\cdot|^2t}| = 1$):
\[ \|u(t)\|_2 = \|g\|_2. \] (11.73)

However, while we have $T_H(t)g \in H^r(\mathbb{R}^n)$ whenever $g \in H^r(\mathbb{R}^n)$, unlike the heat equation, the Schrödinger equation does only preserve but not improve the regularity of the initial condition.

**Theorem 11.28.** The family $T_S(t)$ is a $C_0$-group whose generator is $i\Delta$, $\mathcal{D}(i\Delta) = H^2(\mathbb{R}^n)$. 
As in the case of the heat equation, we would like to express our solution as a convolution with the initial condition. However, now we run into the problem that \( e^{-\|p\|^2 t} \) is not integrable. To overcome this problem we consider

\[
 f_\varepsilon(p) = e^{-(i(t+\varepsilon))p^2}, \quad \varepsilon > 0.
\]

Then, as before we have

\[
 (f_\varepsilon \hat{g})(x) = \frac{1}{(4\pi(\varepsilon t))^n/2} \int_{\mathbb{R}^n} e^{-\|x-y\|^2/4(\varepsilon t)} g(y) d^n y
\]

and hence

\[
 T_S(t)g(x) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{i|x-y|^2/4t} g(y) d^n y
\]

for \( t \neq 0 \) and \( g \in L^2(\mathbb{R}^n) \cap L^1(\mathbb{R}^n) \). In fact, letting \( \varepsilon \downarrow 0 \) the left-hand side converges to \( T_S(t)g \) in \( L^2 \) and the limit of the right-hand side exists pointwise by dominated convergence and its pointwise limit must thus be equal to its \( L^2 \) limit.

Using this explicit form, we can again draw some further consequences. For example, if \( g \in L^2(\mathbb{R}^n) \cap L^1(\mathbb{R}^n) \), then \( g(t) \in C(\mathbb{R}^n) \) for \( t \neq 0 \) (use dominated convergence and continuity of the exponential) and satisfies

\[
 \|u(t)\|_{\infty} \leq \frac{1}{|4\pi t|^{n/2}} \|g\|_1.
\]

Thus we have again spreading of wave functions in this case.

Finally we turn to the wave equation

\[
 u_{tt} - \Delta u = 0, \quad u(0) = g, \quad u_t(0) = f.
\]

This equation will fit into our framework once transform it to a first order system with respect to time:

\[
 u_t = v, \quad v_t = \Delta u, \quad u(0) = g, \quad v(0) = f.
\]

After applying the Fourier transform this system reads

\[
 \hat{u}_t = \hat{v}, \quad \hat{v}_t = -|p|^2 \hat{u}, \quad \hat{u}(0) = \hat{g}, \quad \hat{v}(0) = \hat{f},
\]

and the solution is given by

\[
 \hat{u}(t,p) = \cos(t|p|) \hat{g}(p) + \frac{\sin(t|p|)}{|p|} \hat{f}(p),
\]

\[
 \hat{v}(t,p) = -\sin(t|p|)|p| \hat{g}(p) + \cos(t|p|) \hat{f}(p).
\]

Hence for \((g, f) \in H^2(\mathbb{R}^n) \oplus H^1(\mathbb{R}^n)\) our solution is given by

\[
 \begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = T_W(t) \begin{pmatrix} g \\ f \end{pmatrix}, \quad T_W(t) = \mathcal{F}^{-1} \begin{pmatrix} \cos(t|p|) & \sin(t|p|) \\ -\sin(t|p|)|p| & \cos(t|p|) \end{pmatrix} \mathcal{F}.
\]
Theorem 11.29. The family $T_W(t)$ is a $C_0$-semigroup whose generator is $A = \left( \begin{smallmatrix} 0 & 1 \\ \Delta & 0 \end{smallmatrix} \right)$, $\mathcal{D}(A) = H^2(\mathbb{R}^n) \oplus H^1(\mathbb{R}^n)$.

Note that if we use $w$ defined via $\hat{w}(p) = |p|\hat{v}(p)$ instead of $v$, then
\[
\begin{pmatrix} u(t) \\ w(t) \end{pmatrix} = \hat{T}_W(t) \begin{pmatrix} g \\ h \end{pmatrix}, \quad \hat{T}_W(t) = \mathcal{F}^{-1} \begin{pmatrix} \cos(t|p|) & \sin(t|p|) \\ -\sin(t|p|) & \cos(t|p|) \end{pmatrix} \mathcal{F},
\]
where $h$ is defined via $\hat{h} = |p|\hat{f}$. In this case $\hat{T}_W$ is unitary and thus
\[
||u(t)||^2 + ||w(t)||^2 = ||g||^2 + ||h||^2.
\]

If $n = 1$ we have $\frac{\sin(t|p|)}{|p|} \in L^2(\mathbb{R})$ and hence we can get an expression in terms of convolutions. In fact, since the inverse Fourier transform of $\frac{\sin(t|p|)}{|p|}$ is $\sqrt{2\pi} \chi_{[-1,1]}(p/t)$, we obtain
\[
u(t, x) = \int_{\mathbb{R}} \frac{1}{2} \chi_{[-t,t]}(x-y)f(y)dy = \frac{1}{2} \int_{x-t}^{x+t} f(y)dy
\]
in the case $g = 0$. But the corresponding expression for $f = 0$ is just the time derivative of this expression and thus
\[
u(t, x) = \frac{1}{2} \frac{\partial}{\partial t} \int_{x-t}^{x+t} g(y)dy + \frac{1}{2} \int_{x-t}^{x+t} f(y)dy
\]
\[
= \frac{g(x+t) + g(x-t)}{2} + \frac{1}{2} \int_{x-t}^{x+t} f(y)dy,
\]
which is known as d’Alembert’s formula.

To obtain the corresponding formula in $n = 3$ dimensions we use the following observation
\[
\frac{\partial}{\partial t} \hat{\varphi}_t(p) = \frac{\sin(t|p|)}{|p|}, \quad \hat{\varphi}_t(p) = \frac{1 - \cos(t|p|)}{|p|^2},
\]
where $\hat{\varphi}_t \in L^2(\mathbb{R}^3)$. Hence we can compute its inverse Fourier transform using
\[
\varphi_t(x) = \lim_{R \to \infty} \frac{1}{(2\pi)^{3/2}} \int_{B_R(0)} \hat{\varphi}_t(p)e^{ip \cdot x}d^3p
\]
using spherical coordinates (without loss of generality we can rotate our coordinate system, such that the third coordinate direction is parallel to $x$)
\[
\varphi_t(x) = \lim_{R \to \infty} \frac{1}{(2\pi)^{3/2}} \int_0^R \int_0^\pi \int_0^{2\pi} \frac{1 - \cos(tr)}{r^2}e^{ir|x|\cos(\theta)}r^2 \sin(\theta)d\varphi d\theta dr.
\]
Evalutaing the integrals we obtain

\[
\varphi_t(x) = \lim_{R \to \infty} \frac{1}{\sqrt{2\pi}} \int_0^R (1 - \cos(tr)) \int_{0}^{\pi} e^{i|x|\cos(\theta)} \sin(\theta) d\theta dr
\]

\[
= \lim_{R \to \infty} \sqrt{\frac{2}{\pi}} \int_0^R (1 - \cos(tr)) \frac{\sin(r|x|)}{|x|r} dr
\]

\[
= \lim_{R \to \infty} \frac{1}{\sqrt{2\pi}} \int_0^R \left( \frac{2\sin(r|x|)}{|x|r} + \frac{\sin(r(t - |x|))}{|x|r} - \frac{\sin(r(t + |x|))}{|x|r} \right) dr,
\]

\[
= \lim_{R \to \infty} \frac{1}{\sqrt{2\pi|x|}} (2\text{Si}(R|x|) + \text{Si}(R(t - |x|)) - \text{Si}(R(t + |x|)))
\]

(11.89)

where

\[
\text{Si}(z) = \int_0^z \frac{\sin(x)}{x} dx
\]

(11.90)

is the sine integral. Using \(\text{Si}(-x) = -\text{Si}(x)\) for \(x \in \mathbb{R}\) and (Problem 11.20)

\[
\lim_{x \to \infty} \text{Si}(x) = \frac{\pi}{2}
\]

(11.91)

we finally obtain (since the pointwise limit must equal the \(L^2\) limit)

\[
\varphi_t(x) = \sqrt{\frac{\pi}{2}} \frac{\chi_{[0,t]}(|x|)}{|x|}.
\]

(11.92)

For the wave equation this implies (using Lemma 7.33)

\[
u(t, x) = \frac{1}{4\pi} \frac{\partial}{\partial t} \int_{B_{|x|}(|x|)} \frac{1}{|x - y|} f(y) d^3y
\]

\[
= \frac{1}{4\pi} \frac{\partial}{\partial t} \int_0^{|t|} \int_{S^2} \frac{1}{r} f(x - r\omega)r^2 d\sigma^2(\omega) dr
\]

\[
= \frac{t}{4\pi} \int_{S^2} f(x - t\theta) d\sigma^2(\omega)
\]

(11.93)

and thus finally

\[
u(t, x) = \frac{\partial}{\partial t} \frac{t}{4\pi} \int_{S^2} g(x - t\omega) d\sigma^2(\omega) + \frac{t}{4\pi} \int_{S^2} f(x - t\omega) d\sigma^2(\omega)
\]

(11.94)

which is known as Kirchhoff’s formula.

Finally, to obtain a formula in \(n = 2\) dimensions we use the method of descent: That is we use the fact, that our solution in two dimensions is also a solution in three dimensions which happens to be independent of the third coordinate direction. Hence our solution is given by Kirchhoff’s formula and we can simplify the integral using the fact that \(f\) does not depend on \(x_3\).
Using spherical coordinates we obtain
\[
\frac{t}{4\pi} \int_{S^2} f(x - t\omega) d\sigma^2(\omega) =
\]
\[
= \frac{t}{2\pi} \int_0^{\pi/2} \int_0^{2\pi} f(x_1 - t\cos(\varphi)\sin(\theta), x_2 - t\sin(\theta)) \sin(\theta) d\theta d\varphi
\]
\[
= \frac{t}{2\pi} \int_0^1 \int_0^{2\pi} f(x_1 - t\rho\cos(\varphi), x_2 - t\rho\sin(\varphi)) \rho d\rho d\varphi
\]
\[
= \frac{t}{2\pi} \int_{B_1(0)} \frac{f(x - ty)}{\sqrt{1 - |y|^2}} d^2 y
\]
which gives Poissons formula
\[
u(t, x) = \frac{\partial}{\partial t} \frac{t}{2\pi} \int_{B_1(0)} \frac{g(x - ty)}{\sqrt{1 - |y|^2}} d^2 y + \frac{t}{2\pi} \int_{B_1(0)} \frac{f(x - ty)}{\sqrt{1 - |y|^2}} d^2 y. \tag{11.95}
\]

**Problem 11.16.** Show that \(u(t)\) defined in (11.60) is in \(C^1((0, \infty), L^2(\mathbb{R}^n))\) and solves (11.59). (Hint: \(|e^{-t|p|^2} - 1| \leq t|p|^2\) for \(t \geq 0\).)

**Problem 11.17.** Suppose \(u(t) \in C^1(I, X)\). Show that for \(s, t \in I\)
\[
\|u(t) - u(s)\| \leq M|t - s|, \quad M = \sup_{\tau \in [s,t]} \|\frac{d\nu}{dt}(\tau)\|.
\]
(Hint: Consider \(d(\tau) = \|u(\tau) - u(s)\| - \tilde{M}(\tau - s)\) for \(\tau \in [s, t]\). Suppose \(\tau_0\) is the largest \(\tau\) for which the claim holds with \(\tilde{M} > M\) and find a contradiction if \(\tau_0 < t\).)

**Problem 11.18.** Solve the transport equation
\[
u_t + \nu_x = 0, \quad \nu(0) = g,
\]
using the Fourier transform.

**Problem 11.19.** Suppose \(A \in \mathfrak{L}(X)\). Show that
\[
T(t) = \exp(tA) = \sum_{j=0}^{\infty} \frac{t^j}{j!} A^j
\]
defines a \(C_0\) semigroup. Show that it is fact uniformly continuous: \(T(t) \in C([0, \infty), \mathfrak{L}(X))\).

**Problem 11.20.** Show the Dirichlet integral
\[
\lim_{R \to \infty} \int_0^R \frac{\sin(x)}{x} dx = \frac{\pi}{2}.
\]
Show also that the sine integral is bounded
\[
|\text{Si}(x)| \leq \min(x, \pi(1 + \frac{1}{2ex})), \quad x > 0.
\]
11. The Fourier transform

(Hint: Write $\text{Si}(R) = \int_0^R \int_0^\infty \sin(x)e^{-xt}dt\,dx$ and use Fubini.)

11.5. Tempered distributions

In many situations, in particular when dealing with partial differential equations, it turns out convenient to look at generalized functions, also known as distributions.

To begin with we take a closer look at the Schwartz space $\mathcal{S}(\mathbb{R}^m)$ defined in (11.10) which already turned out to be a convenient class for the Fourier transform. For our purpose it will be crucial to have a notion of convergence in $\mathcal{S}(\mathbb{R}^m)$ and the natural choice is the topology generated by the family of seminorms

$$q_n(f) = \sum_{|\alpha|,|\beta| \leq n} \| x^\alpha (\partial_\beta f)(x) \|_\infty,$$

where the sum runs over all multi indices $\alpha, \beta \in \mathbb{N}^m_0$ of order less than $n$. Unfortunately these seminorms cannot be replaced by a single norm (and hence we do not have a Banach space) but there is at least a metric

$$d(f, g) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{q_n(f - g)}{1 + q_n(f - g)}$$

(11.97)

and $\mathcal{S}(\mathbb{R}^m)$ is complete with respect to this metric and hence a Fréchet space:

**Lemma 11.30.** The Schwartz space $\mathcal{S}(\mathbb{R}^m)$ together with the family of seminorms $\{q_n\}_{n \in \mathbb{N}_0}$ is a Fréchet space.

**Proof.** It suffices to show completeness. Since a Cauchy sequence $f_n$ is in particular a Cauchy sequence in $C^\infty(\mathbb{R}^m)$ there is a limit $f \in C^\infty(\mathbb{R}^m)$ such that all derivatives converge uniformly. Moreover, since Cauchy sequences are bounded $\| x^\alpha (\partial_\beta f_n)(x) \|_\infty \leq C_{\alpha, \beta}$ we obtain $\| x^\alpha (\partial_\beta f)(x) \|_\infty \leq C_{\alpha, \beta}$ and thus $f \in \mathcal{S}(\mathbb{R}^m)$. \hfill \Box

We refer to Section 4.7 for further background on Fréchet space. However, for our present purpose it is sufficient to observe that $f_n \to f$ if and only if $q_k(f_n - f) \to 0$ for every $k \in \mathbb{N}_0$. Moreover, (cf. Corollary 4.42) a linear map $A : \mathcal{S}(\mathbb{R}^m) \to \mathcal{S}(\mathbb{R}^m)$ is continuous if and only if for every $j \in \mathbb{N}_0$ there is some $k \in \mathbb{N}_0$ and a corresponding constant $C_k$ such that $q_j(Af) \leq C_k q_k(f)$ and a linear functional $\ell : \mathcal{S}(\mathbb{R}^m) \to \mathbb{C}$ is continuous if and only if there is some $k \in \mathbb{N}_0$ and a corresponding constant $C_k$ such that $|\ell(f)| \leq C_k q_k(f)$.

Now the set of all continuous linear functionals, that is the dual space $\mathcal{S}^*(\mathbb{R}^m)$, is known as the space of **tempered distributions**. To understand
why this generalizes the concept of a function we begin by observing that any locally integrable function which does not grow too fast gives rise to a distribution.

**Example.** Let \( g \) be a locally integrable function of at most polynomial growth, that is, there is some \( k \in \mathbb{N}_0 \) such that \( C_k = \int_{\mathbb{R}^m} |g(x)|(1 + |x|)^{-k}d^m x < \infty \). Then
\[
T_g(f) = \int_{\mathbb{R}^m} g(x)f(x)d^m x
\]
is a distribution. To see that \( T_g \) is continuous observe that \(|T_g(f)| \leq C_k q_k(f)\). Moreover, note that by Lemma 8.17 the distribution \( T_g \) and the function \( g \) uniquely determine each other.  

Of course the next question is if there are distributions which are not functions.

**Example.** Let \( x_0 \in \mathbb{R}^m \) then
\[
\delta_{x_0}(f) = f(x_0)
\]
is a distribution, the **Dirac delta distribution** centered at \( x_0 \). Continuity follows from \(|\delta_{x_0}(f)| \leq q_0(f)\) Formally \( \delta_{x_0} \) can be written as \( T_{\delta_{x_0}} \) as in the previous example where \( \delta_{x_0}(x) = 0 \) for \( x \neq x_0 \) and \( \delta_{x_0}(x) = \infty \) such that \( \int_{\mathbb{R}^m} \delta_{x_0}(x)f(x)d^m x = f(x_0) \). This is of course nonsense as one can easily see that \( \delta_{x_0} \) cannot be expressed as \( T_g \) with at locally integrable function of at most polynomial growth (show this). However, giving a precise mathematical meaning was the very purpose distribution theory was invented for.

**Example.** Of course this example can be easily generalized: Let \( \mu \) be a Borel measure on \( \mathbb{R}^m \) such that \( C_k = \int_{\mathbb{R}^m}(1 + |x|)^{-k}d\mu(x) < \infty \) for some \( k \), then
\[
T_{\mu}(f) = \int_{\mathbb{R}^m} f(x)d\mu(x)
\]
is a distribution since \(|T_{\mu}(f)| \leq C_k q_k(f)\).

**Example.** Another interesting distribution in \( S^*(\mathbb{R}) \) is given by
\[
(p.v. \frac{1}{x})(f) = \lim_{\varepsilon \downarrow 0} \int_{|x| > \varepsilon} \frac{f(x)}{x} dx.
\]
To see that this is a distribution note that by the mean value theorem
\[
|\left(p.v. \frac{1}{x}\right)(f)| = \int_{\varepsilon < |x| < 1} \left| \frac{f(x) - f(0)}{x} \right| dx + \int_{1 < |x|} \frac{|x f(x)|}{x^2} dx
\leq 2 \sup_{|x| \leq 1} |f'(x)| + 2 \sup_{|x| \geq 1} |x f(x)|.
\]
This shows $|(p.v. \frac{1}{x})(f)| \leq 2q_1(f)$. 

Of course to fill distribution theory with life we need to extend the classical operations for functions to distributions. First of all, addition and multiplication by scalars comes for free, but we can easily do more. The general principle is always the same: For any continuous linear operator $A : \mathcal{S}(\mathbb{R}^m) \to \mathcal{S}(\mathbb{R}^m)$ there is a corresponding adjoint operator $A' : \mathcal{S}^*(\mathbb{R}^m) \to \mathcal{S}^*(\mathbb{R}^m)$ which extends the effect on functions (regarded as distributions of type $T_g$) to all distributions. We start with a simple example illustrating this procedure.

Let $h \in \mathcal{S}(\mathbb{R}^m)$, then the map $A : \mathcal{S}(\mathbb{R}^m) \to \mathcal{S}(\mathbb{R}^m), f \mapsto h \cdot f$ is continuous. In fact, continuity follows from the Leibniz rule
\[
\partial_\alpha (h \cdot f) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} (\partial_\beta h)(\partial_{\alpha-\beta} f),
\]
where $\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha-\beta)!}$, $\alpha! = \prod_{j=1}^{m} (\alpha_j!)$, and $\beta \leq \alpha$ means $\beta_j \leq \alpha_j$ for $1 \leq j \leq m$. In particular, $q_j(h \cdot f) \leq C_j q_j(h) q_j(f)$ which shows that $A$ is continuous and hence the adjoint is well defined via
\[
(A'T)(f) = T(Af). \tag{11.98}
\]

Now what is the effect on distributions? For a distribution $T_g$ given by an integrable function as above we clearly have
\[
(A'T_g)(f) = T_g(hf) = \int_{\mathbb{R}^m} g(x)(h(x)f(x))d^mx
= \int_{\mathbb{R}^m} (g(x)h(x))f(x)d^mx = T_{gh}(f). \tag{11.99}
\]

So the effect of $A'$ on functions is multiplication by $h$ and hence $A'$ generalizes this operation to arbitrary distributions. We will write $A'T = h \cdot T$ for notational simplicity. Note that since $f$ can even compensate a polynomial growth, $h$ could even be a smooth functions all whose derivatives grow at most polynomially (e.g. a polynomial):
\[
C^\infty_{pg}(\mathbb{R}^m) = \{ h \in C^\infty(\mathbb{R}^m) \forall \alpha \in \mathbb{N}_0^m \exists C, n : |\partial_\alpha h(x)| \leq C(1 + |x|^n) \}. \tag{11.100}
\]

In summary we can define
\[
(h \cdot T)(f) = T(h \cdot f), \quad h \in C^\infty_{pg}(\mathbb{R}^m). \tag{11.101}
\]

**Example.** Let $h$ be as above and $\delta_{x_0}(f) = f(x_0)$. Then
\[
h \cdot \delta_{x_0}(f) = \delta_{x_0}(h \cdot f) = h(x_0)f(x_0)
\]
and hence $h \cdot \delta_{x_0} = h(x_0)\delta_{x_0}$. 

\[\diamondsuit\]
Moreover, since Schwartz functions have derivatives of all orders, the same is true for tempered distributions! To this end let $\alpha$ be a multi-index and consider $D_\alpha : \mathcal{S}(\mathbb{R}^m) \to \mathcal{S}(\mathbb{R}^m)$, $f \mapsto (-1)^{|\alpha|}\partial_\alpha f$ (the reason for the exit $(-1)^{|\alpha|}$ will become clear in a moment) which is continuous since $q_n(D_\alpha f) \leq q_{n+|\alpha|}(f)$. Again we let $D'_\alpha$ be the corresponding adjoint operator and compute its effect on distributions given by functions:

$$
(D'_\alpha T_g)(f) = T\hat{g}((-1)^{|\alpha|}\partial_\alpha f) = (-1)^{|\alpha|}\int_{\mathbb{R}^m} \hat{g}(x)\partial_\alpha f(x)d^m x = \int_{\mathbb{R}^m} \hat{g}(x)f(x)d^m x = T\hat{\partial_\alpha g}(f),
$$

(11.102)

where we have used integration by parts in the last step. Hence for every multi-index $\alpha$ we define

$$
(\partial_\alpha T)(f) = (-1)^{|\alpha|}T(\partial_\alpha f).
$$

(11.103)

**Example.** Let $\alpha$ be a multi-index and $\delta_{x_0}(f) = f(x_0)$. Then

$$
\partial_\alpha \delta_{x_0}(f) = (-1)^{|\alpha|}\delta_{x_0}(\partial_\alpha f) = (-1)^{|\alpha|}(\partial_\alpha f)(x_0).
$$

Finally we use the same approach for the Fourier transform $\mathcal{F} : \mathcal{S}(\mathbb{R}^m) \to \mathcal{S}(\mathbb{R}^m)$, which is also continuous since $q_n(f) \leq C_n q_{n+1}(f)$ by Lemma 11.4. Since Fubini implies

$$
\int_{\mathbb{R}^m} g(x)\hat{f}(x)d^m x = \int_{\mathbb{R}^m} \hat{g}(x)f(x)d^m x
$$

(11.104)

for $g \in L^1(\mathbb{R}^m)$ and $f \in \mathcal{S}(\mathbb{R}^m)$ we define the Fourier transform of a distribution to be

$$
(\mathcal{F}T)(f) \equiv \hat{T}(f) = T(\hat{f}).
$$

(11.105)

**Example.** Let us compute the Fourier transform of $\delta_{x_0}(f) = f(x_0)$:

$$
\hat{\delta}_{x_0}(f) = \hat{\delta}_{x_0}(\hat{f}) = \hat{f}(x_0) = \frac{1}{(2\pi)^m/2} \int_{\mathbb{R}^m} e^{-ix_0 x} f(x)d^m x = T\hat{g}(f),
$$

where $g(x) = (2\pi)^{-m/2}e^{-ix_0 x}$. \hfill \diamond
Example. A slightly more involved example is the Fourier transform of \( p.v. \frac{1}{x} \):

\[
((p.v. \frac{1}{x}))^\wedge (f) = \lim_{\varepsilon \downarrow 0} \int_{|x| < \varepsilon} \frac{\hat{f}(x)}{x} \, dx = \lim_{\varepsilon \downarrow 0} \int_{|x| < 1/\varepsilon} \int_{\mathbb{R}} e^{-iyx} \frac{f(y)}{x} \, dy \, dx
\]

\[
= \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}} \int_{|x| < 1/\varepsilon} \frac{e^{-iyx}}{x} \, dx \, f(y) \, dy
\]

\[
= -2i \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}} \text{sign}(y) \int_{|x| < 1/\varepsilon} \frac{\sin(t)}{t} \, dt \, f(y) \, dy
\]

where we have used the sign integral (11.90). Moreover, Problem 11.20 shows that we can use dominated convergence to get

\[
((p.v. \frac{1}{x}))^\wedge (f) = -i\pi \int_{\mathbb{R}} \text{sign}(y) f(y) \, dy,
\]

that is, \((p.v. \frac{1}{x})^\wedge = -i\pi \text{sign}(y)\). ♦

Note that since \( \mathcal{F} : \mathcal{S}(\mathbb{R}^m) \rightarrow \mathcal{S}(\mathbb{R}^m) \) is a homeomorphism, so is its adjoint \( \mathcal{F}^\prime : \mathcal{S}^\ast(\mathbb{R}^m) \rightarrow \mathcal{S}^\ast(\mathbb{R}^m) \). In particular, its inverse is given by

\[
\tilde{T}(f) = T(\hat{f}).
\]  

(11.106)

Moreover, all the operations for \( \mathcal{F} \) carry over to \( \mathcal{F}^\prime \). For example, from Lemma 11.4 we immediately obtain

\[
(\partial_\alpha T)^\wedge = (i\rho)^\alpha \hat{T}, \quad (x^\alpha T)^\wedge = i^{\alpha} \partial_\alpha \hat{T}.
\]  

(11.107)

Similarly one can extend Lemma 11.2 to distributions.

Next we turn to convolutions. Since (Fubini)

\[
\int_{\mathbb{R}^m} (h \ast g)(x) f(x) \, d^m x = \int_{\mathbb{R}^m} g(x) (\tilde{h} \ast f)(x) \, d^m x, \quad \tilde{h}(x) = h(-x),
\]  

(11.108)

for integrable functions \( f, g, h \) we define

\[
(h \ast T)(f) = T(\tilde{h} \ast f), \quad h \in \mathcal{S}(\mathbb{R}^n),
\]  

(11.109)

which is well defined by Corollary 11.13. Moreover, Corollary 11.13 immediately implies

\[
(h \ast T)^\wedge = (2\pi)^{n/2} \hat{\tilde{h}} \hat{T}, \quad (hT)^\wedge = (2\pi)^{-n/2} \hat{\tilde{h}} \ast \hat{T}, \quad h \in \mathcal{S}(\mathbb{R}^n).
\]  

(11.110)

Example. Note that the Dirac delta distribution acts like an identity for convolutions since

\[
(h * \delta_0)(f) = \delta_0(\tilde{h} * f) = (\tilde{h} * f)(0) = \int_{\mathbb{R}^m} h(y) f(y) = T_h(f).
\]
In the last example the convolution is associated with a function. This turns out always to be the case.

**Theorem 11.31.** For every \( T \in \mathcal{S}^*(\mathbb{R}^m) \) and \( h \in \mathcal{S}(\mathbb{R}^m) \) we have that \( h * T \) is associated with the function

\[
h * T = T_g, \quad g(x) = T(h(\cdot - x)) \in C^\infty_p(\mathbb{R}^m).
\] (11.111)

**Proof.** By definition \( (h * T)(f) = T(\tilde{h} * f) \) and since \( (\tilde{h} * f)(x) = \int h(y - x) f(y) dy \) the distribution \( T \) acts on \( h(y - \cdot) \) and we should be able to pull out the integral by linearity. To make this idea work let us replace the integral by a Riemann sum

\[
(\tilde{h} * f)(x) = \lim_{n \to \infty} \sum_{j=1}^{n^2} h(y_j^n - x) f(y_j^n) |Q_j^n|,
\]

where \( Q_j^n \) is a partition of \([-\frac{n}{2}, \frac{n}{2}]^m \) into \( n^{2m} \) cubes of side length \( \frac{1}{n} \) and \( y_j^n \) is the midpoint of \( Q_j^n \). Then, if this Riemann sum converges to \( h * f \) in \( \mathcal{S}(\mathbb{R}^m) \), we have

\[
(h * T)(f) = \lim_{n \to \infty} \sum_{j=1}^{n^2} g(y_j^n) f(y_j^n) |Q_j^n|
\]

and of course we expect this last limit to converge to the corresponding integral. To be able to see this we need some properties of \( g \). Since

\[
|h(z - x) - h(z - y)| \leq q_1|h||x - y|
\]

by the mean value theorem and similarly

\[
q_n(h(z - x) - h(z - y)) \leq C_n q_{n+1}(h)|x - y|
\]

we see that \( x \mapsto h(\cdot - x) \) is continuous in \( \mathcal{S}(\mathbb{R}^m) \). Consequently \( g \) is continuous. Similarly, if \( x = x_0 + \varepsilon e_j \) with \( e_j \) the unit vector in the \( j \)’th coordinate direction,

\[
q_n \left( \frac{1}{\varepsilon} (h(\cdot - x) - h(\cdot - x_0)) - \partial_j h(\cdot - x_0) \right) \leq C_n q_{n+2}(h) \varepsilon
\]

which shows \( \partial_j g(x) = T((\partial_j h)(\cdot - \cdot)) \). Applying this formula iteratively gives

\[
\partial_\alpha g(x) = T((\partial_\alpha h)(\cdot - \cdot)) \tag{11.112}
\]

and hence \( g \in C^\infty(\mathbb{R}^m) \). Furthermore, \( g \) has at most polynomial growth since \( |T(f)| \leq C q_n(f) \) implies

\[
|g(x)| = |T(h(\cdot - x))| \leq C q_n(h(\cdot - x)) \leq \tilde{C}(1 + |x|^m) q(h).
\]

Combining this estimate with (11.112) even gives \( g \in C^\infty_p(\mathbb{R}^m) \).
In particular, since $g \cdot f \in \mathcal{S}(\mathbb{R}^m)$ the corresponding Riemann sum converges and we have $h * T = T_g$.

It remains to show that our first Riemann sum for the convolution converges in $\mathcal{S}(\mathbb{R}^m)$. It suffices to show
\[
\sup_x |x|^N \left| \sum_{j=1}^{2^m} h(y_j^n - x) f(y_j^n) \right| Q_j^n - \int_{\mathbb{R}^m} h(y - x) f(y) d^m y \to 0
\]
since derivatives are automatically covered by replacing $h$ with the corresponding derivative. The above expressions splits into two terms. The first one is
\[
\sup_x |x|^N \int_{|y| > n/2} h(y - x) f(y) d^m y \leq C q_N(h) \int_{|y| > n/2} (1 + |y|^N) |f(y)| d^m y \to 0.
\]
The second one is
\[
\sup_x |x|^N \left| \sum_{j=1}^{2^m} \int_{Q_j^n} (h(y_j^n - x) f(y_j^n) - h(y - x) f(y)) d^m y \right|
\]
and the integrand can be estimated by
\[
|x|^N |h(y_j^n - x) f(y_j^n) - h(y - x) f(y)|
\]
\[
\leq |x|^N |h(y_j^n - x) - h(y - x)| |f(y_j^n)| + |x|^N |h(y - x)| |f(y_j^n) - f(y)|
\]
\[
\leq (q_{N+1}(h)(1 + |y_j^n|^N) |f(y_j^n)| + q_N(h)(1 + |y|^N) |\partial f(y_j^n)|) |y - y_j^n|
\]
and the claim follows since $|f(y)| + |\partial f(y)| \leq C(1 + |y|)^{-N-m-1}$.

As a consequence we get that distributions can be approximated by functions.

**Theorem 11.32.** Let $\phi_\varepsilon$ be the standard mollifier. Then $\phi_\varepsilon * T \to T$ in $\mathcal{S}^*(\mathbb{R}^m)$.

**Proof.** We need to show $\phi_\varepsilon * T(f) = T(\phi_\varepsilon * f)$ for any $f \in \mathcal{S}(\mathbb{R}^m)$. This follows from continuity since $\phi_\varepsilon * f \to f$ in $\mathcal{S}(\mathbb{R}^m)$ as can be easily seen (the derivatives follow from Lemma 8.14 (ii)).

As a consequence we get that distributions can be approximated by functions.

**Theorem 11.32.** Let $\phi_\varepsilon$ be the standard mollifier. Then $\phi_\varepsilon * T \to T$ in $\mathcal{S}^*(\mathbb{R}^m)$.

**Proof.** We need to show $\phi_\varepsilon * T(f) = T(\phi_\varepsilon * f)$ for any $f \in \mathcal{S}(\mathbb{R}^m)$. This follows from continuity since $\phi_\varepsilon * f \to f$ in $\mathcal{S}(\mathbb{R}^m)$ as can be easily seen (the derivatives follow from Lemma 8.14 (ii)).

As a consequence we get that distributions can be approximated by functions.

**Theorem 11.32.** Let $\phi_\varepsilon$ be the standard mollifier. Then $\phi_\varepsilon * T \to T$ in $\mathcal{S}^*(\mathbb{R}^m)$.

**Proof.** We need to show $\phi_\varepsilon * T(f) = T(\phi_\varepsilon * f)$ for any $f \in \mathcal{S}(\mathbb{R}^m)$. This follows from continuity since $\phi_\varepsilon * f \to f$ in $\mathcal{S}(\mathbb{R}^m)$ as can be easily seen (the derivatives follow from Lemma 8.14 (ii)).

As a consequence we get that distributions can be approximated by functions.

**Theorem 11.32.** Let $\phi_\varepsilon$ be the standard mollifier. Then $\phi_\varepsilon * T \to T$ in $\mathcal{S}^*(\mathbb{R}^m)$.

**Proof.** We need to show $\phi_\varepsilon * T(f) = T(\phi_\varepsilon * f)$ for any $f \in \mathcal{S}(\mathbb{R}^m)$. This follows from continuity since $\phi_\varepsilon * f \to f$ in $\mathcal{S}(\mathbb{R}^m)$ as can be easily seen (the derivatives follow from Lemma 8.14 (ii)).
11.5. Tempered distributions

An example of a distribution supported at 0 is the Dirac delta distribution \( \delta_0 \) as well as all of its derivatives. It turns out that these are in fact the only examples.

**Lemma 11.33.** Suppose \( T \) is a distribution supported at \( x_0 \). Then

\[
T = \sum_{|\alpha| \leq n} c_\alpha \partial_\alpha \delta_{x_0}.
\]  

(11.114)

**Proof.** For simplicity of notation we suppose \( x_0 = 0 \). First of all there is some \( n \) such that \( |T(f)| \leq C q_n(f) \). Write

\[
T = \sum_{|\alpha| \leq n} c_\alpha \partial_\alpha \delta_0 + \tilde{T},
\]

where \( c_\alpha = \frac{T(x^\alpha)}{\alpha!} \). Then \( \tilde{T} \) vanishes on every polynomial of degree at most \( n \), has support at 0, and still satisfies \( |\tilde{T}(f)| \leq C q_n(f) \). Now let \( \phi(x) = 1 \) in a neighborhood of 0.

Then \( \tilde{T}(f) = \tilde{T}(\phi \cdot g) \), where \( g(x) = f(x) - \sum_{|\alpha| \leq n} \frac{\hat{f}(\alpha)(0)}{\alpha!} x^\alpha \). Since \( |\partial_\beta g(x)| \leq C \beta \epsilon^{n+1-|\beta|} \) for \( x \in B_\epsilon(0) \) Leibniz’ rule implies \( q_n(\phi \cdot g) \leq C \epsilon \). Hence \( |\tilde{T}(f)| = |\tilde{T}(\phi \cdot g)| \leq C q_n(\phi \cdot g) \leq C \epsilon \) and since \( \epsilon > 0 \) is arbitrary we have \( |\tilde{T}(f)| = 0 \), that is, \( \tilde{T} = 0 \). \( \square \)

**Example.** Let us try to solve the Poisson equation in the sense of distributions. We begin with solving

\[-\Delta T = \delta_0.\]

Taking the Fourier transform we obtain

\[|p|^2 \hat{T} = (2\pi)^{-m/2}\]

and since \( |p|^{-1} \) is a bounded locally integrable function in \( \mathbb{R}^m \) for \( m \geq 2 \) the above equation will be solved by

\[\Phi = (2\pi)^{-m/2} \left( \frac{1}{|p|^2} \right)^\vee.\]

Explicitly, \( \Phi \) must be determined from

\[\Phi(f) = (2\pi)^{-m/2} \int_{\mathbb{R}^m} \frac{\hat{f}(p)}{|p|^2} d^m p = \int_{\mathbb{R}^m} \frac{\hat{f}(p)}{|p|^2} d^m p\]

and evaluating Lemma 11.17 at \( x = 0 \) we obtain that \( \Phi = (2\pi)^{-m/2} I_2 \) where \( I_2 \) is the Riesz potential. Note, that \( \Phi \) is not unique since we could add a polynomial of degree two corresponding to a solution of the homogeneous equation \( |p|^2 \hat{T} = 0 \) implies that \( \text{supp}(\hat{T}) = 0 \) and comparing with Lemma 11.33 shows \( \hat{T} = \sum_{|\alpha| \leq 2} c_\alpha \partial_\alpha \delta_0 \) and hence \( T = \sum_{|\alpha| \leq 2} \tilde{c}_\alpha x^\alpha \).
Given $\Phi$ we can now consider $h \ast \Phi$ which solves
\[ -\Delta (h \ast \Phi) = h \ast (-\Delta \Phi) = h \ast \delta_0 = h, \]
where in the last equality we have identified $h$ with $T_h$. This gives the formal calculations with the Dirac delta function found in many physics textbooks a solid mathematical meaning.

Note that while we have been quite successful in generalizing many basic operations to distributions, our approach is limited to linear operations! In particular, it is not possible to define nonlinear operations, for example the product of two distributions within this framework. In fact, there is no associative product of two distributions extending the product of a distribution by a function from above (see Problem 11.23). This is known as Schwartz’ impossibility result. If one is content with preserving the product of functions, Colombeau algebras will do the trick.

**Problem 11.21.** Compute the derivative of $g(x) = \text{sign}(x)$ in $S^*(\mathbb{R})$.

**Problem 11.22.** Let $h \in C^\infty_{pg}(\mathbb{R}^m)$ and $T \in S^*(\mathbb{R}^m)$. Show
\[ \partial_\alpha (h \cdot T) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} (\partial_{\beta} h)(\partial_{\alpha - \beta} T). \]

**Problem 11.23.** Consider the distributions $\delta_0$, $x$, and $p.v. \frac{1}{x}$ in $S^*(\mathbb{R})$. Then
\[ x \cdot \delta_0 = 0, \quad x \cdot p.v. \frac{1}{x} = 1. \]
Hence if there would be product of distributions we would get $0 = (x \cdot \delta_0) \cdot p.v. \frac{1}{x} = \delta_0 \cdot (x \cdot p.v. \frac{1}{x}) = \delta_0$.

**Problem 11.24.** Show that $\text{supp}(T_g) = \text{supp}(g)$ for locally integrable functions.
12.1. Interpolation and the Fourier transform on \( L^p \)

We will fix some measure space \((X, \mu)\) and abbreviate \( L^p = L^p(X, d\mu) \) for notational simplicity. If \( f \in L^{p_0} \cap L^{p_1} \) for some \( p_0 < p_1 \) then it is not hard to see that \( f \in L^p \) for every \( p \in [p_0, p_1] \) and we have (Problem 12.1) the Lyapunov inequality

\[
\|f\|_p \leq \|f\|^{1-\theta}_{p_0} \|f\|^\theta_{p_1},
\]

(12.1)

where \( \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1} \), \( \theta \in (0, 1) \). Note that \( L^{p_0} \cap L^{p_1} \) contains all integrable simple functions which are dense in \( L^p \) for \( 1 \leq p < \infty \) (for \( p = \infty \) this is only true if the measure is finite — cf. Problem 8.10).

This is a first occurrence of an interpolation technique. Next we want to turn to operators. For example, we have defined the Fourier transform as an operator from \( L^1 \to L^\infty \) as well as from \( L^2 \to L^2 \) and the question is if this can be used to extend the Fourier transform to the spaces in between.

Denote by \( L^{p_0} + L^{p_1} \) the space of (equivalence classes) of measurable functions \( f \) which can be written as a sum \( f = f_0 + f_1 \) with \( f_0 \in L^{p_0} \) and \( f_1 \in L^{p_1} \) (clearly such a decomposition is not unique and different decompositions will differ by elements from \( L^{p_0} \cap L^{p_1} \)). Then we have

\[
L^p \subseteq L^{p_0} + L^{p_1}, \quad p_0 < p < p_1,
\]

(12.2)

since we can always decompose a function \( f \in L^p \), \( 1 \leq p < \infty \), as \( f = f\chi_{\{|f(x)| \leq 1\}} + f\chi_{\{|f(x)| > 1\}} \) with \( f\chi_{\{|f(x)| \leq 1\}} \in L^p \cap L^\infty \) and \( f\chi_{\{|f(x)| > 1\}} \in L^1 \cap L^p \). Hence, if we have two operators \( A_0 : L^{p_0} \to L^{q_0} \) and \( A_1 : L^{p_1} \to L^{q_1} \) which coincide on the intersection, \( A_0|_{L^{p_0} \cap L^{p_1}} = A_1|_{L^{p_0} \cap L^{p_1}} \), we can extend
them by virtue of
\[ A : L^{p_0} + L^{p_1} \to L^{q_0} + L^{q_1}, \quad f_0 + f_1 \mapsto A_0 f_0 + A_1 f_1 \]  
(12.3)
(check that A is indeed well-defined, i.e., independent of the decomposition of f into \( f_0 + f_1 \)). In particular, this defines A on \( L^p \) for every \( p \in (p_0, p_1) \) and the question is if A restricted to \( L^p \) will be a bounded operator into some \( L^q \) provided \( A_0 \) and \( A_1 \) are bounded.

To answer this question we begin with a result from complex analysis.

**Theorem 12.1** (Hadamard three-lines theorem). Let \( S \) be the open strip \( \{ z \in \mathbb{C} | 0 < \text{Re}(z) < 1 \} \) and let \( F : \overline{S} \to \mathbb{C} \) be continuous and bounded on \( \overline{S} \) and holomorphic in \( S \). If
\[
|F(z)| \leq \begin{cases} 
M_0, & \text{Re}(z) = 0, \\
M_1, & \text{Re}(z) = 1,
\end{cases} 
\]  
(12.4)
then
\[
|F(z)| \leq M_0^{1-\text{Re}(z)}M_1^{\text{Re}(z)} 
\]  
(12.5)
for every \( z \in \overline{S} \).

**Proof.** Without loss of generality we can assume \( M_0, M_1 > 0 \) and after the transformation \( F(z) \to M_0^{-1}M_1^{-z}F(z) \) even \( M_0 = M_1 = 1 \). Now we consider the auxiliary function
\[
F_n(z) = e^{(z^2-1)/n}F(z)
\]
which still satisfies \( |F_n(z)| \leq 1 \) for \( \text{Re}(z) = 0 \) and \( \text{Re}(z) = 1 \) since \( \text{Re}(z^2 - 1) \leq -\text{Im}(z)^2 \leq 0 \) for \( z \in \overline{S} \). Moreover, by assumption \( |F(z)| \leq M \) implying \( |F_n(z)| \leq 1 \) for \( |\text{Im}(z)| \geq \sqrt{\log(M)n} \). Since we also have \( |F_n(z)| \leq 1 \) for \( |\text{Im}(z)| \leq \sqrt{\log(M)n} \) by the maximum modulus principle we see \( |F_n(z)| \leq 1 \) for all \( z \in \overline{S} \). Finally, letting \( n \to \infty \) the claim follows. \( \square \)

Now we are able to show the **Riesz–Thorin interpolation theorem**

**Theorem 12.2** (Riesz–Thorin). Let \( (X, d\mu) \) and \( (Y, d\nu) \) measure spaces and \( 1 \leq p_0, p_1, q_0, q_1 \leq \infty \). If \( q_0 = q_1 = \infty \) assume additionally that \( \nu \) is \( \sigma \)-finite. If \( A \) is a linear operator on
\[
A : L^{p_0}(X, d\mu) + L^{p_1}(X, d\mu) \to L^{q_0}(Y, d\nu) + L^{q_1}(Y, d\nu)
\]  
(12.6)
satisfying
\[
\|Af\|_{q_0} \leq M_0 \|f\|_{p_0}, \quad \|Af\|_{q_1} \leq M_1 \|f\|_{p_1},
\]  
(12.7)
then \( A \) has continuous restrictions
\[
A_\theta : L^{p_0}(X, d\mu) \to L^{q_0}(Y, d\nu), \quad \frac{1}{p_0} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \frac{1}{q_0} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}
\]  
(12.8)
satisfying \( \|A_\theta\| \leq M_0^{1-\theta}M_1^\theta \) for every \( \theta \in (0, 1) \).
12.1. Interpolation and the Fourier transform on $L^p$ 277

**Proof.** In the case $p_0 = p_1 = \infty$ the claim is immediate from (12.1) and hence we can assume $p_0 < \infty$ in which case the space of integrable simple functions is dense in $L^{p_0}$. We will also temporarily assume $q_0 < \infty$. Then, by Lemma 8.5 it suffices to show
\[
\left| \int (Af)(y)g(y)d\nu(y) \right| \leq M_0^{1-\theta}M_1^\theta,
\]
where $f, g$ are simple functions with $\|f\|_{p_0} = \|g\|_{q_0} = 1$ and $\frac{1}{q_0} + \frac{1}{q_0} = 1$.

Now choose simple functions $f(x) = \sum_j \alpha_j \chi_{A_j}(x)$, $g(x) = \sum_k \beta_k \chi_{B_k}(x)$ with $\|f\|_1 = \|g\|_1 = 1$ and set $f_z(x) = \sum_j |\alpha_j|^{-p_0} \text{sign}(\alpha_j) \chi_{A_j}(x)$ and $g_z(x) = \sum_k |\beta_k|^{-1/q_0} \text{sign}(\beta_k) \chi_{B_k}(y)$ such that $\|f_z\|_{p_0} = \|g_z\|_{q_0} = 1$ for $\theta = \text{Re}(z) \in [0, 1]$. Moreover, note that both functions are entire and thus the function $F(z) = \int (Af_z)(y)g_zd\nu(y)$ satisfies the assumptions of the three-lines theorem. Hence we have the required estimate for integrable simple functions. Now let $f \in L^{p_0}$ and split it according to $f = f_0 + f_1$ with $f_0 \in L^{p_0} \cap L^{p_1}$ and $f_1 \in L^{p_1} \cap L^{p_0}$ and approximate both by integrable simple functions (cf. Problem 8.10).

It remains to consider the case $p_0 < p_1$ and $q_0 = q_1 = \infty$. In this case we can proceed as before using again Lemma 8.5 and a simple function for $g = g_z$.

Note that the proof shows even a bit more

**Corollary 12.3.** Let $A$ be an operator defined on the space of integrable simple functions satisfying (12.7). Then $A$ has continuous extensions $A_\theta$ as in the Riesz–Thorin theorem which will agree on $L^{p_0}(X,d\mu) \cap L^{p_1}(X,d\mu)$.

As a consequence we get two important inequalities:

**Corollary 12.4** (Hausdorff–Young inequality). The Fourier transform extends to a continuous map $\mathcal{F} : L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^n)$, for $1 \leq p \leq 2$, $\frac{1}{p} + \frac{1}{q} = 1$, satisfying
\[
(2\pi)^{-n/(2q)}\|\hat{f}\|_q \leq (2\pi)^{-n/(2p)}\|f\|_p.
\]

We remark that the Fourier transform does not extend to a continuous map $\mathcal{F} : L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^n)$, for $p > 2$ (Problem 12.2). Moreover, its range is dense for $1 < p \leq 2$ but not all of $L^q(\mathbb{R}^n)$ unless $p = q = 2$.

**Corollary 12.5** (Young inequality). Let $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$ with $\frac{1}{p} + \frac{1}{q} \geq 1$. Then $(f(y)g(x - y)$ is integrable with respect to $y$ for a.e. $x$ and the convolution satisfies $f \ast g \in L^r(\mathbb{R}^n)$ with
\[
\|f \ast g\|_r \leq \|f\|_p \|g\|_q,
\]
where \( \frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1 \).

**Proof.** We consider the operator \( A_g f = f * g \) which satisfies \( \|A_g f\|_q \leq \|g\|_q \|f\|_p \) for every \( f \in L^1 \) by Lemma 8.14. Similarly, Hölder’s inequality implies \( \|A_g f\| \leq \|g\| \|f\|_q \) for every \( f \in L^q \), where \( \frac{1}{q} + \frac{1}{q} = 1 \). Hence the Riesz–Thorin theorem implies that \( A_g \) extends to an operator \( A_g : L^p \rightarrow L^q \), where \( \frac{1}{p} = 1 - \frac{1}{q} = 1 - \frac{1}{q} \) and \( \frac{1}{r} = \frac{1}{p} + \frac{1}{q} = 1 - \frac{1}{1-q} - 1 \). To see that \( f(y)g(x - y) \) is integrable a.e. consider \( f_n(x) = \chi_{|x| \leq n}(x) \max(n, |f(x)|) \). Then the convolution \( (f_n * g)(x) \) is finite and converges for every \( x \) by monotone convergence. Moreover, since \( f_n \rightarrow |f| \) in \( L^p \) we have \( f_n * |g| \rightarrow A_g f \) in \( L^r \), which finishes the proof.

Combining the last two corollaries we obtain:

**Corollary 12.6.** Let \( f \in L^p(\mathbb{R}^n) \) and \( g \in L^q(\mathbb{R}^n) \) with \( \frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1 \geq 0 \) and \( 1 \leq r, p, q \leq 2 \). Then \( \|f \ast g\| = (2\pi)^{n/2} \hat{f} \hat{g} \).

**Proof.** By Corollary 11.13 the claim holds for \( f, g \in S(\mathbb{R}^n) \). Now take a sequence of Schwartz functions \( f_m \rightarrow f \) in \( L^p \) and a sequence of Schwartz functions \( g_m \rightarrow g \) in \( L^q \). Then the left-hand side converges in \( L^r \), where \( \frac{1}{r} = 2 - \frac{1}{p} - \frac{1}{q} \), by the Young and Hausdorff-Young inequalities. Similarly, the right-hand side converges in \( L^r \) by the generalized Hölder (Problem 8.5) and Hausdorff-Young inequalities.

**Problem 12.1.** Show (12.1). (Hint: Generalized Hölder inequality from Problem 8.5.)

**Problem 12.2.** Show that the Fourier transform does not extend to a continuous map \( \mathcal{F} : L^p(\mathbb{R}^n) \rightarrow L^q(\mathbb{R}^n) \), for \( p > 2 \). Use the closed graph theorem to conclude that \( \mathcal{F} \) is not onto for \( 1 \leq p \leq 2 \). (Hint for the case \( n = 1 \): Consider \( \phi_z(x) = \exp(-z x^2/2) \) for \( z = \lambda + i \omega \) with \( \lambda > 0 \).)

**Problem 12.3 (Young inequality).** Let \( K(x, y) \) be measurable and suppose

\[
\sup_x \|K(x, \cdot)\|_{L^r(Y, d\nu)} \leq C, \quad \sup_y \|K(\cdot, y)\|_{L^r(X, d\mu)} \leq C.
\]

where \( \frac{1}{r} = 1 - (\frac{1}{p} - \frac{1}{q}) \geq 1 \) for some \( 1 \leq p, q \leq \infty \). Then the operator \( K : L^p(Y, d\nu) \rightarrow L^q(X, d\mu) \), defined by

\[
(Kf)(x) = \int_Y K(x, y)f(y)d\nu(y),
\]

for \( \mu \)-almost every \( x \) is bounded with \( \|K\| \leq C \). (Hint: Show \( \|Kf\|_r \leq C \|f\|_{p', \|Kf\|_r} \leq C \|f\|_1 \) and use interpolation.)
12.2. The Marcinkiewicz interpolation theorem

In this section we are going to look at another interpolation theorem which might be helpful in situations where the Riesz–Thorin interpolation theorem does not apply. In this respect recall, that \( f(x) = \frac{1}{x} \) just fails to be integrable over \( \mathbb{R} \). To include such functions we begin by slightly weakening the \( L^p \) norms. To this end we consider the distribution function

\[
E_f(r) = \mu(\{x \in X | |f(x)| > r\})
\]

of a measurable function \( f : X \to \mathbb{C} \) with respect to \( \mu \). Given, the distribution function we can compute the \( L^p \) norm via

\[
\|f\|_p^p = p \int_0^\infty r^{p-1}E_f(r)dr, \quad 1 \leq p < \infty.
\]

(12.11)

In fact, to see this, start with the elementary integral

\[
|f(x)|^p = p \int_0^{\|f(x)\|} r^{p-1} dr
\]

and use Fubini. In the case \( p = \infty \) we have

\[
\|f\|_\infty = \inf\{r \geq 0 | E_f(r) = 0\}.
\]

(12.12)

Another relationship follows from the observation

\[
\|f\|_p^p = \int_X |f|^p d\mu \geq \int_{|f|>r} r^p d\mu = r^p E_f(r)
\]

(12.14)

which yields Chebyshev’s inequality

\[
E_f(r) \leq r^{-p} \|f\|_p^p.
\]

(12.15)

Motivated by this we define the weak \( L_p \) norm

\[
\|f\|_{p,w} = \sup_{r>0} r E_f(r)^{1/p}, \quad 1 \leq p < \infty,
\]

(12.16)

and the corresponding spaces \( L^{p,w}(X, d\mu) \) consist of all equivalence classes of functions which are equal a.e. for which the above norm is finite. Clearly the distribution function and hence the weak \( L_p \) norm depend only on the equivalence class. Despite its name the weak \( L_p \) norm turns out to be only a quasinorm (Problem 12.4). By construction we have

\[
\|f\|_{p,w} \leq \|f\|_p
\]

(12.17)

and thus \( L^p(X, d\mu) \subseteq L^{p,w}(X, d\mu) \). In the case \( p = \infty \) we set \( \|\cdot\|_{\infty,w} = \|\cdot\|_\infty \).

**Example.** Consider \( f(x) = \frac{1}{x} \) in \( \mathbb{R} \). Then clearly \( f \notin L^1(\mathbb{R}) \) but

\[
E_f(r) = \{|x| > r\} = \{|x| < r^{-1}\} = \frac{2}{r}
\]

shows that \( f \in L^{1,w}(\mathbb{R}) \) with \( \|f\|_{1,w} = 2 \). Slightly more general the function \( f(x) = |x|^{-n/p} \notin L^p(\mathbb{R}^n) \) but \( f \in L^{p,w}(\mathbb{R}^n) \). Hence \( L^{p,w}(\mathbb{R}^n) \) is strictly larger than \( L^p(\mathbb{R}^n) \).
Now we are ready for our interpolation result. We call an operator $T : L^p(X, d\mu) \to L^q(X, d\nu)$ **subadditive** if it satisfies

$$
\|T(f + g)\|_q \leq \|T(f)\|_q + \|T(g)\|_q.
$$

(12.18)

It is said to be of **strong type** $(p, q)$ if

$$
\|T(f)\|_q \leq C_{p,q} \|f\|_p
$$

(12.19)

and of **weak type** $(p, q)$ if

$$
\|T(f)\|_{q,w} \leq C_{p,q,w} \|f\|_p.
$$

(12.20)

By (12.17) strong type $(p, q)$ is indeed stronger than weak type $(p, q)$ and we have $C_{p,q,w} \leq C_{p,q}$.

**Theorem 12.7** (Marcinkiewicz). Let $(X, d\mu)$ and $(Y, d\nu)$ measure spaces and $1 \leq p_0 < p_1 \leq \infty$. Let $T$ be a subadditive operator defined for all $f \in L^p(X, d\mu)$, $p \in [p_0, p_1]$. If $T$ is of weak type $(p_0, p_0)$ and $(p_1, p_1)$ then it is also of strong type $(p, p)$ for every $p_0 < p < p_1$.

**Proof.** We begin by assuming $p_1 < \infty$. Fix $f \in L^p$ as well as some number $s > 0$ and decompose $f = f_0 + f_1$ according to

$$
f_0 = f \chi_{\{|f| > s\}} \in L^{p_0} \cap L^p, \quad f_1 = f \chi_{\{|f| \leq s\}} \in L^p \cap L^{p_1}.
$$

Next we use (12.12),

$$
\|T(f)\|_p^p = p \int_0^\infty r^{p-1} E_T(f)(r) dr = p 2^p \int_0^\infty r^{p-1} E_T(f)(2r) dr
$$

and observe

$$
E_T(f)(2r) \leq E_T(f_0)(r) + E_T(f_1)(r)
$$

since $|T(f)| \leq |T(f_0)| + |T(f_1)|$ implies $|T(f)| > 2r$ only if $|T(f_0)| > r$ or $|T(f_1)| > r$. Now using (12.15) our assumption implies

$$
E_T(f_0)(r) \leq \left( \frac{C_0 \|f_0\|_{p_0}}{r} \right)^{p_0}, \quad E_T(f_1)(r) \leq \left( \frac{C_1 \|f_1\|_{p_1}}{r} \right)^{p_1}
$$

and choosing $s = r$ we obtain

$$
E_T(f)(2r) \leq \frac{C_0^{p_0}}{r^{p_0}} \int_{\{|f| > r\}} |f|^{p_0} d\mu + \frac{C_1^{p_1}}{r^{p_1}} \int_{\{|f| \leq r\}} |f|^{p_1} d\mu.
$$

In summary we have $\|T(f)\|_p^p \leq p 2^p (C_0^{p_0} I_1 + C_1^{p_1} I_2)$ with

$$
I_0 = \int_0^\infty \int_X r^{p-p_0-1} \chi_{\{|f(x)| > r\}} |f(x)|^{p_0} d\mu(x) dr
$$

$$
= \int_X |f(x)|^{p_0} \int_0^\infty r^{p-p_0-1} dr d\mu(x) = \frac{1}{p - p_0} \|f\|_p^p
$$
12.2. The Marcinkiewicz interpolation theorem

and

\[ I_1 = \int_0^\infty \int_X r^{p-p_1-1} \chi_{\{|f(x)| \leq r\}} |f(x)|^{p_1} d\mu(x) \, dr \]

\[ = \int_X |f(x)|^{p_1} \int_{\{|f(x)| \leq r\}} r^{p-p_1-1} dr \, d\mu(x) = \frac{1}{p_1 - p} \|f\|_p^p. \]

This is the desired estimate

\[ \|T(f)\|_p \leq 2\left(\frac{p}{p - p_0} C_0^{p_0} + \frac{p}{p_1 - p} C_1^{p_1}\right)^{1/p} \|f\|_p. \]

The case \( p_1 = \infty \) is similar: Split \( f \in L^{p_0} \) according to

\[ f_0 = f \chi_{\{|f| > s/C_1\}} \in L^{p_0} \cap L^p, \quad f_1 = f \chi_{\{|f| \leq s/C_1\}} \in L^p \cap L^\infty \]

(if \( C_1 = 0 \) there is nothing to prove). Then \( \|T(f_1)\|_\infty \leq s/C_1 \) and hence \( E_T(f_1)(s) = 0 \). Thus

\[ E_T(f)(2r) \leq \frac{C_0^{p_0}}{r^{p_0}} \int_{\{|f| > r/C_1\}} |f|^{p_0} d\mu \]

and we can proceed as before to obtain

\[ \|T(f)\|_p \leq 2\left(\frac{p}{p - p_0}/p\right) \frac{1}{C_0^{p_0/p}} C_1^{1-p_0/p} \|f\|_p, \]

which is again the desired estimate.

As with the Riesz–Thorin theorem there is also a version for operators which are of weak type \((p_0, q_0)\) and \((p_1, q_1)\) but the proof is slightly more involved and the above diagonal version will be sufficient for our purpose.

As a first application we will use it to investigate the **Hardy–Littlewood maximal function** defined for any locally integrable function in \( \mathbb{R}^n \) via

\[ M(f)(x) = \sup_{r > 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)| dy. \]  

(12.21)

By the dominated convergence theorem, the integral is continuous with respect to \( x \) and consequently (Problem 7.10) \( M(f) \) is lower semicontinuous (and hence measurable). Moreover, its value is unchanged if we change \( f \) on sets of measure zero, so \( M \) is well defined for functions in \( L^p(\mathbb{R}^n) \). However, it is unclear if \( M(f)(x) \) is finite a.e. at this point. If \( f \) is bounded we of course have the trivial estimate

\[ \|M(f)\|_\infty \leq \|f\|_\infty. \]  

(12.22)

**Theorem 12.8** (Hardy–Littlewood maximal inequality). The maximal function is of weak type \((1,1), \)

\[ E_{M(f)}(r) \leq \frac{3^n}{r} \|f\|_1, \]  

(12.23)
and of strong type \((p, p)\),
\[
\|\mathcal{M}(f)\|_p \leq 2 \left( \frac{3^n p}{p - 1} \right)^{1/p} \|f\|_p,
\]  
(12.24)
for every \(1 < p \leq \infty\).

**Proof.** The first estimate follows literally as in the proof of Lemma 9.5 and combining this estimate with the trivial one (12.22) the Marcinkiewicz interpolation theorem yields the second. \(\square\)

Using this fact, our next aim is to prove the Hardy–Littlewood–Sobolev inequality. As a preparation we show

**Lemma 12.9.** Let \(\varphi \in L^1(\mathbb{R}^n)\) be a radial, \(\varphi(x) = \phi_0(|x|)\) with \(\phi_0\) positive and non increasing. Then we have the following estimate for convolutions with integrable functions:
\[
| (\varphi * f)(x) | \leq \| \varphi \|_1 \mathcal{M}(f)(x).
\]  
(12.25)

**Proof.** By approximating \(\varphi_0\) with simple functions of the same type, it suffices to prove that case where \(\phi_0 = \sum_{j=1}^{p} \alpha_j \chi_{[0, r_j]}\) with \(\alpha_j > 0\). Then
\[
(\varphi * f)(x) = \sum_j \alpha_j |B_{r_j}(0)| \frac{1}{|B_{r_j}(x)|} \int_{B_{r_j}(x)} f(y) d^n y
\]
and the estimate follows upon taking absolute values and observing \(\| \varphi \|_1 = \sum_j \alpha_j |B_{r_j}(0)|\). \(\square\)

Now we will apply this to the Riesz potential (11.29) of order \(\alpha\):
\[
\mathcal{I}_\alpha f = I_\alpha * f.
\]  
(12.26)

**Theorem 12.10** (Hardy–Littlewood–Sobolev inequality). Let \(0 < \alpha < n\), \(p \in (1, \frac{n}{\alpha})\), and \(q = \frac{pn}{n - p\alpha} \in (\frac{n}{n - \alpha}, \infty)\) (i.e, \(\frac{n}{n} = \frac{1}{p} - \frac{1}{q}\)). Then \(\mathcal{I}_\alpha\) is of strong type \((p, q)\),
\[
\|\mathcal{I}_\alpha f\|_q \leq C_{p, \alpha, n} \|f\|_p.
\]  
(12.27)

**Proof.** We split the Riesz potential into two parts
\[
I_\alpha = I_\alpha^0 + I_\alpha^\infty, \quad I_\alpha^0 = I_\alpha \chi_{(0, \varepsilon)}, \quad I_\alpha^\infty = I_\alpha \chi_{[\varepsilon, \infty)},
\]
where \(\varepsilon > 0\) will be determined later. Note that \(I_\alpha^0(|.|) \in L^1(\mathbb{R}^n)\) and \(I_\alpha^\infty(|.|) \in L^r(\mathbb{R}^n)\) for every \(r \in (\frac{n}{n - \alpha}, \infty)\). In particular, since \(p' = \frac{p}{p - 1} \in (\frac{n}{n - \alpha}, \infty)\), both integrals converge absolutely by the Young inequality (12.10). Next we will estimate both parts individually. Using Lemma 12.9 we obtain
\[
|\mathcal{I}_\alpha^0 f(x)| \leq \int_{|y| < \varepsilon} \frac{d^n y}{|y|^{n - \alpha}} \mathcal{M}(f)(x) = \frac{(n - 1)V_n}{\alpha - 1} \varepsilon^{n - \alpha} \mathcal{M}(f)(x).
\]
On the other hand, using Hölder’s inequality we infer
\[ |\mathcal{I}_{\alpha}^\infty f(x)| \leq \left( \int_{|y| \geq \varepsilon} \frac{d^n y}{|y|^{(n-\alpha)p'}} \right)^{1/p'} \|f\|_p = \left( \frac{(n-1)V_n}{p'(n-\alpha) - n} \right)^{1/p'} \varepsilon^{\alpha - n/p} \|f\|_p. \]

Now we choose \( \varepsilon = \left( \frac{\|f\|_p}{M(f)(x)} \right)^{p/n} \) such that
\[ |\mathcal{I}\alpha f(x)| \leq \tilde{C} \|f\|_p \theta M(f)(x)^{1-\theta}, \quad \theta = \frac{\alpha p}{n} \in (\frac{\alpha}{n}, 1), \]
where \( \tilde{C}/2 \) is the larger of the two constants in the estimates for \( \mathcal{I}_0 f \) and \( \mathcal{I}_\alpha f \). Taking the \( L^q \) norm in the above expression gives
\[ \|\mathcal{I}_\alpha f\|_q \leq \tilde{C} \|f\|_p \theta M(f)\|_q = \tilde{C} \|f\|_p \theta |M(f)|_q^{1-\theta} = \tilde{C} \|f\|_p \theta |M(f)|_q^{1-\theta} \]
and the claim follows from the Hardy–Littlewood maximal inequality. \( \square \)

**Problem 12.4.** Show that \( E_f = 0 \) if and only if \( f = 0 \). Moreover, show \( E_{f+g}(r) \leq E_f(r/2) + E_g(r/2) \) and \( E_{\alpha f}(t) = E_f(r/|\alpha|) \) for \( \alpha \neq 0 \). Conclude that \( L^{p,w}(X, d\mu) \) is a quasinormed space with
\[ \|f + g\|_{p,w} \leq 2(\|f\|_{p,w} + \|g\|_{p,w}), \quad \|\alpha f\|_{p,w} = |\alpha| \|f\|_{p,w}. \]

**Problem 12.5.** Show that the maximal function of an integrable function is finite at every Lebesgue point.

**Problem 12.6.** Let \( \phi \) be a nonnegative decreasing radial function with \( \|\phi\|_1 = 1 \). Set \( \phi_\varepsilon(x) = \varepsilon^{-n} \phi(\frac{x}{\varepsilon}) \). Show that \( (\phi_\varepsilon \ast f)(x) \to f(x) \) at every Lebesgue point. (Hint: Split \( \phi = \phi^\delta + \tilde{\phi}^\delta \) into a part with compact support \( \phi^\delta \) and a rest by setting \( \tilde{\phi}^\delta(x) = \min(\delta, \phi(x)) \). To handle the compact part use Problem 8.13. To control the contribution of the rest use Lemma 12.9.)

**Problem 12.7.** Show \( f(x) = |x|^{-n/p} \in L^{p,w}(\mathbb{R}^n) \). Compute \( \|f\|_{p,w} \).

**Problem 12.8.** For \( f \in L^1(0,1) \) define
\[ T(f)(x) = e^{-i \arg \int_0^1 f(y) dy} f(x). \]
Show that \( T \) is subadditive and norm preserving. Show that \( T \) is not continuous.
Part 3

Nonlinear Functional Analysis
Chapter 13

Analysis in Banach spaces

13.1. Differentiation and integration in Banach spaces

We first review some basic facts from calculus in Banach spaces.

Let $X$ and $Y$ be two Banach spaces and denote by $C(X,Y)$ the set of continuous functions from $X$ to $Y$ and by $\mathcal{L}(X,Y) \subset C(X,Y)$ the subset of (bounded) linear functions. Let $U$ be an open subset of $X$. Then a function $F : U \to Y$ is called differentiable at $x \in U$ if there exists a linear function $dF(x) \in \mathcal{L}(X,Y)$ such that

$$F(x + u) = F(x) + dF(x)u + o(u),\quad (13.1)$$

where $o, O$ are the Landau symbols. The linear map $dF(x)$ is called derivative of $F$ at $x$. If $F$ is differentiable for all $x \in U$ we call $F$ differentiable. In this case we get a map

$$dF : U \to \mathcal{L}(X,Y)$$

$$x \mapsto dF(x) \quad (13.2)$$

If $dF$ is continuous, we call $F$ continuously differentiable and write $F \in C^1(U,Y)$.

Let $Y = \prod_{j=1}^m Y_j$ and let $F : X \to Y$ be given by $F = (F_1, \ldots, F_m)$ with $F_j : X \to Y_j$. Then $F \in C^1(X,Y)$ if and only if $F_j \in C^1(X,Y_j)$, $1 \leq j \leq m$, and in this case $dF = (dF_1, \ldots, dF_m)$. Similarly, if $X = \prod_{i=1}^n X_i$, then one can define the partial derivative $\partial_i F \in \mathcal{L}(X_i,Y)$, which is the derivative of $F$ considered as a function of the $i$-th variable alone (the other variables
being fixed). We have \( dF v = \sum_{i=1}^{n} \partial_i F v_i, \ v = (v_1, \ldots, v_n) \in X, \) and \( F \in C^1(X, Y) \) if and only if all partial derivatives exist and are continuous.

In the case of \( X = \mathbb{R}^m \) and \( Y = \mathbb{R}^n \), the matrix representation of \( dF \) with respect to the canonical basis in \( \mathbb{R}^m \) and \( \mathbb{R}^n \) is given by the partial derivatives \( \partial_i F_j(x) \) and is called Jacobi matrix of \( F \) at \( x \).

We can iterate the procedure of differentiation and write \( F \in C^r(U, Y) \), \( r \geq 1 \), if the \( r \)-th derivative of \( F \) (i.e., the derivative of the \( (r - 1) \)-th derivative of \( F \)), exists and is continuous. Finally, we set \( C^\infty(U, Y) = \bigcap_{r \in \mathbb{N}} C^r(U, Y) \) and, for notational convenience, \( C^0(U, Y) = C(U, Y) \) and \( d^0 F = F \).

It is often necessary to equip \( C^r(U, Y) \) with a norm. A suitable choice is

\[
|F| = \max_{0 \leq j \leq r} \sup_{x \in U} |d^j F(x)|. \tag{13.3}
\]

The set of all \( r \) times continuously differentiable functions for which this norm is finite forms a Banach space which is denoted by \( C^r_b(U, Y) \).

If \( F \) is bijective and \( F^{-1} \) are both of class \( C^r \), \( r \geq 1 \), then \( F \) is called a \textbf{diffeomorphism} of class \( C^r \).

Note that if \( F \in \mathcal{L}(X, Y) \), then \( dF(x) = F \) (independent of \( x \)) and \( d^r F(x) = 0, \ r > 1 \).

For the composition of mappings we note the following result (which is easy to prove).

\textbf{Lemma 13.1} (Chain rule). Let \( F \in C^r(X, Y) \) and \( G \in C^r(Y, Z) \), \( r \geq 1 \). Then \( G \circ F \in C^r(X, Z) \) and

\[
d(G \circ F)(x) = dG(F(x)) \circ dF(x), \quad x \in X. \tag{13.4}
\]

In particular, if \( \lambda \in Y^* \) is a linear functional, then \( d(\lambda \circ F) = d\lambda \circ dF = \lambda \circ dF \). In addition, we have the following mean value theorem.

\textbf{Theorem 13.2} (Mean value). Suppose \( U \subseteq X \) and \( F \in C^1(U, Y) \). If \( U \) is convex, then

\[
|F(x) - F(y)| \leq M|x - y|, \quad M = \max_{0 \leq t \leq 1} |dF((1 - t)x + ty)|. \tag{13.5}
\]

Conversely, (for any open \( U \)) if

\[
|F(x) - F(y)| \leq M|x - y|, \quad x, y \in U, \quad (13.6)
\]

then

\[
\sup_{x \in U} |dF(x)| \leq M. \tag{13.7}
\]
If we take \( n \) copies of the same space, the set of multilinear functions \( F : X^n \to Y \) will be denoted by \( \mathcal{L}^n(X,Y) \). A multilinear function is called symmetric provided its value remains unchanged if any two arguments are switched. With the norm from above it is a Banach space and in fact there is a canonical isometric isomorphism between \( \mathcal{L}^n(X,Y) \) and \( \mathcal{L}(X,\mathcal{L}^{n-1}(X,Y)) \) given by \( F : (x_1,\ldots,x_n) \mapsto F(x_1,\ldots,x_n) \) maps to \( x_1 \mapsto F(x_1,\ldots) \). In addition, note that to each \( F \in \mathcal{L}^n(X,Y) \) we can assign its polar form.
$F \in C(X,Y)$ using $F(x) = F(x, \ldots, x)$, $x \in X$. If $F$ is symmetric it can be reconstructed from its polar form using

$$F(x_1, \ldots, x_n) = \frac{1}{n!} \partial_{t_1} \cdots \partial_{t_n} F \left( \sum_{i=1}^{n} t_i x_i \right) \bigg|_{t_1 = \cdots = t_n = 0}. \quad (13.12)$$

Moreover, the $r$-th derivative of $F \in C^r(X,Y)$ is symmetric since,

$$d^r F(x_1, \ldots, x_r) = \partial_{t_1} \cdots \partial_{t_r} F \left( x + \sum_{i=1}^{r} t_i v_i \right) \bigg|_{t_1 = \cdots = t_r = 0}, \quad (13.13)$$

where the order of the partial derivatives can be shown to be irrelevant.

Now we turn to integration. We will only consider the case of mappings $f : I \to X$ where $I = [a, b] \subset \mathbb{R}$ is a compact interval and $X$ is a Banach space. A function $f : I \to X$ is called simple if the image of $f$ is finite, $f(I) = \{ x_i \}_{i=1}^{n}$, and if each inverse image $f^{-1}(x_i)$, $1 \leq i \leq n$ is a Borel set. The set of simple functions $S(I,X)$ forms a linear space and can be equipped with the sup norm. The corresponding Banach space obtained after completion is called the set of regulated functions $R(I,X)$.

Observe that $C(I,X) \subset R(I,X)$. In fact, consider the functions $f_n = \sum_{i=0}^{n-1} f(t_i) \chi_{[t_i, t_{i+1}]} \in S(I,X)$, where $t_i = a + \frac{i(b-a)}{n}$ and $\chi$ is the characteristic function. Since $f \in C(I,X)$ is uniformly continuous, we infer that $f_n$ converges uniformly to $f$.

For $f \in S(I,X)$ we can define a linear map $\int : S(I,X) \to X$ by

$$\int_a^b f(t) dt = \sum_{i=1}^{n} x_i \mu(f^{-1}(x_i)), \quad (13.14)$$

where $\mu$ denotes the Lebesgue measure on $I$. This map satisfies

$$\left| \int_a^b f(t) dt \right| \leq |f|(b-a). \quad (13.15)$$

and hence it can be extended uniquely to a linear map $\int : R(I,X) \to X$ with the same norm $(b-a)$. We even have

$$\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt. \quad (13.16)$$

In addition, if $\lambda \in X^*$ is a continuous linear functional, then

$$\lambda(\int_a^b f(t) dt) = \int_a^b \lambda(f(t)) dt, \quad f \in R(I,X). \quad (13.17)$$

We will use the usual conventions $\int_a^{t_2} f(s) ds = \int_a^{t_1} \chi(t_1,t_2)(s)f(s) ds$ and $\int_{t_1}^{t_2} f(s) ds = -\int_{t_2}^{t_1} f(s) ds$. 


If $I \subseteq \mathbb{R}$, we have an isomorphism $\mathcal{L}(I, X) \cong X$ and if $F : I \to X$ we will write $\dot{F}(t)$ instead of $dF(t)$ if we regard $dF(t)$ as an element of $X$. In particular, if $f \in C(I, X)$, then $F(t) = \int_a^t f(s) \, ds \in C^1(I, X)$ and $\dot{F}(t) = f(t)$ as can be seen from

$$\int_a^{t+\varepsilon} f(s) \, ds - \int_a^t f(s) \, ds - f(t) \varepsilon = \int_t^{t+\varepsilon} (f(s) - f(t)) \, ds \leq \varepsilon \sup_{s \in [t, t+\varepsilon]} |f(s) - f(t)|.$$  

(13.18)

This even shows that $F(t) = F(a) + \int_a^t (\dot{F}(s)) \, ds$ for any $F \in C^1(I, X)$.

13.2. Contraction principles

A fixed point of a mapping $F : C \subseteq X \to C$ is an element $x \in C$ such that $F(x) = x$. Moreover, $F$ is called a contraction if there is a contraction constant $\theta \in [0, 1)$ such that

$$|F(x) - F(\tilde{x})| \leq \theta |x - \tilde{x}|, \quad x, \tilde{x} \in C.$$  

(13.19)

Note that a contraction is continuous. We also recall the notation $F^n(x) = F(F^{n-1}(x))$, $F^0(x) = x$.

**Theorem 13.4** (Contraction principle). *Let $C$ be a closed subset of a Banach space $X$ and let $F : C \to C$ be a contraction, then $F$ has a unique fixed point $\bar{x} \in C$ such that

$$|F^n(x) - \bar{x}| \leq \frac{\theta^n}{1 - \theta} |F(x) - x|, \quad x \in C.$$  

(13.20)*

**Proof.** If $x = F(x)$ and $\tilde{x} = F(\tilde{x})$, then $|x - \tilde{x}| = |F(x) - F(\tilde{x})| \leq \theta |x - \tilde{x}|$ shows that there can be at most one fixed point.

Concerning existence, fix $x_0 \in C$ and consider the sequence $x_n = F^n(x_0)$. We have

$$|x_{n+1} - x_n| \leq \theta |x_n - x_{n-1}| \leq \cdots \leq \theta^n |x_1 - x_0|$$  

(13.21)

and hence by the triangle inequality (for $n > m$)

$$|x_n - x_m| \leq \sum_{j=m+1}^n |x_j - x_{j-1}| \leq \theta^m \sum_{j=0}^{n-m-1} \theta^j |x_1 - x_0|$$

$$\leq \frac{\theta^m}{1 - \theta} |x_1 - x_0|.$$  

(13.22)

Thus $x_n$ is Cauchy and tends to a limit $\bar{x}$. Moreover,

$$|F(\bar{x}) - \bar{x}| = \lim_{n \to \infty} |x_{n+1} - x_n| = 0$$  

(13.23)

shows that $\bar{x}$ is a fixed point and the estimate (13.20) follows after taking the limit $n \to \infty$ in (13.22).\qed
Next, we want to investigate how fixed points of contractions vary with respect to a parameter. Let \( U \subseteq X, \ V \subseteq Y \) be open and consider \( F : \overline{U} \times V \to U \). The mapping \( F \) is called a uniform contraction if there is a \( \theta \in (0, 1) \) such that
\[
|F(x, y) - F(\tilde{x}, y)| \leq \theta|x - \tilde{x}|, \quad x, \tilde{x} \in \overline{U}, \ y \in V. \tag{13.24}
\]

**Theorem 13.5 (Uniform contraction principle).** Let \( U, \ V \) be open subsets of Banach spaces \( X, \ Y \), respectively. Let \( F : \overline{U} \times V \to U \) be a uniform contraction and denote by \( \varphi(y) \in U \) the unique continuous solution of\( F(\cdot, y) \). If \( F \in C^r(U \times V, U), \ r \geq 0 \), then \( \varphi(\cdot) \in C^r(V, U) \).

**Proof.** Let us first show that \( \varphi(y) \) is continuous. From
\[
|\varphi(y + v) - \varphi(y)| = |F(\varphi(y + v), y + v) - F(\varphi(y), y + v) + F(\varphi(y), y + v) - F(\varphi(y), y)|
\[
\leq \theta|\varphi(y + v) - \varphi(y)| + |F(\varphi(y), y + v) - F(\varphi(y), y)| \tag{13.25}
\]
we infer
\[
|\varphi(y + v) - \varphi(y)| \leq \frac{1}{1 - \theta}|F(\varphi(y), y + v) - F(\varphi(y), y)| \tag{13.26}
\]
and hence \( \varphi(y) \in C(V, U) \). Now let \( r = 1 \) and let us formally differentiate \( \varphi(y) = F(\varphi(y), y) \) with respect to \( y \),
\[
d\varphi(y) = \partial_y F(\varphi(y), y)\, dy + \partial_y F(\varphi(y), y). \tag{13.27}
\]
Considering this as a fixed point equation \( T(x', y) = x' \), where \( T(\cdot, y) : \mathcal{L}(Y, X) \to \mathcal{L}(Y, X), \ x' \to \partial_y F(\varphi(y), y) x' + \partial_y F(\varphi(y), y) \) is a uniform contraction since we have \( |\partial_y F(\varphi(y), y)| \leq \theta \) by Theorem 13.2. Hence we get a unique continuous solution \( \varphi'(y) \). It remains to show
\[
\varphi(y + v) - \varphi(y) - \varphi'(y)v = o(v). \tag{13.28}
\]
Let us abbreviate \( u = \varphi(y + v) - \varphi(y) \), then using (13.27) and the fixed point property of \( \varphi(y) \) we see
\[
(1 - \partial_y F(\varphi(y), y))(u - \varphi'(y)v) =
\[
= F(\varphi(y) + u, y + v) - F(\varphi(y), y) - \partial_y F(\varphi(y), y) u - \partial_y F(\varphi(y), y) v
\[
= o(u) + o(v) \tag{13.29}
\]
since \( F \in C^1(U \times V, U) \) by assumption. Moreover, \(|(1 - \partial_y F(\varphi(y), y))^{-1}| \leq (1 - \theta)^{-1}\) and \( u = O(v) \) (by (13.26)) implying \( u - \varphi'(y)v = o(v) \) as desired.

Finally, suppose that the result holds for some \( r - 1 \geq 1 \). Thus, if \( F \) is \( C^r \), then \( \varphi(y) \) is at least \( C^{r-1} \) and the fact that \( d\varphi(y) \) satisfies (13.27) implies \( \varphi(y) \in C^r(V, U) \).

As an important consequence we obtain the implicit function theorem.
Theorem 13.6 (Implicit function). Let $X$, $Y$, and $Z$ be Banach spaces and let $U, V$ be open subsets of $X$, $Y$, respectively. Let $F \in C^r(U \times V, Z)$, $r \geq 1$, and fix $(x_0, y_0) \in U \times V$. Suppose $\partial_x F(x_0, y_0) \in \mathcal{L}(X, Z)$ is an isomorphism. Then there exists an open neighborhood $U_1 \times V_1 \subseteq U \times V$ of $(x_0, y_0)$ such that for each $y \in V_1$ there exists a unique point $(\xi(y), y) \in U_1 \times V_1$ satisfying $F(\xi(y), y) = F(x_0, y_0)$. Moreover, the map $\xi$ is in $C^r(V_1, Z)$ and fulfills

$$d\xi(y) = - (\partial_x F(\xi(y), y))^{-1} \circ \partial_y F(\xi(y), y).$$

(13.30)

Proof. Using the shift $F \rightarrow F - F(x_0, y_0)$ we can assume $F(x_0, y_0) = 0$. Next, the fixed points of $G(x, y) = x - (\partial_x F(x_0, y_0))^{-1} F(x, y)$ are the solutions of $F(x, y) = 0$. The function $G$ has the same smoothness properties as $F$ and since $|\partial_x G(x_0, y_0)| = 0$, we can find balls $U_1$ and $V_1$ around $x_0$ and $y_0$ such that $|\partial_x G(x, y)| \leq \theta < 1$. Thus $G(., y)$ is a uniform contraction and in particular, $G(U_1, y) \subset U_1$, that is, $G : U_1 \times V_1 \rightarrow U_1$. The rest follows from the uniform contraction principle. Formula (13.30) follows from differentiating $F(\xi(y), y) = 0$ using the chain rule. \qed

Note that our proof is constructive, since it shows that the solution $\xi(y)$ can be obtained by iterating $x - (\partial_x F(x_0, y_0))^{-1} F(x, y)$.

Moreover, as a corollary of the implicit function theorem we also obtain the inverse function theorem.

Theorem 13.7 (Inverse function). Suppose $F \in C^r(U, Y)$, $U \subseteq X$, and let $dF(x_0)$ be an isomorphism for some $x_0 \in U$. Then there are neighborhoods $U_1$, $V_1$ of $x_0$, $F(x_0)$, respectively, such that $F \in C^r(U_1, V_1)$ is a diffeomorphism.

Proof. Apply the implicit function theorem to $G(x, y) = y - F(x)$. \qed

13.3. Ordinary differential equations

As a first application of the implicit function theorem, we prove (local) existence and uniqueness for solutions of ordinary differential equations in Banach spaces. Let $X$ be an Banach space and $U \subseteq X$. Denote by $C_b(I, U)$ the Banach space of bounded continuous functions equipped with the sup norm.

The following lemma will be needed in the proof.

Lemma 13.8. Suppose $I \subseteq \mathbb{R}$ is a compact interval and $f \in C^r(U, Y)$. Then $f_* \in C^r(C_b(I, U), C_b(I, Y))$, where

$$(f_* x)(t) = f(x(t)).$$

(13.31)
Proof. Fix \(x_0 \in C_b(I,U)\) and \(\varepsilon > 0\). For each \(t \in I\) we have a \(\delta(t) > 0\) such that \(B_{\delta(t)}(x_0(t)) \subset U\) and \(|f(x) - f(x_0(t))| \leq \varepsilon/2\) for all \(x\) with \(|x - x_0(t)| \leq 2\delta(t)\). The balls \(B_{\delta(t)}(x_0(t)), t \in I\), cover the set \(\{x_0(t)\}_{t \in I}\) and since \(I\) is compact, there is a finite subcover \(B_{\delta(t)}(x_0(t)), 1 \leq j \leq n\). Let \(|x - x_0| \leq \delta = \min_{1 \leq j \leq n} \delta(t_j)\). Then for each \(t \in I\) there is \(t_i\) such that \(|x_0(t) - x_0(t_j)| \leq \delta(t_j)\) and hence \(|f(x(t)) - f(x_0(t))| \leq \varepsilon\) since \(|x(t) - x_0(t)| \leq |x(t) - x_0(t)| + |x_0(t) - x_0(t_j)| \leq 2\delta(t_j)\). This settles the case \(r = 0\).

Next let us turn to \(r = 1\). We claim that \(df_{\ast}\) is given by \((df_{\ast}(x_0))x(t) = df(x_0(t))x(t)\). Hence we need to show that for each \(\varepsilon > 0\) we can find a \(\delta > 0\) such that
\[
\sup_{t \in I} |f(x_0(t) + x(t)) - f(x_0(t)) - df(x_0(t))x(t)| \leq \varepsilon \sup_{t \in I} |x(t)| \tag{13.32}
\]
whenever \(|x| = \sup_{t \in I} |x(t)| \leq \delta\). By assumption we have
\[
|f(x_0(t) + x(t)) - f(x_0(t)) - df(x_0(t))x(t)| \leq \varepsilon |x(t)| \tag{13.33}
\]
whenever \(|x(t)| \leq \delta(t)\). Now argue as before to show that \(\delta(t)\) can be chosen independent of \(t\). It remains to show that \(df_{\ast}\) is continuous. To see this we use the linear map
\[
\lambda : C_b(I, \mathcal{L}(X,Y)) \rightarrow \mathcal{L}(C_b(I, X), C_b(I, Y)) \tag{13.34}
\]
where \((T_{\ast} x)(t) = T(t)x(t)\). Since we have
\[
|T_{\ast} x| = \sup_{t \in I} |T(t)x(t)| \leq \sup_{t \in I} |T(t)||x(t)| \leq |T||x|, \tag{13.35}
\]
we infer \(|\lambda| \leq 1\) and hence \(\lambda\) is continuous. Now observe \(df_{\ast} = \lambda \circ (df)_{\ast}\).

The general case \(r > 1\) follows from induction. \(\square\)

Now we come to our existence and uniqueness result for the initial value problem in Banach spaces.

**Theorem 13.9.** Let \(I\) be an open interval, \(U\) an open subset of a Banach space \(X\) and \(\Lambda\) an open subset of another Banach space. Suppose \(F \in C^r(I \times U \times \Lambda, X)\), then the initial value problem
\[
\dot{x}(t) = F(t,x,\lambda), \quad x(t_0) = x_0, \quad (t_0, x_0, \lambda) \in I \times U \times \Lambda, \tag{13.36}
\]
has a unique solution \(x(t, t_0, x_0, \lambda) \in C^r(I, I_1 \times U_1 \times \Lambda_1, X)\), where \(I_1, U_1,\) and \(\Lambda_1\) are open subsets of \(I, U,\) and \(\Lambda\), respectively. The sets \(I_1, U_1,\) and \(\Lambda_1\) can be chosen to contain any point \(t_0 \in I, x_0 \in U,\) and \(\lambda_0 \in \Lambda,\) respectively.
Proof. If we shift $t \to t - t_0$, $x \to x - x_0$, and hence $F \to F(. + t_0, . + x_0, \lambda)$, we see that it is no restriction to assume $x_0 = 0$, $t_0 = 0$ and to consider $(t_0, x_0)$ as part of the parameter $\lambda$ (i.e., $\lambda \to (t_0, x_0, \lambda)$). Moreover, using the standard transformation $\dot{x} = F(\tau, x, \lambda)$, $\dot{\tau} = 1$, we can even assume that $F$ is independent of $t$.

Our goal is to invoke the implicit function theorem. In order to do this we introduce an additional parameter $\varepsilon \in \mathbb{R}$ and consider

$$\dot{x} = \varepsilon F(x, \lambda), \quad x \in D^{r+1} = \{x \in C^{r+1}_b((-1, 1), U) | x(0) = 0\}, \quad (13.37)$$

such that we know the solution for $\varepsilon = 0$. The implicit function theorem will show that solutions still exist as long as $\varepsilon$ remains small. At first sight this doesn’t seem to be good enough for us since our original problem corresponds to $\varepsilon = 1$. But since $\varepsilon$ corresponds to a scaling $t \to \varepsilon t$, the solution for one $\varepsilon > 0$ suffices. Now let us turn to the details.

Our problem (13.37) is equivalent to looking for zeros of the function

$$G : D^{r+1} \times \Lambda \times (-\varepsilon_0, \varepsilon_0) \to C^{r}_b((-1, 1), X),$$

$$G(x, \lambda, \varepsilon) \mapsto \dot{x} - \varepsilon F(x, \lambda). \quad (13.38)$$

Lemma 13.8 ensures that this function is $C^r$. Now fix $\lambda_0$, then $G(0, \lambda_0, 0) = 0$ and $\partial_x G(0, \lambda_0, 0) = T$, where $Tx = \dot{x}$. Since $(T^{-1}x)(t) = \int_0^t x(s)ds$ we can apply the implicit function theorem to conclude that there is a unique solution $x(\lambda, \varepsilon) \in C^r(\Lambda_1 \times (-\varepsilon_0, \varepsilon_0), D^{r+1})$. In particular, the map $(\lambda, t) \mapsto x(\lambda, \varepsilon)(t/\varepsilon)$ is in $C^r(\Lambda_1, C^{r+1}((-\varepsilon, \varepsilon), X)) \hookrightarrow C^r(\Lambda \times (-\varepsilon, \varepsilon), X)$. Hence it is the desired solution of our original problem. \qed
Chapter 14

The Brouwer mapping degree

14.1. Introduction

Many applications lead to the problem of finding all zeros of a mapping \( f : U \subseteq X \rightarrow X \), where \( X \) is some (real) Banach space. That is, we are interested in the solutions of

\[
f(x) = 0, \quad x \in U.
\]

(14.1)

In most cases it turns out that this is too much to ask for, since determining the zeros analytically is in general impossible.

Hence one has to ask some weaker questions and hope to find answers for them. One such question would be "Are there any solutions, respectively, how many are there?". Luckily, these questions allow some progress.

To see how, let’s consider the case \( f \in \mathcal{H}(\mathbb{C}) \), where \( \mathcal{H}(U) \) denotes the set of holomorphic functions on a domain \( U \subseteq \mathbb{C} \). Recall the concept of the winding number from complex analysis. The winding number of a path \( \gamma : [0, 1] \rightarrow \mathbb{C} \) around a point \( z_0 \in \mathbb{C} \) is defined by

\[
n(\gamma, z_0) = \frac{1}{2\pi i} \int_\gamma \frac{dz}{z - z_0} \in \mathbb{Z}.
\]

(14.2)

It gives the number of times \( \gamma \) encircles \( z_0 \) taking orientation into account. That is, encirclings in opposite directions are counted with opposite signs.

In particular, if we pick \( f \in \mathcal{H}(\mathbb{C}) \) one computes (assuming \( 0 \notin f(\gamma) \))

\[
n(f(\gamma), 0) = \frac{1}{2\pi i} \int_\gamma \frac{f'(z)}{f(z)} dz = \sum_k n(\gamma, z_k) \alpha_k,
\]

(14.3)
where \( z_k \) denote the zeros of \( f \) and \( \alpha_k \) their respective multiplicity. Moreover, if \( \gamma \) is a Jordan curve encircling a simply connected domain \( U \subset \mathbb{C} \), then 
\[ n(\gamma, z_k) = 0 \text{ if } z_k \not\in U \text{ and } n(\gamma, z_k) = 1 \text{ if } z_k \in U. \]
Hence \( n(f(\gamma), 0) \) counts the number of zeros inside \( U \).

However, this result is useless unless we have an efficient way of computing \( n(f(\gamma), 0) \) (which does not involve the knowledge of the zeros \( z_k \)). This is our next task.

Now, let's recall how one would compute complex integrals along complicated paths. Clearly, one would use homotopy invariance and look for a simpler path along which the integral can be computed and which is homotopic to the original one. In particular, if \( f : \gamma \to \mathbb{C} \setminus \{0\} \) and \( g : \gamma \to \mathbb{C} \setminus \{0\} \) are homotopic, we have \( n(f(\gamma), 0) = n(g(\gamma), 0) \) (which is known as Rouché's theorem).

More explicitly, we need to find a mapping \( g \) for which \( n(g(\gamma), 0) \) can be computed and a homotopy \( H : [0, 1] \times \gamma \to \mathbb{C} \setminus \{0\} \) such that \( H(0, z) = f(z) \) and \( H(1, z) = g(z) \) for \( z \in \gamma \). For example, how many zeros of \( f(z) = \frac{1}{2}z^6 + z - \frac{1}{3} \) lie inside the unit circle? Consider \( g(z) = z \), then \( H(t, z) = (1-t)f(z) + tg(z) \) is the required homotopy since \( |f(z) - g(z)| < |g(z)| \), \( |z| = 1 \), implying \( H(t, z) \neq 0 \) on \([0, 1] \times \gamma\). Hence \( f(z) \) has one zero inside the unit circle.

Summarizing, given a (sufficiently smooth) domain \( U \) with enclosing Jordan curve \( \partial U \), we have defined a degree \( \deg(f, U, z_0) = n(f(\partial U), z_0) = n(f(\partial U) - z_0, 0) \in \mathbb{Z} \) which counts the number of solutions of \( f(z) = z_0 \) inside \( U \). The invariance of this degree with respect to certain deformations of \( f \) allowed us to explicitly compute \( \deg(f, U, z_0) \) even in nontrivial cases.

Our ultimate goal is to extend this approach to continuous functions \( f : \mathbb{R}^n \to \mathbb{R}^n \). However, such a generalization runs into several problems. First of all, it is unclear how one should define the multiplicity of a zero.

But even more severe is the fact, that the number of zeros is unstable with respect to small perturbations. For example, consider \( f_\varepsilon : [-1, 2] \to \mathbb{R}, x \mapsto x^2 - \varepsilon \). Then \( f_\varepsilon \) has no zeros for \( \varepsilon < 0 \), one zero for \( \varepsilon = 0 \), two zeros for \( 0 < \varepsilon \leq 1 \), one for \( 1 < \varepsilon \leq \sqrt{2} \), and none for \( \varepsilon > \sqrt{2} \). This shows the following facts.

(i) Zeros with \( f' \neq 0 \) are stable under small perturbations.

(ii) The number of zeros can change if two zeros with opposite sign change (i.e., opposite signs of \( f' \)) run into each other.

(iii) The number of zeros can change if a zero drops over the boundary.

Hence we see that we cannot expect too much from our degree. In addition, since it is unclear how it should be defined, we will first require some basic
14.2. Definition of the mapping degree and the determinant formula

To begin with, let us introduce some useful notation. Throughout this section \( U \) will be a bounded open subset of \( \mathbb{R}^n \). For \( f \in C^1(U, \mathbb{R}^n) \) the Jacobi matrix of \( f \) at \( x \in U \) is
\[
 df(x) = (\partial_x f_j(x))_{1 \leq i,j \leq n}
\]
The Jacobi determinant of \( f \) at \( x \in U \) is
\[
 J_f(x) = \det df(x). \tag{14.4}
\]

The set of \textbf{regular values}
is
\[
 RV(f) = \{ y \in \mathbb{R}^n | \forall x \in f^{-1}(y) : J_f(x) \neq 0 \}. \tag{14.5}
\]
Its complement \( CV(f) = \mathbb{R}^n \setminus RV(f) \) is called the set of \textbf{critical values}. We set \( C^r(U, \mathbb{R}^n) = \{ f \in C^r(U, \mathbb{R}^n) | df\ f \in C(\mathbb{U}, \mathbb{R}^n), 0 \leq j \leq r \} \) and
\[
 D_y^r(\mathbb{U}, \mathbb{R}^n) = \{ f \in C^r(\mathbb{U}, \mathbb{R}^n) | y \notin f(\partial U) \}, \quad D_y(\mathbb{U}, \mathbb{R}^n) = D_y^0(\mathbb{U}, \mathbb{R}^n) \tag{14.6}
\]
for \( y \in \mathbb{R}^n \). We will use the topology induced by the sup norm for \( C^r(\mathbb{U}, \mathbb{R}^n) \) such that it becomes a Banach space (cf. Section 13.1).

Note that, since \( U \) is bounded, \( \partial U \) is compact and so is \( f(\partial U) \) if \( f \in C(\mathbb{U}, \mathbb{R}^n) \). In particular,
\[
 \text{dist}(y, f(\partial U)) = \inf_{x \in \partial U} |y - f(x)| \tag{14.7}
\]
is positive for \( f \in D_y(\mathbb{U}, \mathbb{R}^n) \) and thus \( D_y(\mathbb{U}, \mathbb{R}^n) \) is an open subset of \( C^r(\mathbb{U}, \mathbb{R}^n) \).

Now that these things are out of the way, we come to the formulation of the requirements for our degree.

A function \( \text{deg} \) which assigns each \( f \in D_y(\mathbb{U}, \mathbb{R}^n), y \in \mathbb{R}^n \), a real number \( \text{deg}(f, U, y) \) will be called degree if it satisfies the following conditions.

(D1). \( \text{deg}(f, U, y) = \text{deg}(f - y, U, 0) \) (translation invariance).

(D2). \( \text{deg}(1, U, y) = 1 \) if \( y \in U \) (normalization).

(D3). If \( U_{1,2} \) are open, disjoint subsets of \( U \) such that \( y \notin f(\mathbb{U} \setminus (U_1 \cup U_2)) \), then \( \text{deg}(f, U, y) = \text{deg}(f, U_1, y) + \text{deg}(f, U_2, y) \) (additivity).

(D4). If \( H(t) = (1 - t)f + tg \in D_y(\mathbb{U}, \mathbb{R}^n), t \in [0, 1] \), then \( \text{deg}(f, U, y) = \text{deg}(g, U, y) \) (homotopy invariance).

Before we draw some first conclusions form this definition, let us discuss the properties (D1)–(D4) first. (D1) is natural since \( \text{deg}(f, U, y) \) should have something to do with the solutions of \( f(x) = y, x \in U \), which is the same
as the solutions of \( f(x) - y = 0, \ x \in U \). (D2) is a normalization since any multiple of \( \text{deg} \) would also satisfy the other requirements. (D3) is also quite natural since it requires \( \text{deg} \) to be additive with respect to components. In addition, it implies that sets where \( f \neq y \) do not contribute. (D4) is not that natural since it already rules out the case where \( \text{deg} \) is the cardinality of \( f^{-1}(U) \). On the other hand it will give us the ability to compute \( \text{deg}(f,U,y) \) in several cases.

**Theorem 14.1.** Suppose \( \text{deg} \) satisfies (D1)–(D4) and let \( f, g \in D_y(\overline{U}, \mathbb{R}^n) \), then the following statements hold.

(i). We have \( \text{deg}(f,\emptyset,y) = 0 \). Moreover, if \( U_i, \ 1 \leq i \leq N \), are disjoint open subsets of \( U \) such that \( y \not\in f(\overline{U}\setminus \bigcup_{i=1}^{N} U_i) \), then \( \text{deg}(f,U,y) = \sum_{i=1}^{N} \text{deg}(f,U_i,y) \).

(ii). If \( y \not\in f(U) \), then \( \text{deg}(f,U,y) = 0 \) (but not the other way round). Equivalently, if \( \text{deg}(f,U,y) \neq 0 \), then \( y \in f(U) \).

(iii). If \( |f(x) - g(x)| < \text{dist}(y,f(\partial U)), \ x \in \partial U \), then \( \text{deg}(f,U,y) = \text{deg}(g,U,y) \). In particular, this is true if \( f(x) = g(x) \) for \( x \in \partial U \).

**Proof.** For the first part of (i) use (D3) with \( U_1 = U \) and \( U_2 = \emptyset \). For the second part use \( U_2 = \emptyset \) in (D3) if \( i = 1 \) and the rest follows from induction. For (ii) use \( i = 1 \) and \( U_1 = \emptyset \) in (ii). For (iii) note that \( H(t,x) = (1 - t)f(x) + tg(x) \) satisfies \( |H(t,x) - y| \geq \text{dist}(y,f(\partial U)) - |f(x) - g(x)| \) for \( x \) on the boundary. \( \square \)

Next we show that (D.4) implies several at first sight much stronger looking facts.

**Theorem 14.2.** We have that \( \text{deg}(.,U,y) \) and \( \text{deg}(f,.,.) \) are both continuous. In fact, we even have

(i). \( \text{deg}(.,U,y) \) is constant on each component of \( D_y(\overline{U}, \mathbb{R}^n) \).

(ii). \( \text{deg}(f,.,.) \) is constant on each component of \( \mathbb{R}^n \setminus f(\partial U) \).

Moreover, if \( H : [0,1] \times \overline{U} \rightarrow \mathbb{R}^n \) and \( y : [0,1] \rightarrow \mathbb{R}^n \) are both continuous such that \( H(t) \in D_y(\overline{U}, \mathbb{R}^n), \ t \in [0,1], \) then \( \text{deg}(H(0),U,y(0)) = \text{deg}(H(1),U,y(1)) \).

**Proof.** For (i) let \( C \) be a component of \( D_y(\overline{U}, \mathbb{R}^n) \) and let \( d_0 \in \text{deg}(C,U,y) \). It suffices to show that \( \text{deg}(.,U,y) \) is locally constant. But if \( |g - f| < \text{dist}(y,f(\partial U)) \), then \( \text{deg}(f,U,y) = \text{deg}(g,U,y) \) by (D.4) since \( |H(t) - y| \geq |f - y| - |g - f| > 0 \), \( H(t) = (1 - t)f + tg \). The proof of (ii) is similar. For the remaining part observe, that if \( H : [0,1] \times \overline{U} \rightarrow \mathbb{R}^n, \ (t,x) \mapsto H(t,x), \) is continuous, then so is \( H : [0,1] \rightarrow C(\overline{U}, \mathbb{R}^n) \), \( t \mapsto H(t) \), since \( \overline{U} \) is compact. Hence, if in addition \( H(t) \in D_y(\overline{U}, \mathbb{R}^n) \), then \( \text{deg}(H(t),U,y) \) is independent
of \( t \) and if \( y = y(t) \) we can use \( \deg(H(0), U, y(0)) = \deg(H(t) - y(t), U, 0) = \deg(H(1), U, y(1)). \)

Note that this result also shows why \( \deg(f, U, y) \) cannot be defined meaningful for \( y \in f(\partial D) \). Indeed, approaching \( y \) from within different components of \( \mathbb{R}^n \setminus f(\partial U) \) will result in different limits in general!

In addition, note that if \( Q \) is a closed subset of a locally pathwise connected space \( X \), then the components of \( X \setminus Q \) are open (in the topology of \( X \)) and pathwise connected (the set of points for which a path to a fixed point \( x_0 \) exists is both open and closed).

Now let us try to compute \( \deg \) using its properties. Lets start with a simple case and suppose \( f \in C^1(U, \mathbb{R}^n) \) and \( y \notin CV(f) \cup f(\partial U) \). Without restriction we consider \( y = 0 \). In addition, we avoid the trivial case \( f^{-1}(0) = \emptyset \). Since the points of \( f^{-1}(0) \) inside \( U \) are isolated (use \( Jf(x) \neq 0 \) and the inverse function theorem) they can only cluster at the boundary \( \partial U \).

But this is also impossible since \( f \) would equal \( y \) at the limit point on the boundary by continuity. Hence \( f^{-1}(0) = \{x^i\}_{i=1}^N \). Picking sufficiently small neighborhoods \( U(x^i) \) around \( x^i \) we consequently get

\[
\deg(f, U, 0) = \sum_{i=1}^N \deg(f, U(x^i), 0). \quad (14.8)
\]

It suffices to consider one of the zeros, say \( x^1 \). Moreover, we can even assume \( x^1 = 0 \) and \( U(x^1) = B_\delta(0) \). Next we replace \( f \) by its linear approximation around \( 0 \). By the definition of the derivative we have

\[
f(x) = df(0)x + |x|r(x), \quad r \in C(B_\delta(0), \mathbb{R}^n), \quad r(0) = 0. \quad (14.9)
\]

Now consider the homotopy \( H(t, x) = df(0)x + (1 - t)|x|r(x) \). In order to conclude \( \deg(f, B_\delta(0), 0) = \deg(df(0), B_\delta(0), 0) \) we need to show \( 0 \notin H(t, \partial B_\delta(0)) \). Since \( Jf(0) \neq 0 \) we can find a constant \( \lambda \) such that \( |df(0)x| \geq \lambda|x| \) and since \( r(0) = 0 \) we can decrease \( \delta \) such that \( |r| < \lambda \). This implies \( |H(t, x)| \geq ||df(0)x| - (1 - t)|x||r(x)|| \geq \lambda\delta - \delta|r| > 0 \) for \( x \in \partial B_\delta(0) \) as desired.

In summary we have

\[
\deg(f, U, 0) = \sum_{i=1}^N \deg(df(x^i), U(x^i), 0) \quad (14.10)
\]

and it remains to compute the degree of a nonsingular matrix. To this end we need the following lemma.

**Lemma 14.3.** Two nonsingular matrices \( M_{1,2} \in \text{GL}(n) \) are homotopic in \( \text{GL}(n) \) if and only if \( \text{sign det } M_1 = \text{sign det } M_2 \).
Theorem 14.4. Suppose would have to vanish somewhere in between (i.e., we would leave $GL(n)$ cannot change the sign of the determinant since otherwise the determinant shows that $\text{diag}(\pm 1)$ cannot change the sign of the determinant since otherwise the determinant shows that $\text{diag}(\pm 1, 1)$ are homotopic. Now we apply this result to all two by two subblocks as follows. For each $i$ starting from $n$ and going down to 2 transform the subblock $\text{diag}(m_{i-1}, m_i)$ into $\text{diag}(1, 1)$ respectively $\text{diag}(-1, 1)$. The result is the desired form for $M$.

To conclude the proof note that a continuous deformation within $GL(n)$ cannot change the sign of the determinant since otherwise the determinant would have to vanish somewhere in between (i.e., we would leave $GL(n)$). □

Proof. We will show that any given nonsingular matrix $M$ is homotopic to $\text{diag}(\text{sign det } M, 1, \ldots, 1)$, where $\text{diag}(m_1, \ldots, m_n)$ denotes a diagonal matrix with diagonal entries $m_i$.

In fact, note that adding one row to another and multiplying a row by a positive constant can be realized by continuous deformations such that all intermediate matrices are nonsingular. Hence we can reduce $M$ to a diagonal matrix $\text{diag}(m_1, \ldots, m_n)$ with $(m_i)^2 = 1$. Next,

$$
\begin{pmatrix}
\pm \cos(\pi t) & \mp \sin(\pi t) \\
\sin(\pi t) & \cos(\pi t)
\end{pmatrix},
$$

(14.11)

shows that $\text{diag}(\pm 1, 1)$ and $\text{diag}(\mp 1, -1)$ are homotopic. Now we apply this result to all two by two subblocks as follows. For each $i$ starting from $n$ and going down to 2 transform the subblock $\text{diag}(m_{i-1}, m_i)$ into $\text{diag}(1, 1)$ respectively $\text{diag}(-1, 1)$. The result is the desired form for $M$.

Using this lemma we can now show the main result of this section.

Theorem 14.4. Suppose $f \in D^1_y(U, \mathbb{R}^n)$ and $y \notin CV(f)$, then a degree satisfying (D1)–(D4) satisfies

$$
\text{deg}(f, U, y) = \sum_{x \in f^{-1}(y)} \text{sign } J_f(x),
$$

(14.12)

where the sum is finite and we agree to set $\sum_{x \in \emptyset} = 0$.

Proof. By the previous lemma we obtain

$$
\text{deg}(df(0), B_\delta(0), 0) = \text{deg}(\text{diag}(\text{sign } J_f(0), 1, \ldots, 1), B_\delta(0), 0)
$$

(14.13)

since $\text{det } M \neq 0$ is equivalent to $Mx \neq 0$ for $x \in \partial B_\delta(0)$. Hence it remains to show $\text{deg}(df(0), B_\delta(0), 0) = \text{sign } J_f(0)$.

If $\text{sign } J_f(0) = 1$ this is true by (D2). Otherwise we can replace $df(0)$ by $M_- = \text{diag}(-1, 1, \ldots, 1)$ and it remains to show $\text{deg}(M_-, B_1(0), 0) = -1$.

Abbreviate $U_1 = B_1(0) = \{x \in \mathbb{R}^n| |x_i| < 1, 1 \leq i \leq n\}$, $U_2 = \{x \in \mathbb{R}^n| 1 < x_1 < 3, |x_i| < 1, 2 \leq i \leq n\}$, $U = \{x \in \mathbb{R}^n|-1 < x_1 < 3, |x_i| < 1, 2 \leq i \leq n\}$, and $g(r) = 2 - |r - 1|$, $h(r) = 1 - r^2$. Now consider the two functions $f_1(x) = (1 - g(x_1)h(x_2) \cdots h(x_n), x_2, \ldots, x_n)$ and $f_2(x) = (1, x_2, \ldots, x_n)$. Clearly $f_1^{-1}(0) = \{x^1, x^2\}$ with $x^1 = 0$, $x^2 = (2, \ldots, 0)$ and $f_2^{-1}(0) = \emptyset$. Since $f_1(x) = f_2(x)$ for $x \in \partial U$ we infer $\text{deg}(f_1, U, 0) = \text{deg}(f_2, U, 0) = 0$. Moreover, we have $\text{deg}(f_1, U_1, 0) = \text{deg}(f_1, U_1, 0) + \text{deg}(f_1, U_2, 0)$ and hence $\text{deg}(M_-, U_1, 0) = \text{deg}(df_1(x^1)) = \text{deg}(f_1, U_1, 0) = -\text{deg}(f_1, U_2, 0) = -\text{deg}(df_1(x^2)) = -\text{deg}(I, U_2, 0) = -1$ as claimed. □
14.3. Extension of the determinant formula

Up to this point we have only shown that a degree (provided there is one at all) necessarily satisfies (14.12). Once we have shown that regular values are dense, it will follow that the degree is uniquely determined by (14.12) since the remaining values follow from point (iii) of Theorem 14.1. On the other hand, we don’t even know whether a degree exists. Hence we need to show that (14.12) can be extended to \( f \in D_y(U, \mathbb{R}^n) \) and that this extension satisfies our requirements (D1)–(D4).

14.3. Extension of the determinant formula

Our present objective is to show that the determinant formula (14.12) can be extended to all \( f \in D_y(U, \mathbb{R}^n) \). This will be done in two steps, where we will show that \( \text{deg}(f, U, y) \) as defined in (14.12) is locally constant with respect to both \( y \) (step one) and \( f \) (step two).

Before we work out the technical details for these two steps, we prove that the set of regular values is dense as a warm up. This is a consequence of a special case of Sard’s theorem which says that \( CV(f) \) has zero measure.

Lemma 14.5 (Sard). Suppose \( f \in C^1(U, \mathbb{R}^n) \), then the Lebesgue measure of \( CV(f) \) is zero.

Proof. Since the claim is easy for linear mappings our strategy is as follows. We divide \( U \) into sufficiently small subsets. Then we replace \( f \) by its linear approximation in each subset and estimate the error.

Let \( CP(f) = \{ x \in U | J_f(x) = 0 \} \) be the set of critical points of \( f \). We first pass to cubes which are easier to divide. Let \( \{ Q_i \}_{i \in \mathbb{N}} \) be a countable cover for \( U \) consisting of open cubes such that \( \overline{Q}_i \subset U \). Then it suffices to prove that \( f(CP(f) \cap Q_i) \) has zero measure since \( CV(f) = f(CP(f)) = \bigcup_i f(CP(f) \cap Q_i) \) (the \( Q_i \)’s are a cover).

Let \( Q \) be any of these cubes and denote by \( \rho \) the length of its edges. Fix \( \varepsilon > 0 \) and divide \( Q \) into \( N^n \) cubes \( Q_i \) of length \( \rho/N \). Since \( df(x) \) is uniformly continuous on \( Q \) we can find an \( N \) (independent of \( i \)) such that

\[
|f(x) - f(\tilde{x}) - df(\tilde{x})(x - \tilde{x})| \leq \int_0^1 |df(\tilde{x} + t(x - \tilde{x})) - df(\tilde{x})||\tilde{x} - x|dt \leq \frac{\varepsilon \rho}{N}
\]

(14.14)

for \( \tilde{x}, x \in Q_i \). Now pick a \( Q_i \) which contains a critical point \( \tilde{x}_i \in CP(f) \). Without restriction we assume \( \tilde{x}_i = 0 \), \( f(\tilde{x}_i) = 0 \) and set \( M = df(\tilde{x}_i) \). By \( \det M = 0 \) there is an orthonormal basis \( \{ b^i \}_{1 \leq i \leq n} \) of \( \mathbb{R}^n \) such that \( b^n \) is orthogonal to the image of \( M \). In addition,

\[
Q_i \subseteq \left\{ \sum_{i=1}^n \lambda_i b^i \left| \sqrt{\sum_{i=1}^n |\lambda_i|^2} \leq \sqrt{n} \frac{\rho}{N} \right\} \subseteq \left\{ \sum_{i=1}^n \lambda_i b^i \left| |\lambda_i| \leq \sqrt{n} \frac{\rho}{N} \right\}
\]
and hence there is a constant (again independent of \(i\)) such that
\[
MQ_i \subseteq \{ \sum_{i=1}^{n-1} \lambda_i b^i |\lambda_i| \leq C \frac{\rho}{N} \}
\]  
(14.15)
(e.g., \(C = \sqrt{n} \max_{x \in Q} |df(x)|\)). Next, by our estimate (14.14) we even have
\[
f(Q_i) \subseteq \{ \sum_{i=1}^{n} \lambda_i b^i |\lambda_i| \leq (C + \varepsilon) \frac{\rho}{N}, |\lambda_n| \leq \varepsilon \frac{\rho}{N} \}
\]  
(14.16)
and hence the measure of \(f(Q_i)\) is smaller than \(\tilde{C}\varepsilon\). Since there are at most \(N^n\) such \(Q_i\)’s, we see that the measure of \(f(CP(f) \cap Q)\) is smaller than \(\tilde{C}\varepsilon\).

Having this result out of the way we can come to step one and two from above.

**Step 1: Admitting critical values**

By (ii) of Theorem 14.2, \(\deg(f, U, y)\) should be constant on each component of \(\mathbb{R}^n \setminus f(\partial U)\). Unfortunately, if we connect \(y\) and a nearby regular value \(\tilde{y}\) by a path, then there might be some critical values in between. To overcome this problem we need a definition for \(\deg\) which works for critical values as well. Let us try to look for an integral representation. Formally (14.12) can be written as \(\deg(f, U, y) = \int_U \delta_y(f(x)) J_f(x) dx\), where \(\delta_y(.)\) is the Dirac distribution at \(y\). But since we don’t want to mess with distributions, we replace \(\delta_y(.)\) by \(\phi_\varepsilon(-y)\), where \(\{\phi_\varepsilon\}_{\varepsilon>0}\) is a family of functions such that \(\phi_\varepsilon\) is supported on the ball \(B_\varepsilon(0)\) of radius \(\varepsilon\) around 0 and satisfies \(\int_{\mathbb{R}^n} \phi_\varepsilon(x) dx = 1\).

**Lemma 14.6.** Let \(f \in D^1_y(U, \mathbb{R}^n), y \notin CV(f)\). Then
\[
\deg(f, U, y) = \int_U \phi_\varepsilon(f(x) - y) J_f(x) dx
\]  
(14.17)
for all positive \(\varepsilon\) smaller than a certain \(\varepsilon_0\) depending on \(f\) and \(y\). Moreover, \(\text{supp}(\phi_\varepsilon(f, -y)) \subseteq U\) for \(\varepsilon < \text{dist}(y, f(\partial U))\).

**Proof.** If \(f^{-1}(y) = \emptyset\), we can set \(\varepsilon_0 = \text{dist}(y, f(\overline{U}))\), implying \(\phi_\varepsilon(f(x)-y) = 0\) for \(x \in \overline{U}\).

If \(f^{-1}(y) = \{x^i\}_{1 \leq i \leq N}\), we can find an \(\varepsilon_0 > 0\) such that \(f^{-1}(B_{\varepsilon_0}(y))\) is a union of disjoint neighborhoods \(U(x^i)\) of \(x^i\) by the inverse function theorem. Moreover, after possibly decreasing \(\varepsilon_0\) we can assume that \(f|_{U(x^i)}\) is a bijection and that \(J_f(x)\) is nonzero on \(U(x^i)\). Again \(\phi_\varepsilon(f(x) - y) = 0\)
for \( x \in \bigcup_{i=1}^N U(x^i) \) and hence
\[
\int_U \phi_\varepsilon(f(x) - y)J_f(x) dx = \sum_{i=1}^N \int_{U(x^i)} \phi_\varepsilon(f(x) - y)J_f(x) dx
\]
\[
= \sum_{i=1}^N \text{sign}(J_f(x)) \int_{B_{c_0}(0)} \phi_\varepsilon(\tilde{x}) d\tilde{x} = \text{deg}(f, U, y),
\]
(14.18)
where we have used the change of variables \( \tilde{x} = f(x) \) in the second step. □

Our new integral representation makes sense even for critical values. But since \( \varepsilon \) depends on \( y \), continuity with respect to \( y \) is not clear. This will be shown next at the expense of requiring \( f \in C^2 \) rather than \( f \in C^1 \).

The key idea is to rewrite \( \text{deg}(f, U, y^2) - \text{deg}(f, U, y^1) \) as an integral over a divergence (here we will need \( f \in C^2 \)) supported in \( U \) and then apply Stokes theorem. For this purpose the following result will be used.

Lemma 14.7. Suppose \( f \in C^2(U, \mathbb{R}^n) \) and \( u \in C^1(\mathbb{R}^n, \mathbb{R}^n) \), then
\[
(\text{div } u)(f)J_f = \text{div } D_f(u),
\]
(14.19)
where \( D_f(u)_j \) is the determinant of the matrix obtained from \( df \) by replacing the \( j \)-th column by \( u(f) \).

Proof. We compute
\[
\text{div } D_f(u) = \sum_{j=1}^n \partial_{x_j} D_f(u)_j = \sum_{j,k=1}^n D_f(u)_{j,k},
\]
(14.20)
where \( D_f(u)_{j,k} \) is the determinant of the matrix obtained from the matrix associated with \( D_f(u)_j \) by applying \( \partial_{x_j} \) to the \( k \)-th column. Since \( \partial_{x_j} \partial_{x_k} f = \partial_{x_k} \partial_{x_j} f \) we infer \( D_f(u)_{j,k} = -D_f(u)_{k,j}, j \neq k \), by exchanging the \( k \)-th and the \( j \)-th column. Hence
\[
\text{div } D_f(u) = \sum_{i=1}^n D_f(u)_{i,i}.
\]
(14.21)
Now let \( J_f^{(i,j)}(x) \) denote the \((i, j)\) minor of \( df(x) \) and recall \( \sum_{i=1}^n J_f^{(i,j)} \partial_{x_i} f_k = \delta_{j,k}J_f \). Using this to expand the determinant \( D_f(u)_{i,i} \) along the \( i \)-th column shows
\[
\text{div } D_f(u) = \sum_{i,j=1}^n J_f^{(i,j)} \partial_{x_j} u_i(f) = \sum_{i,j=1}^n J_f^{(i,j)} \sum_{k=1}^n (\partial_{x_k} u_j(f)) \partial_{x_i} f_k
\]
\[
= \sum_{j,k=1}^n (\partial_{x_k} u_j(f)) \sum_{i=1}^n J_f^{(i,j)} \partial_{x_i} f_k = \sum_{j=1}^n (\partial_{x_j} u_j(f)) J_f
\]
(14.22)
as required. □
Now we can prove

**Lemma 14.8.** Suppose \( f \in C^2(\overline{U}, \mathbb{R}^n) \). Then \( \deg(f, U, \cdot) \) is constant in each ball contained in \( \mathbb{R}^n \setminus f(\partial U) \), whenever defined.

**Proof.** Fix \( \tilde{y} \in \mathbb{R}^n \setminus f(\partial U) \) and consider the largest ball \( B_\rho(\tilde{y}) \), \( \rho = \text{dist}(\tilde{y}, f(\partial U)) \) around \( \tilde{y} \) contained in \( \mathbb{R}^n \setminus f(\partial U) \). Pick \( y^i \in B_\rho(\tilde{y}) \cap RV(f) \) and consider

\[
\deg(f, U, y^2) - \deg(f, U, y^1) = \int_U (\phi_\varepsilon(f(x) - y^2) - \phi_\varepsilon(f(x) - y^1)) J_f(x) \, dx
\]

(14.23)

for suitable \( \phi_\varepsilon \in C^2(\mathbb{R}^n, \mathbb{R}) \) and suitable \( \varepsilon > 0 \). Now observe

\[
(\text{div } u)(y) = \int_0^1 z_j \partial y_j \phi(y + tz) \, dt
\]

\[
= \int_0^1 \left( \frac{d}{dt} \phi(y + tz) \right) dt = \phi_\varepsilon(y - y^2) - \phi_\varepsilon(y - y^1),
\]

(14.24)

where

\[
u(y) = z \int_0^1 \phi(y + tz) \, dt, \quad \phi(y) = \phi_\varepsilon(y - y^1), \quad z = y^1 - y^2,
\]

(14.25)

and apply the previous lemma to rewrite the integral as \( \int_U \text{div } Df(u) \, dx \). Since the integrand vanishes in a neighborhood of \( \partial U \) it is no restriction to assume that \( \partial U \) is smooth such that we can apply Stokes theorem. Hence we have \( \int_U \text{div } Df(u) \, dx = \int_{\partial U} Df(u) \, dF = 0 \) since \( u \) is supported inside \( B_\rho(\tilde{y}) \) provided \( \varepsilon \) is small enough (e.g., \( \varepsilon < \rho - \max \{|y^i - \tilde{y}| : i = 1, 2\} \)). \( \square \)

As a consequence we can define

\[
\deg(f, U, y) = \deg(f, U, \tilde{y}), \quad y \notin f(\partial U), \quad f \in C^2(\overline{U}, \mathbb{R}^n),
\]

(14.26)

where \( \tilde{y} \) is a regular value of \( f \) with \( |\tilde{y} - y| < \text{dist}(y, f(\partial U)) \).

**Remark 14.9.** Let me remark a different approach due to Kronecker. For \( U \) with sufficiently smooth boundary we have

\[
\deg(f, U, 0) = \frac{1}{|S^{n-1}|} \int_{\partial U} Df(x) \, dF = \frac{1}{|S^n|} \int_{\partial U} \frac{1}{|f|^n} Df(x) \, dF, \quad \tilde{f} = \frac{f}{|f|},
\]

(14.27)

for \( f \in C^2_y(\overline{U}, \mathbb{R}^n) \). Explicitly we have

\[
\deg(f, U, 0) = \frac{1}{|S^{n-1}|} \int_{\partial U} \sum_{j=1}^n (-1)^{j-1} \frac{f_j}{|f|^n} df_1 \wedge \cdots \wedge df_{j-1} \wedge df_{j+1} \wedge \cdots \wedge df_n.
\]

(14.28)

Since \( \tilde{f} : \partial U \to S^{n-1} \) the integrand can also be written as the pull back \( \tilde{f}^*dS \) of the canonical surface element \( dS \) on \( S^{n-1} \).
This coincides with the boundary value approach for complex functions (note that holomorphic functions are orientation preserving).

Step 2: Admitting continuous functions

Our final step is to remove the condition $f \in C^2$. As before we want the degree to be constant in each ball contained in $D_y(\mathcal{U}, \mathbb{R}^n)$. For example, fix $f \in D_y(\mathcal{U}, \mathbb{R}^n)$ and set $\rho = \text{dist}(y, f(\partial \mathcal{U})) > 0$. Choose $f^i \in C^2(\mathcal{U}, \mathbb{R}^n)$ such that $|f^i - f| < \rho$, implying $f^i \in D_y(\mathcal{U}, \mathbb{R}^n)$. Then $H(t, x) = (1 - t)f^1(x) + tf^2(x) \in D_y(\mathcal{U}, \mathbb{R}^n) \cap C^2(U, \mathbb{R}^n)$, $t \in [0, 1]$, and $|H(t) - f| < \rho$. If we can show that $\deg(H(t), U, y)$ is locally constant with respect to $t$, then it is continuous with respect to $t$ and hence constant (since $[0, 1]$ is connected).

Consequently we can define

$$\deg(f, U, y) = \deg(\tilde{f}, U, y), \quad f \in D_y(\mathcal{U}, \mathbb{R}^n),$$

where $\tilde{f} \in C^2(\mathcal{U}, \mathbb{R}^n)$ with $|\tilde{f} - f| < \text{dist}(y, f(\partial \mathcal{U}))$.

It remains to show that $t \mapsto \deg(H(t), U, y)$ is locally constant.

**Lemma 14.10.** Suppose $f \in C_y^2(\mathcal{U}, \mathbb{R}^n)$. Then for each $\tilde{f} \in C^2(\mathcal{U}, \mathbb{R}^n)$ there is an $\varepsilon > 0$ such that $\deg(f + t \tilde{f}, U, y) = \deg(f, U, y)$ for all $t \in (-\varepsilon, \varepsilon)$.

**Proof.** If $f^{-1}(y) = \emptyset$ the same is true for $f + t \tilde{f}$ if $|t| < \text{dist}(y, f(\overline{\mathcal{U}}))/|g|$. Hence we can exclude this case. For the remaining case we use our usual strategy of considering $y \in \text{RV}(f)$ first and then approximating general $y$ by regular ones.

Suppose $y \in \text{RV}(f)$ and let $f^{-1}(y) = \{x^i\}_{j=1}^N$. By the implicit function theorem we can find disjoint neighborhoods $U(x^i)$ such that there exists a unique solution $x^i(t) \in U(x^i)$ of $(f + t g)(x) = y$ for $|t| < \varepsilon_1$. By reducing $U(x^i)$ if necessary, we can even assume that the sign of $J_{f + t g}$ is constant on $U(x^i)$. Finally, let $\varepsilon_2 = \text{dist}(y, f(U \setminus \bigcup_{i=1}^N U(x^i)))/|g|$. Then $|f + t g - y| > 0$ for $|t| < \varepsilon_2$ and $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$ is the quantity we are looking for.

It remains to consider the case $y \in \text{CV}(f)$. Pick a regular value $\tilde{y} \in B_{\rho/3}(y)$, where $\rho = \text{dist}(y, f(\partial U))$, implying $\deg(f, U, y) = \deg(f, U, \tilde{y})$. Then we can find an $\varepsilon > 0$ such that $\deg(f, U, \tilde{y}) = \deg(f + t g, U, \tilde{y})$ for $|t| < \varepsilon$. Setting $\varepsilon = \min(\varepsilon, \rho/(3|g|))$ we infer $\tilde{y} - (f + t g)(x) \geq \rho/3$ for $x \in \partial U$, that is $|\tilde{y} - y| < \text{dist}(\tilde{y}, (f + t g)(\partial U))$, and thus $\deg(f + t g, U, \tilde{y}) = \deg(f + t g, U, y)$. Putting it all together implies $\deg(f, U, y) = \deg(f + t g, U, y)$ for $|t| < \varepsilon$ as required. \quad \square

Now we can finally prove our main theorem.

**Theorem 14.11.** There is a unique degree $\deg$ satisfying (D1)-(D4). Moreover, $\deg(., U, y) : D_y(\mathcal{U}, \mathbb{R}^n) \rightarrow \mathbb{Z}$ is constant on each component and given
\[ f \in D_y(\overline{U}, \mathbb{R}^n) \] we have
\[
\deg(f, U, y) = \sum_{x \in f^{-1}(y)} \text{sign } J_f(x)
\] (14.30)
where \( \tilde{f} \in D_y^2(\overline{U}, \mathbb{R}^n) \) is in the same component of \( D_y(\overline{U}, \mathbb{R}^n) \), say \(|f - \tilde{f}| < \text{dist}(y, f(\partial U))\), such that \( y \in \text{RV}(\tilde{f}) \).

**Proof.** Our previous considerations show that \( \deg \) is well-defined and locally constant with respect to the first argument by construction. Hence \( \deg(., U, y) : D_y(\overline{U}, \mathbb{R}^n) \to \mathbb{Z} \) is continuous and thus necessarily constant on components since \( \mathbb{Z} \) is discrete. (D2) is clear and (D1) is satisfied since it holds for \( \tilde{f} \) by construction. Similarly, taking \( U_{1,2} \) as in (D3) we can require \(|f - \tilde{f}| < \text{dist}(y, f(\partial U))\). Then (D3) is satisfied since it also holds for \( \tilde{f} \) by construction. Finally, (D4) is a consequence of continuity. \( \square \)

To conclude this section, let us give a few simple examples illustrating the use of the Brouwer degree.

First, let’s investigate the zeros of
\[
f(x_1, x_2) = (x_1 - 2x_2 + \cos(x_1 + x_2), x_2 + 2x_1 + \sin(x_1 + x_2)).
\] (14.31)
Denote the linear part by
\[
g(x_1, x_2) = (x_1 - 2x_2, x_2 + 2x_1).
\] (14.32)
Then we have \(|g(x)| = \sqrt{5}|x|\) and \(|f(x) - g(x)| = 1\) and hence \( h(t) = (1-t)g + tf = g + t(f - g) \) satisfies \(|h(t)| \geq |g| - t|f - g| > 0\) for \(|x| > 1/\sqrt{5}\) implying
\[
\deg(f, B_r(0), 0) = \deg(g, B_r(0), 0) = 1, \quad r > 1/\sqrt{5}.
\] (14.33)
Moreover, since \( J_f(x) = 5 + 3 \cos(x_1 + x_2) + \sin(x_1 + x_2) > 1\) the determinant formula (14.12) for the degree implies that \( f(x) = 0 \) has a unique solution in \( \mathbb{R}^2 \). This solution even has to lie on the circle \(|x| = 1/\sqrt{5}\) since \( f(x) = 0 \) implies \( 1 = |f(x) - g(x)| = |g(x)| = \sqrt{5}|x|\).

Next let us prove the following result which implies the hairy ball (or hedgehog) theorem.

**Theorem 14.12.** Suppose \( U \) contains the origin and let \( f : \partial U \to \mathbb{R}^n \setminus \{0\} \) be continuous. If \( n \) is odd, then there exists a \( x \in \partial U \) and a \( \lambda \neq 0 \) such that \( f(x) = \lambda x \).

**Proof.** By Theorem 14.15 we can assume \( f \in C(\overline{U}, \mathbb{R}^n) \) and since \( n \) is odd we have \( \deg(-I, U, 0) = -1 \). Now if \( \deg(f, U, 0) \neq -1 \), then \( H(t, x) = (1-t)f(x) - tx \) must have a zero \((t_0, x_0) \in (0, 1) \times \partial U \) and hence \( f(x_0) = \frac{t_0}{1-t_0}x_0 \). Otherwise, if \( \deg(f, U, 0) = -1 \) we can apply the same argument to \( H(t, x) = (1-t)f(x) + tx \). \( \square \)
In particular, this result implies that a continuous tangent vector field on the unit sphere \( f : S^{n-1} \to \mathbb{R}^n \) (with \( f(x)x = 0 \) for all \( x \in S^n \)) must vanish somewhere if \( n \) is odd. Or, for \( n = 3 \), you cannot smoothly comb a hedgehog without leaving a bald spot or making a parting. It is however possible to comb the hair smoothly on a torus and that is why the magnetic containers in nuclear fusion are toroidal.

Another simple consequence is the fact that a vector field on \( \mathbb{R}^n \), which points outwards (or inwards) on a sphere, must vanish somewhere inside the sphere.

**Theorem 14.13.** Suppose \( f : \overline{B_R(0)} \to \mathbb{R}^n \) is continuous and satisfies
\[
f(x)x > 0, \quad |x| = R.
\] Then \( f(x) \) vanishes somewhere inside \( B_R(0) \).

**Proof.** If \( f \) does not vanish, then \( H(t, x) = (1 - t)x + tf(x) \) must vanish at some point \( (t_0, x_0) \in (0, 1) \times \partial B_R(0) \) and thus
\[
0 = H(t_0, x_0)x_0 = (1 - t_0)R^2 + t_0f(x_0)x_0.
\] But the last part is positive by assumption, a contradiction. \( \square \)

### 14.4. The Brouwer fixed-point theorem

Now we can show that the famous Brouwer fixed-point theorem is a simple consequence of the properties of our degree.

**Theorem 14.14** (Brouwer fixed point). Let \( K \) be a topological space homeomorphic to a compact, convex subset of \( \mathbb{R}^n \) and let \( f \in C(K, K) \), then \( f \) has at least one fixed point.

**Proof.** Clearly we can assume \( K \subset \mathbb{R}^n \) since homeomorphisms preserve fixed points. Now lets assume \( K = \overline{B_r(0)} \). If there is a fixed-point on the boundary \( \partial B_r(0) \) we are done. Otherwise \( H(t, x) = x - tf(x) \) satisfies \( 0 \not\in H(t, \partial B_r(0)) \) since \( |H(t, x)| \geq |x| - t|f(x)| \geq (1 - t)r > 0, 0 \leq t < 1 \). And the claim follows from \( \deg(x - f(x), B_r(0), 0) = \deg(x, B_r(0), 0) = 1 \).

Now let \( K \) be convex. Then \( K \subset B_r(0) \) and, by Theorem 14.15 below, we can find a continuous retraction \( R : \mathbb{R}^n \to K \) (i.e., \( R(x) = x \) for \( x \in K \)) and consider \( \tilde{f} = f \circ R \in C(\overline{B_r(0)}, \overline{B_r(0)}) \). By our previous analysis, there is a fixed point \( x = \tilde{f}(x) \in \text{hull}(f(K)) \subseteq K \). \( \square \)

Note that any compact, convex subset of a finite dimensional Banach space (complex or real) is isomorphic to a compact, convex subset of \( \mathbb{R}^n \) since linear transformations preserve both properties. In addition, observe that all assumptions are needed. For example, the map \( f : \mathbb{R} \to \mathbb{R}, \ x \mapsto x + 1 \), has no fixed point (\( \mathbb{R} \) is homeomorphic to a bounded set but not to a compact one).
The same is true for the map $f : \partial B_1(0) \to \partial B_1(0)$, $x \mapsto -x$ ($\partial B_1(0) \subset \mathbb{R}^n$ is simply connected for $n \geq 3$ but not homeomorphic to a convex set).

It remains to prove the result from topology needed in the proof of the Brouwer fixed-point theorem. It is a variant of the Tietze extension theorem.

**Theorem 14.15.** Let $X$ be a metric space, $Y$ a Banach space and let $K$ be a closed subset of $X$. Then $F \in C(K, Y)$ has a continuous extension $F \in C(X, Y)$ such that $F(X) \subseteq \text{hull}(F(K))$.

**Proof.** Consider the open cover $\{B_{\rho(x)}(x)\}_{x \in X \setminus K}$ for $X \setminus K$, where $\rho(x) = \text{dist}(x, K)/2$. Choose a (locally finite) partition of unity $\{\phi_\lambda\}_{\lambda \in \Lambda}$ subordinate to this cover and set

$$F(x) = \sum_{\lambda \in \Lambda} \phi_\lambda(x) F(x_\lambda) \text{ for } x \in X \setminus K,$$

(14.36)

where $x_\lambda \in K$ satisfies $\text{dist}(x_\lambda, \text{supp} \phi_\lambda) \leq 2 \text{dist}(K, \text{supp} \phi_\lambda)$. By construction, $F$ is continuous except for possibly at the boundary of $K$. Fix $x_0 \in \partial K$, $\varepsilon > 0$ and choose $\delta > 0$ such that $|F(x) - F(x_0)| \leq \varepsilon$ for all $x \in K$ with $|x - x_0| < 4\delta$. We will show that $|F(x) - F(x_0)| \leq \varepsilon$ for all $x \in X$ with $|x - x_0| < \delta$. Suppose $x \notin K$, then $|F(x) - F(x_0)| \leq \sum_{\lambda \in \Lambda} \phi_\lambda(x)|F(x_\lambda) - F(x_0)|$. By our construction, $x_\lambda$ should be close to $x$ for all $\lambda$ with $x \in \text{supp} \phi_\lambda$ since $x$ is close to $K$. In fact, if $x \in \text{supp} \phi_\lambda$ we have

$$|x - x_\lambda| \leq \text{dist}(x_\lambda, \text{supp} \phi_\lambda) + d(\text{supp} \phi_\lambda) \leq 2 \text{dist}(K, \text{supp} \phi_\lambda) + d(\text{supp} \phi_\lambda),$$

(14.37)

where $d(\text{supp} \phi_\lambda) = \sup_{x,y \in \text{supp} \phi_\lambda} |x - y|$. Since our partition of unity is subordinate to the cover $\{B_{\rho(x)}(x)\}_{x \in X \setminus K}$ we can find a $\tilde{x} \in X \setminus K$ such that $\text{supp} \phi_\lambda \subset B_{\rho(\tilde{x})}(\tilde{x})$ and hence $d(\text{supp} \phi_\lambda) \leq \rho(\tilde{x}) \leq \text{dist}(K, B_{\rho(\tilde{x})}(\tilde{x})) \leq \text{dist}(K, \text{supp} \phi_\lambda)$. Putting it all together implies that we have $|x - x_\lambda| \leq 3 \text{dist}(K, \text{supp} \phi_\lambda) \leq 3|x_0 - x|$ whenever $x \in \text{supp} \phi_\lambda$ and thus

$$|x_0 - x_\lambda| \leq |x_0 - x| + |x - x_\lambda| \leq 4|x_0 - x| \leq 4\delta$$

(14.38)

as expected. By our choice of $\delta$ we have $|F(x_\lambda) - F(x_0)| \leq \varepsilon$ for all $\lambda$ with $\phi_\lambda(x) \neq 0$. Hence $|F(x) - F(x_0)| \leq \varepsilon$ whenever $|x - x_0| \leq \delta$ and we are done. \qed

Let me remark that the Brouwer fixed point theorem is equivalent to the fact that there is no continuous retraction $R : \overline{B}_1(0) \to \partial B_1(0)$ (with $R(x) = x$ for $x \in \partial B_1(0)$) from the unit ball to the unit sphere in $\mathbb{R}^n$.

In fact, if $R$ would be such a retraction, $-R$ would have a fixed point $x_0 \in \partial B_1(0)$ by Brouwer’s theorem. But then $x_0 = -f(x_0) = -x_0$ which is impossible. Conversely, if a continuous function $f : \overline{B}_1(0) \to \overline{B}_1(0)$ has no
fixed point we can define a retraction $R(x) = f(x) + t(x)(x - f(x))$, where $t(x) \geq 0$ is chosen such that $|R(x)|^2 = 1$ (i.e., $R(x)$ lies on the intersection of the line spanned by $x$, $f(x)$ with the unit sphere).

Using this equivalence the Brouwer fixed point theorem can also be derived easily by showing that the homology groups of the unit ball $B_1(0)$ and its boundary (the unit sphere) differ (see, e.g., [24] for details).

Finally, we also derive the following important consequence known as invariance of domain theorem.

**Theorem 14.16 (Brouwer).** Let $U \subseteq \mathbb{R}^n$ be open and let $f : U \to \mathbb{R}^n$ be continuous and injective. Then $f(U)$ is also open.

**Proof.** By scaling and translation it suffices to show that if $f : B_1(0) \to \mathbb{R}^n$ is injective, then $f(0)$ is an inner point for $f(B_1(0))$. Abbreviate $C = B_1(0)$. Since $C$ is compact so is $f(C)$ and thus $f : C \to f(C)$ is a homeomorphism. In particular, $f^{-1} : f(C) \to C$ is continuous and can be extended to a continuous left inverse $g : \mathbb{R}^n \to \mathbb{R}^n$ (i.e., $g(f(x)) = x$ for all $x \in C$.

Note that $g$ has a zero in $f(C)$, namely $f(0)$, which is stable in the sense that any perturbation $\tilde{g} : f(C) \to \mathbb{R}^n$ satisfying $|\tilde{g}(y) - g(y)| \leq 1$ for all $y \in f(C)$ also has a zero. To see this apply the Brower fixed point theorem to the function $F(x) = x - t\tilde{g}(f(x)) = g(f(x)) - t\tilde{g}(f(x))$ which maps $C$ to $C$ by assumption.

Our strategy is to find a contradiction to this fact. Since $g(f(0)) = 0$ vanishes there is some $\varepsilon$ such that $|g(y)| \leq \frac{1}{3}$ for $y \in B_{2\varepsilon}(f(0))$. If $f(0)$ were not in the interior of $f(C)$ we can find some $z \in B_2(f(0))$ which is not in $f(C)$. After a translation we can assume $z = 0$ without loss of generality, that is, $0 \notin f(C)$ and $|f(0)| < \varepsilon$. In particular, we also have $|g(y)| \leq \frac{1}{3}$ for $y \in B_{\varepsilon}(0)$.

Next consider the map $\varphi : f(C) \to \mathbb{R}^n$ given by

$$\varphi(y) = \begin{cases} y, & |y| > \varepsilon, \\ \frac{y}{|y|}, & |y| \leq \varepsilon. \end{cases}$$

It is continuous away from 0 and its range is contained in $\Sigma_1 \cup \Sigma_2$, where $\Sigma_1 = \{y \in f(C) | |y| \geq \varepsilon\}$ and $\Sigma_2 = \{y \in \mathbb{R}^n | |y| = \varepsilon\}$.

Since $f$ is injective, $g$ does not vanish on $\Sigma_1$ and since $\Sigma_1$ is compact there is a $\delta$ such that $|g(y)| \geq \delta$ for $y \in \Sigma_1$. We may even assume $\delta < \frac{1}{3}$.

Next, by the Stone–Weierstraß theorem we can find a polynomial $P$ such that

$$|P(y) - g(y)| < \delta$$

for all $y \in \Sigma$. In particular, $P$ does not vanish on $\Sigma_1$. However, it could vanish on $\Sigma_2$. But since $\Sigma_2$ has measure zero, so has $P(\Sigma_2)$ and we can find
an arbitrarily small value which is not in \( P(\Sigma_2) \). Shifting \( P \) by such a value we can assume that \( P \) does not vanish on \( \Sigma_1 \cup \Sigma_2 \).

Now chose \( \tilde{g} : f(C) \to \mathbb{R}^n \) according to

\[
\tilde{g}(y) = P(\varphi(y)).
\]

Then \( \tilde{g} \) is a continuous function which does not vanish. Moreover, if \( |y| \geq \varepsilon \) we have

\[
|g(y) - \tilde{g}(y)| = |g(y) - P(y)| < \delta < \frac{1}{3}.
\]

And if \( |y| < \varepsilon \) we have \( |g(y)| \leq \frac{1}{3} \) and \( |g(\varphi(y))| \leq \frac{1}{3} \) implying

\[
|g(y) - \tilde{g}(y)| \leq |g(y) - g(\varphi(y))| + |g(\varphi(y)) - P(\varphi(y))| \leq \frac{2}{3} + \delta \leq 1.
\]

Thus \( \tilde{g} \) contradicts our above observation. \( \square \)

An easy consequence worth while noting is the topological invariance of dimension:

**Corollary 14.17.** If \( m < n \) and \( U \) is a nonempty open subset of \( \mathbb{R}^n \), then there is no continuous injective mapping from \( U \) to \( \mathbb{R}^m \).

**Proof.** Suppose there where such a map and extend it to a map from \( U \) to \( \mathbb{R}^n \) by setting the additional coordinates equal to zero. The resulting map contradicts the invariance of domain theorem. \( \square \)

In particular, \( \mathbb{R}^m \) and \( \mathbb{R}^n \) are not homeomorphic for \( m \neq n \).

### 14.5. Kakutani’s fixed-point theorem and applications to game theory

In this section we want to apply Brouwer’s fixed-point theorem to show the existence of Nash equilibria for \( n \)-person games. As a preparation we extend Brouwer’s fixed-point theorem to set valued functions. This generalization will be more suitable for our purpose.

Denote by \( CS(K) \) the set of all nonempty convex subsets of \( K \).

**Theorem 14.18** (Kakutani). Suppose \( K \) is a compact convex subset of \( \mathbb{R}^n \) and \( f : K \to CS(K) \). If the set

\[
\Gamma = \{(x,y) | y \in f(x)\} \subseteq K^2
\]

is closed, then there is a point \( x \in K \) such that \( x \in f(x) \).

**Proof.** Our strategy is to apply Brouwer’s theorem, hence we need a function related to \( f \). For this purpose it is convenient to assume that \( K \) is a simplex

\[
K = \langle v_1, \ldots, v_m \rangle, \quad m \leq n,
\]

(14.40)
14.5. Kakutani’s fixed-point theorem and applications to game theory

where \( v_i \) are the vertices. If we pick \( y_i \in f(v_i) \) we could set

\[
f^1(x) = \sum_{i=1}^{m} \lambda_i y_i,
\]

(14.41)

where \( \lambda_i \) are the barycentric coordinates of \( x \) (i.e., \( \lambda_i \geq 0 \), \( \sum_{i=1}^{m} \lambda_i = 1 \) and \( x = \sum_{i=1}^{m} \lambda_i v_i \)). By construction, \( f^1 \in C(K,K) \) and there is a fixed point \( x^1 \). But unless \( x^1 \) is one of the vertices, this doesn’t help us too much. So lets choose a better function as follows. Consider the \( k \)-th barycentric subdivision and for each vertex \( v_i \) in this subdivision pick an element \( y_i \in f(v_i) \).

Now define

\[
x_k = \sum_{i=1}^{m} \lambda_i^k v_i = \sum_{i=1}^{m} \lambda_i^k y_i^k, \quad y_i^k = f^k(v_i^k),
\]

(14.42)

in the subsimplex \( \langle v_1^k, \ldots, v_m^k \rangle \). Since \( (x^k, \lambda_1^0, \ldots, \lambda_m^0, y_1^0, \ldots, y_m^0) \in K \times [0,1]^m \times K^m \) we can assume that this sequence converges to some limit \( (x^0, \lambda_1^0, \ldots, \lambda_m^0, y_1^0, \ldots, y_m^0) \) after passing to a subsequence. Since the sub-simplices shrink to a point, this implies \( y_i^k \to y_i^0 \) and hence \( y_i^0 \in f(x^0) \) since \( (v_i^k, y_i^k) \in \Gamma \to (v_i^0, y_i^0) \in \Gamma \) by the closedness assumption. Now (14.42) tells us

\[
x^0 = \sum_{i=1}^{m} \lambda_i^0 y_i^0 \in f(x^0)
\]

(14.43)

since \( f(x^0) \) is convex and the claim holds if \( K \) is a simplex.

If \( K \) is not a simplex, we can pick a simplex \( S \) containing \( K \) and proceed as in the proof of the Brouwer theorem.

If \( f(x) \) contains precisely one point for all \( x \), then Kakutani’s theorem reduces to the Brouwer’s theorem.

Now we want to see how this applies to game theory.

An \( n \)-person game consists of \( n \) players who have \( m_i \) possible actions to choose from. The set of all possible actions for the \( i \)-th player will be denoted by \( \Phi_i = \{1, \ldots, m_i\} \). An element \( \varphi_i \in \Phi_i \) is also called a pure strategy for reasons to become clear in a moment. Once all players have chosen their move \( \varphi_i \), the payoff for each player is given by the payoff function

\[
R_i(\varphi), \quad \varphi = (\varphi_1, \ldots, \varphi_n) \in \Phi = \prod_{i=1}^{n} \Phi_i
\]

(14.44)

of the \( i \)-th player. We will consider the case where the game is repeated a large number of times and where in each step the players choose their action according to a fixed strategy. Here a strategy \( s_i \) for the \( i \)-th player is a probability distribution on \( \Phi_i \), that is, \( s_i = (s_{i1}, \ldots, s_{im_i}) \) such that \( s_{ik} \geq 0 \)
and \( \sum_{k=1}^{m_i} s_i^k = 1 \). The set of all possible strategies for the \( i \)-th player is denoted by \( S_i \). The number \( s_i^k \) is the probability for the \( k \)-th pure strategy to be chosen. Consequently, if \( s = (s_1, \ldots, s_n) \in S = \prod_{i=1}^{n} S_i \) is a collection of strategies, then the probability that a given collection of pure strategies gets chosen is

\[
s(\varphi) = \prod_{i=1}^{n} s_i(\varphi), \quad s_i(\varphi) = s_i^k, \quad \varphi = (k_1, \ldots, k_n) \in \Phi \quad (14.45)
\]

(assuming all players make their choice independently) and the expected payoff for player \( i \) is

\[
R_i(s) = \sum_{\varphi \in \Phi} s(\varphi)R_i(\varphi). \quad (14.46)
\]

By construction, \( R_i(s) \) is continuous.

The question is of course, what is an optimal strategy for a player? If the other strategies are known, a **best reply** of player \( i \) against \( s \) would be a strategy \( \tilde{s}_i \) satisfying

\[
R_i(s\setminus \tilde{s}_i) = \max_{\tilde{s}_i \in \tilde{S}_i} R_i(s\setminus \tilde{s}_i) \quad (14.47)
\]

Here \( s\setminus \tilde{s}_i \) denotes the strategy combination obtained from \( s \) by replacing \( s_i \) by \( \tilde{s}_i \). The set of all best replies against \( s \) for the \( i \)-th player is denoted by \( B_i(s) \). Explicitly, \( \pi_i \in B(s) \) if and only if \( \pi_i = 0 \) whenever \( R_i(s\setminus k) < \max_{1 \leq l \leq m_i} R_i(s\setminus l) \) (in particular \( B_i(s) \neq \emptyset \)).

Let \( s, \pi \in S \), we call \( \pi \) a best reply against \( s \) if \( \pi_i \) is a best reply against \( s \) for all \( i \). The set of all best replies against \( s \) is \( B(s) = \prod_{i=1}^{n} B_i(s) \).

A strategy combination \( \pi \in S \) is a **Nash equilibrium** for the game if it is a best reply against itself, that is,

\[
\pi \in B(\pi). \quad (14.48)
\]

Or, put differently, \( \pi \) is a Nash equilibrium if no player can increase his payoff by changing his strategy as long as all others stick to their respective strategies. In addition, if a player sticks to his equilibrium strategy, he is assured that his payoff will not decrease no matter what the others do.

To illustrate these concepts, let us consider the famous **prisoners dilemma**. Here we have two players which can choose to defect or cooperate. The payoff is symmetric for both players and given by the following diagram

\[
\begin{array}{c|cc}
R_1 & d_2 & c_2 \\
\hline
d_1 & 0 & 2 \\
c_1 & -1 & 1 \\
\end{array} \quad \begin{array}{c|cc}
R_2 & d_2 & c_2 \\
\hline
d_1 & 0 & -1 \\
c_1 & 2 & 1 \\
\end{array}
\quad (14.49)
\]

where \( c_i \) or \( d_i \) means that player \( i \) cooperates or defects, respectively. It is easy to see that the (pure) strategy pair \( (d_1, d_2) \) is the only Nash equilibrium for this game and that the expected payoff is 0 for both players. Of course,
both players could get the payoff 1 if they both agree to cooperate. But if one would break this agreement in order to increase his payoff, the other one would get less. Hence it might be safer to defect.

Now that we have seen that Nash equilibria are a useful concept, we want to know when such an equilibrium exists. Luckily we have the following result.

**Theorem 14.19 (Nash).** Every n-person game has at least one Nash equilibrium.

**Proof.** The definition of a Nash equilibrium begs us to apply Kakutani’s theorem to the set valued function \( s \mapsto B(s) \). First of all, \( S \) is compact and convex and so are the sets \( B(s) \). Next, observe that the closedness condition of Kakutani’s theorem is satisfied since if \( s_m \in S \) and \( \tilde{s}_m \in B(s_m) \) both converge to \( s \) and \( \tilde{s} \), respectively, then (14.47) for \( s_m, \tilde{s}_m \)

\[
R_i(s_m \setminus \tilde{s}_i) \leq R_i(s_m \setminus \tilde{s}_i), \quad \tilde{s}_i \in S_i, \ 1 \leq i \leq n,
\]

implies (14.47) for the limits \( s, \tilde{s} \)

\[
R_i(s \setminus \tilde{s}_i) \leq R_i(s \setminus \tilde{s}_i), \quad \tilde{s}_i \in S_i, \ 1 \leq i \leq n,
\]

by continuity of \( R_i(s) \). \( \square \)

### 14.6. Further properties of the degree

We now prove some additional properties of the mapping degree. The first one will relate the degree in \( \mathbb{R}^n \) with the degree in \( \mathbb{R}^m \). It will be needed later on to extend the definition of degree to infinite dimensional spaces. By virtue of the canonical embedding \( \mathbb{R}^m \hookrightarrow \mathbb{R}^m \times \{0\} \subset \mathbb{R}^n \) we can consider \( \mathbb{R}^m \) as a subspace of \( \mathbb{R}^n \).

**Theorem 14.20 (Reduction property).** Let \( f \in C(U, \mathbb{R}^m) \) and \( y \in \mathbb{R}^m \setminus (\mathbb{I} + f)(\partial U) \), then

\[
\deg(\mathbb{I} + f, U, y) = \deg(\mathbb{I} + f_m, U_m, y), \quad (14.52)
\]

where \( f_m = f|_{U_m} \), where \( U_m \) is the projection of \( U \) to \( \mathbb{R}^m \).

**Proof.** Choose a \( \tilde{f} \in C^2(U, \mathbb{R}^m) \) sufficiently close to \( f \) such that \( y \in RV(\tilde{f}) \). Let \( x \in (\mathbb{I} + \tilde{f})^{-1}(y) \), then \( x = y - f(x) \in \mathbb{R}^m \) implies \( (\mathbb{I} + f)^{-1}(y) = (\mathbb{I} + \tilde{f})^{-1}(y) \). Moreover,

\[
J_{\mathbb{I} + f}(x) = \det(\mathbb{I} + \tilde{f}')(x) = \det \begin{pmatrix} \delta_{ij} + \partial_j \tilde{f}_i(x) & \partial_j \tilde{f}_j(x) \\
0 & \delta_{ij} \end{pmatrix} = \det(\delta_{ij} + \partial_j \tilde{f}_i) = J_{\mathbb{I} + \tilde{f}}(x)
\]

shows \( \deg(\mathbb{I} + f, U, y) = \deg(\mathbb{I} + \tilde{f}, U, y) = \deg(\mathbb{I} + \tilde{f}_m, U_m, y) = \deg(\mathbb{I} + f_m, U_m, y) \) as desired. \( \square \)
Let $U \subseteq \mathbb{R}^n$ and $f \in C(\overline{U}, \mathbb{R}^n)$ be as usual. By Theorem 14.2 we know that $\deg(f, U, y)$ is the same for every $y$ in a connected component of $\mathbb{R}^n \setminus f(\partial U)$. We will denote these components by $K_j$ and write $\deg(f, U, y) = \deg(f, U, K_j)$ if $y \in K_j$.

**Theorem 14.21** (Product formula). Let $U \subseteq \mathbb{R}^n$ be a bounded and open set and denote by $G_j$ the connected components of $\mathbb{R}^n \setminus f(\partial U)$. If $g \circ f \in D_y(U, \mathbb{R}^n)$, then

$$\deg(g \circ f, U, y) = \sum_j \deg(f, U, G_j) \deg(g, G_j, y),$$

(14.54)

where only finitely many terms in the sum are nonzero.

**Proof.** Since $f(\overline{U})$ is is compact, we can find an $r > 0$ such that $f(\overline{U}) \subseteq B_r(0)$. Moreover, since $g^{-1}(y)$ is closed, $g^{-1}(y) \cap B_r(0)$ is compact and hence can be covered by finitely many components $\{G_j\}_{j=1}^m$. In particular, the others will have $\deg(f, G_k, y) = 0$ and hence only finitely many terms in the above sum are nonzero.

We begin by computing $\deg(g \circ f, U, y)$ in the case where $f, g \in C^1$ and $y \notin CV(g \circ f)$. Since $d(g \circ f)(x) = g'(f(x)) df(x)$ the claim is a straightforward calculation

$$\deg(g \circ f, U, y) = \sum_{x \in (g \circ f)^{-1}(y)} \text{sign}(J_{g \circ f}(x))$$

$$= \sum_{x \in (g \circ f)^{-1}(y)} \text{sign}(J_g(f(x))) \text{sign}(J_f(x))$$

$$= \sum_{z \in g^{-1}(y)} \text{sign}(J_g(z)) \sum_{x \in f^{-1}(z)} \text{sign}(J_f(x))$$

$$= \sum_{z \in g^{-1}(y)} \text{sign}(J_g(z)) \deg(f, U, z)$$

and, using our cover $\{G_j\}_{j=1}^m$,

$$\deg(g \circ f, U, y) = \sum_{j=1}^m \sum_{z \in g^{-1}(y) \cap G_j} \text{sign}(J_g(z)) \deg(f, U, z)$$

$$= \sum_{j=1}^m \deg(f, U, G_j) \sum_{z \in g^{-1}(y) \cap G_j} \text{sign}(J_g(z))$$

(14.55)

$$= \sum_{j=1}^m \deg(f, U, G_j) \deg(g, G_j, y).$$

(14.56)

Moreover, this formula still holds for $y \in CV(g \circ f)$ and for $g \in C$ by construction of the Brouwer degree. However, the case $f \in C$ will need a
closer investigation since the sets $G_j$ depend on $f$. To overcome this problem we will introduce the sets

\[ L_l = \{ z \in \mathbb{R}^n \setminus f(\partial U) | \deg(f, U, z) = l \}. \tag{14.57} \]

Observe that $L_l$, $l > 0$, must be a union of some sets of \{ $G_j$ \}_{j=1}^m.

Now choose $\tilde{f} \in C^1$ such that $|f(x) - \tilde{f}(x)| < 2^{-1} \text{dist}(g^{-1}(y), f(\partial U))$ for $x \in U$ and define $\tilde{K}_j$, $\tilde{L}_l$ accordingly. Then we have $U_l \cap g^{-1}(y) = \tilde{U}_l \cap g^{-1}(y)$ by Theorem 14.1 (iii). Moreover,

\[
\deg(g \circ f, U, y) = \deg(g \circ \tilde{f}, U, y) = \sum_j \deg(\tilde{f}, U, \tilde{K}_j) \deg(g, \tilde{K}_j, y)
\]

\[
= \sum_{l>0} l \deg(g, \tilde{U}_l, y) = \sum_{l>0} l \deg(g, U_l, y)
\]

\[
= \sum_j \deg(f, U, G_j) \deg(g, G_j, y) \tag{14.58}
\]

which proves the claim. \[ \square \]

14.7. The Jordan curve theorem

In this section we want to show how the product formula (14.54) for the Brouwer degree can be used to prove the famous **Jordan curve theorem** which states that a homeomorphic image of the circle dissects $\mathbb{R}^2$ into two components (which necessarily have the image of the circle as common boundary). In fact, we will even prove a slightly more general result.

**Theorem 14.22.** Let $C_j \subset \mathbb{R}^n$, $j = 1, 2$, be homeomorphic compact sets. Then $\mathbb{R}^n \setminus C_1$ and $\mathbb{R}^n \setminus C_2$ have the same number of connected components.

**Proof.** Denote the components of $\mathbb{R}^n \setminus C_1$ by $H_j$ and those of $\mathbb{R}^n \setminus C_2$ by $K_j$. Let $h : C_1 \to C_2$ be a homeomorphism with inverse $k : C_2 \to C_1$. By Theorem 14.15 we can extend both to $\mathbb{R}^n$. Then Theorem 14.1 (iii) and the product formula imply

\[ 1 = \deg(k \circ h, H_j, y) = \sum_l \deg(h, H_j, G_l) \deg(k, G_l, y) \tag{14.59} \]

for any $y \in H_j$. Now we have

\[ \bigcup_i K_i = \mathbb{R}^n \setminus C_2 \subseteq \mathbb{R}^n \setminus h(\partial H_j) \subseteq \bigcup_l G_l \tag{14.60} \]

and hence for every $i$ we have $K_i \subseteq G_l$ for some $l$ since components are maximal connected sets. Let $N_l = \{ i | K_i \subseteq G_l \}$ and observe that we have
deg(k, G_l, y) = \sum_{i \in N_l} deg(k, K_i, y) and deg(h, H_j, G_l) = deg(h, H_j, K_i) for every \( i \in N_l \). Therefore,

\[ 1 = \sum_l \sum_{i \in N_l} \text{deg}(h, H_j, K_i) \text{deg}(k, K_i, y) = \sum_i \text{deg}(h, H_j, K_i) \text{deg}(k, K_i, H_j) \]  
\hspace{10cm} (14.61)

By reversing the role of \( C_1 \) and \( C_2 \), the same formula holds with \( H_j \) and \( K_i \) interchanged.

Hence

\[ \sum_i 1 = \sum_i \sum_j \text{deg}(h, H_j, K_i) \text{deg}(k, K_i, H_j) = \sum_j 1 \]  
\hspace{10cm} (14.62)

shows that if either the number of components of \( \mathbb{R}^n \setminus C_1 \) or the number of components of \( \mathbb{R}^n \setminus C_2 \) is finite, then so is the other and both are equal. Otherwise there is nothing to prove. \( \square \)
Chapter 15

The Leray–Schauder mapping degree

15.1. The mapping degree on finite dimensional Banach spaces

The objective of this section is to extend the mapping degree from $\mathbb{R}^n$ to general Banach spaces. Naturally, we will first consider the finite dimensional case.

Let $X$ be a (real) Banach space of dimension $n$ and let $\phi$ be any isomorphism between $X$ and $\mathbb{R}^n$. Then, for $f \in D_y(U, X)$, $U \subset X$ open, $y \in X$, we can define

$$\text{deg}(f, U, y) = \text{deg}(\phi \circ f \circ \phi^{-1}, \phi(U), \phi(y))$$

(15.1)

provided this definition is independent of the basis chosen. To see this let $\psi$ be a second isomorphism. Then $A = \psi \circ \phi^{-1} \in \text{GL}(n)$. Abbreviate $f^* = \phi \circ f \circ \phi^{-1}$, $y^* = \phi(y)$ and pick $\tilde{f}^* \in C^1_y(\phi(U), \mathbb{R}^n)$ in the same component of $D_y(\phi(U), \mathbb{R}^n)$ as $f^*$ such that $y^* \in \text{RV}(\tilde{f}^*)$. Then $A \circ \tilde{f}^* \circ A^{-1} \in C^1_y(\psi(U), \mathbb{R}^n)$ is the same component of $D_y(\psi(U), \mathbb{R}^n)$ as $A \circ f^* \circ A^{-1} = \psi \circ f \circ \psi^{-1}$ (since $A$ is also a homeomorphism) and

$$J_{A \circ \tilde{f}^* \circ A^{-1}}(A y^*) = \det(A) J_{\tilde{f}^*}(y^*) \det(A^{-1}) = J_{\tilde{f}^*}(y^*)$$

(15.2)

by the chain rule. Thus we have $\text{deg}(\psi \circ f \circ \psi^{-1}, \psi(U), \psi(y)) = \text{deg}(\phi \circ f \circ \phi^{-1}, \phi(U), \phi(y))$ and our definition is independent of the basis chosen. In addition, it inherits all properties from the mapping degree in $\mathbb{R}^n$. Note also that the reduction property holds if $\mathbb{R}^m$ is replaced by an arbitrary subspace $X_1$ since we can always choose $\phi : X \to \mathbb{R}^n$ such that $\phi(X_1) = \mathbb{R}^m$. 

319
Our next aim is to tackle the infinite dimensional case. The general idea is to approximate $F$ by finite dimensional maps (in the same spirit as we approximated continuous $f$ by smooth functions). To do this we need to know which maps can be approximated by finite dimensional operators. Hence we have to recall some basic facts first.

15.2. Compact maps

Let $X, Y$ be Banach spaces and $U \subset X$. A map $F : U \subset X \rightarrow Y$ is called **finite dimensional** if its range is finite dimensional. In addition, it is called **compact** if it is continuous and maps bounded sets into relatively compact ones. The set of all compact maps is denoted by $C(U,Y)$ and the set of all compact, finite dimensional maps is denoted by $F(U,Y)$. Both sets are normed linear spaces and we have $F(U,Y) \subseteq C(U,Y) \subseteq C_b(U,Y)$ (recall that compact sets are automatically bounded).

If $U$ is compact, then $C(U,Y) = C(U,Y)$ (since the continuous image of a compact set is compact) and if $\dim(Y) < \infty$, then $F(U,Y) = C(U,Y)$. In particular, if $U \subset \mathbb{R}^n$ is bounded, then $F(\overline{U}, \mathbb{R}^n) = C(\overline{U}, \mathbb{R}^n) = C(\overline{U}, \mathbb{R}^n)$.

Now let us collect some results needed in the sequel.

**Lemma 15.1.** If $K \subset X$ is compact, then for every $\varepsilon > 0$ there is a finite dimensional subspace $X_\varepsilon \subseteq X$ and a continuous map $P_\varepsilon : K \rightarrow X_\varepsilon$ such that $|P_\varepsilon(x) - x| \leq \varepsilon$ for all $x \in K$.

**Proof.** Pick $\{x_i\}_{i=1}^n \subseteq K$ such that $\bigcup_{i=1}^n B_\varepsilon(x_i)$ covers $K$. Let $\{\phi_i\}_{i=1}^n$ be a partition of unity (restricted to $K$) subordinate to $\{B_\varepsilon(x_i)\}_{i=1}^n$, that is, $\phi_i \in C(K, [0, 1])$ with $\text{supp}(\phi_i) \subset B_\varepsilon(x_i)$ and $\sum_{i=1}^n \phi_i(x) = 1$, $x \in K$. Set

$$P_\varepsilon(x) = \sum_{i=1}^n \phi_i(x) x_i,$$  \hspace{1cm} (15.3)

then

$$|P_\varepsilon(x) - x| = \left| \sum_{i=1}^n \phi_i(x) - \sum_{i=1}^n \phi_i(x_i) x_i \right| \leq \sum_{i=1}^n \phi_i(x) |x - x_i| \leq \varepsilon. \hspace{1cm} (15.4)$$

This lemma enables us to prove the following important result.

**Theorem 15.2.** Let $U$ be bounded, then the closure of $F(U,Y)$ in $C(U,Y)$ is $C(U,Y)$.

**Proof.** Suppose $F_N \in C(U,Y)$ converges to $F$. If $F \notin C(U,Y)$ then we can find a sequence $x_n \in U$ such that $|F(x_n) - F(x_m)| \geq \rho > 0$ for $n \neq m$. If $N$
15.3. The Leray–Schauder mapping degree

is so large that $|F - F_N| \leq \rho/4$, then

$$|F_N(x_n) - F_N(x_m)| \geq |F(x_n) - F(x_m)| - |F_N(x_n) - F(x_n)|$$

$$\geq \rho - 2\frac{\rho}{4} = \frac{\rho}{2}$$ (15.5)

This contradiction shows $\overline{F(U, y)} \subseteq \mathcal{C}(U, Y)$. Conversely, let $K = \overline{F(U)}$ and choose $P_\varepsilon$ according to Lemma 15.1, then $F_\varepsilon = P_\varepsilon \circ F \in \mathcal{F}(U, Y)$ converges to $F$. Hence $\mathcal{C}(U, Y) \subseteq \overline{F(U, y)}$ and we are done. \qed

Finally, let us show some interesting properties of mappings $\mathbb{I} + F$, where $F \in \mathcal{C}(U, Y)$.

**Lemma 15.3.** Let $U$ be bounded and closed. Suppose $F \in \mathcal{C}(U, X)$, then $\mathbb{I} + F$ is proper (i.e., inverse images of compact sets are compact) and maps closed subsets to closed subsets.

**Proof.** Let $A \subseteq U$ be closed and $y_n = (\mathbb{I} + F)(x_n) \in (\mathbb{I} + F)(A)$ converges to some point $y$. Since $y_n - x_n = F(x_n) \in F(U)$ we can assume that $y_n - x_n \to z$ after passing to a subsequence and hence $x_n \to x = y - z \in A$. Since $y = x + F(x) \in (\mathbb{I} + F)(A)$, $(\mathbb{I} + F)(A)$ is closed.

Next, let $U$ be closed and $K \subseteq Y$ be compact. Let $\{x_n\} \subseteq (\mathbb{I} + F)^{-1}(K)$. Then we can pass to a subsequence $y_{n_m} = x_{n_m} + F(x_{n_m})$ such that $y_{n_m} \to y$. As before this implies $x_{n_m} \to x$ and thus $(\mathbb{I} + F)^{-1}(K)$ is compact. \qed

Now we are all set for the definition of the Leray–Schauder degree, that is, for the extension of our degree to infinite dimensional Banach spaces.

### 15.3. The Leray–Schauder mapping degree

For $U \subset X$ we set

$$\mathcal{D}_y(\overline{U}, X) = \{F \in \mathcal{C}(\overline{U}, X) | y \notin (\mathbb{I} + F)(\partial U)\}$$ (15.6)

and $\mathcal{F}_y(\overline{U}, X) = \{F \in \mathcal{F}(\overline{U}, X) | y \notin (\mathbb{I} + F)(\partial U)\}$. Note that for $F \in \mathcal{D}_y(\overline{U}, X)$ we have $\text{dist}(y, (\mathbb{I} + F)(\partial U)) > 0$ since $\mathbb{I} + F$ maps closed sets to closed sets.

Abbreviate $\rho = \text{dist}(y, (\mathbb{I} + F)(\partial U))$ and pick $F_1 \in \mathcal{F}(\overline{U}, X)$ such that $|F - F_1| < \rho$ implying $F_1 \in \mathcal{F}_y(\overline{U}, X)$. Next, let $X_1$ be a finite dimensional subspace of $X$ such that $F_1(U) \subset X_1$, $y \in X_1$ and set $U_1 = U \cap X_1$. Then we have $F_1 \in \mathcal{F}_y(\overline{U_1}, X_1)$ and might define

$$\deg(\mathbb{I} + F, U, y) = \deg(\mathbb{I} + F_1, U_1, y)$$ (15.7)

provided we show that this definition is independent of $F_1$ and $X_1$ (as above). Pick another map $F_2 \in \mathcal{F}(\overline{U}, X)$ such that $|F - F_2| < \rho$ and let $X_2$ be a
corresponding finite dimensional subspace as above. Consider \( X_0 = X_1 + X_2, U_0 = U \cap X_0 \), then \( F_i \in \mathcal{F}_y(U_0, X_0) \), \( i = 1, 2 \), and
\[
\deg(\mathbb{I} + F_i, U_0, y) = \deg(\mathbb{I} + F_i, U_i, y), \quad i = 1, 2, \tag{15.8}
\]
by the reduction property. Moreover, set \( H(t) = \mathbb{I} + (1 - t)F_1 + t F_2 \) implying \( H(t) \subset U \), \( t \in [0,1] \), since \( |H(t) - (\mathbb{I} + F)| < \rho \) for \( t \in [0,1] \). Hence homotopy invariance
\[
\deg(\mathbb{I} + F_1, U_0, y) = \deg(\mathbb{I} + F_2, U_0, y) \tag{15.9}
\]
shows that (15.7) is independent of \( F_1, X_1 \).

**Theorem 15.4.** Let \( U \) be a bounded open subset of a (real) Banach space \( X \) and let \( F \in \mathcal{D}_y(U, X) \), \( y \in X \). Then the following hold true.

(i). \( \deg(\mathbb{I} + F, U, y) = \deg(\mathbb{I} + F - y, U, 0) \).

(ii). \( \deg(\mathbb{I}, U, y) = 1 \) if \( y \in U \).

(iii). If \( U_{1, 2} \) are open, disjoint subsets of \( U \) such that \( y \not\in f(\mathcal{U} \setminus (U_1 \cup U_2)) \), then \( \deg(\mathbb{I} + F, U, y) = \deg(\mathbb{I} + F, U_1, y) + \deg(\mathbb{I} + F, U_2, y) \).

(iv). If \( H : [0,1] \times \mathcal{U} \to X \) and \( y : [0,1] \to X \) are both continuous such that \( H(t) \in \mathcal{D}_{y(t)}(U, \mathbb{R}^n), t \in [0,1] \), then \( \deg(\mathbb{I} + H(0), U, y(0)) = \deg(\mathbb{I} + H(1), U, y(1)) \).

**Proof.** Except for (iv) all statements follow easily from the definition of the degree and the corresponding property for the degree in finite dimensional spaces. Considering \( H(t, x) - y(t) \), we can assume \( y(t) = 0 \) by (i). Since \( H([0,1], \partial U) \) is compact, we have \( \rho = \text{dist}(y, H([0,1], \partial U)) > 0 \). By Theorem 15.2 we can pick \( H_1 \in \mathcal{F}([0,1] \times U, X) \) such that \( |H(t) - H_1(t)| < \rho \), \( t \in [0,1] \). this implies \( \deg(\mathbb{I} + H(t), U, 0) = \deg(\mathbb{I} + H_1(t), U, 0) \) and the rest follows from Theorem 14.2. \( \square \)

In addition, Theorem 14.1 and Theorem 14.2 hold for the new situation as well (no changes are needed in the proofs).

**Theorem 15.5.** Let \( F, G \in \mathcal{D}_y(U, X) \), then the following statements hold.

(i). \( \text{We have } \deg(\mathbb{I} + F, 0, y) = 0 \). Moreover, if \( U_i, 1 \leq i \leq N \), are disjoint open subsets of \( U \) such that \( y \not\in (\mathbb{I} + F)(\mathcal{U} \setminus \bigcup_{i=1}^N U_i) \), then \( \deg(\mathbb{I} + F, U_i, y) = \sum_{i=1}^N \deg(\mathbb{I} + F, U_i, y) \).

(ii). \( \text{If } y \not\in (\mathbb{I} + F)(U), \text{ then } \deg(\mathbb{I} + F, U, y) = 0 \) (but not the other way round). Equivalently, if \( \deg(\mathbb{I} + F, U, y) \neq 0 \), then \( y \in (\mathbb{I} + F)(U) \).

(iii). \( \text{If } |f(x) - g(x)| < \text{dist}(y, f(\partial U)), x \in \partial U, \text{ then } \deg(f, U, y) = \deg(g, U, y) \). In particular, this is true if \( f(x) = g(x) \) for \( x \in \partial U \).

(iv). \( \deg(\mathbb{I} + ., U, y) \) is constant on each component of \( \mathcal{D}_y(U, X) \).

(v). \( \deg(\mathbb{I} + F, ., ) \) is constant on each component of \( X \setminus f(\partial U) \).
15.4. The Leray–Schauder principle and the Schauder fixed-point theorem

As a first consequence we note the Leray–Schauder principle which says that a priori estimates yield existence.

**Theorem 15.6** (Leray–Schauder principle). Suppose $F \in C(X, X)$ and any solution $x$ of $x = tF(x)$, $t \in [0, 1]$ satisfies the a priori bound $|x| \leq M$ for some $M > 0$, then $F$ has a fixed point.

**Proof.** Pick $\rho > M$ and observe $\deg(I - F, B_\rho(0), 0) = \deg(I, B_\rho(0), 0) = 1$ using the compact homotopy $H(t, x) = -tF(x)$. Here $H(t) \in D_0(B_\rho(0), X)$ due to the a priori bound. \hfill \Box

Now we can extend the Brouwer fixed-point theorem to infinite dimensional spaces as well.

**Theorem 15.7** (Schauder fixed point). Let $K$ be a closed, convex, and bounded subset of a Banach space $X$. If $F \in C(K, K)$, then $F$ has at least one fixed point. The result remains valid if $K$ is only homeomorphic to a closed, convex, and bounded subset.

**Proof.** Since $K$ is bounded, there is a $\rho > 0$ such that $K \subseteq B_\rho(0)$. By Theorem 14.15 we can find a continuous retraction $R : X \to K$ (i.e., $R(x) = x$ for $x \in K$) and consider $\tilde{F} = F \circ R \in C(B_\rho(0), B_\rho(0))$. The compact homotopy $H(t, x) = -t\tilde{F}(x)$ shows that $\deg(I - \tilde{F}, B_\rho(0), 0) = \deg(I, B_\rho(0), 0) = 1$. Hence there is a point $x_0 = \tilde{F}(x_0) \in K$. Since $\tilde{F}(x_0) = F(x_0)$ for $x_0 \in K$ we are done. \hfill \Box

Finally, let us prove another fixed-point theorem which covers several others as special cases.

**Theorem 15.8.** Let $U \subset X$ be open and bounded and let $F \in C(U, X)$. Suppose there is an $x_0 \in U$ such that

\[ F(x) - x_0 \neq \alpha(x - x_0), \quad x \in \partial U, \alpha \in (1, \infty). \quad (15.10) \]

Then $F$ has a fixed point.

**Proof.** Consider $H(t, x) = x - x_0 - t(F(x) - x_0)$, then we have $H(t, x) \neq 0$ for $x \in \partial U$ and $t \in [0, 1]$ by assumption. If $H(1, x) = 0$ for some $x \in \partial U$, then $x$ is a fixed point and we are done. Otherwise we have $\deg(I - F, U, 0) = \deg(I - x_0, U, 0) = \deg(I, U, x_0) = 1$ and hence $F$ has a fixed point. \hfill \Box

Now we come to the anticipated corollaries.

**Corollary 15.9.** Let $U \subset X$ be open and bounded and let $F \in C(U, X)$. Then $F$ has a fixed point if one of the following conditions holds.
15. The Leray–Schauder mapping degree

(i) \( U = B_\rho(0) \) and \( F(\partial U) \subseteq \overline{U} \) (Rothe).

(ii) \( U = B_\rho(0) \) and \( |F(x) - x|^2 \geq |F(x)|^2 - |x|^2 \) for \( x \in \partial U \) (Altman).

(iii) \( X \) is a Hilbert space, \( U = B_\rho(0) \) and \( |F(x), x| \leq |x|^2 \) for \( x \in \partial U \) (Krasnosel’skii).

**Proof.** (1). \( F(\partial U) \subseteq \overline{U} \) and \( F(x) = ax \) for \( |x| = \rho \) implies \( |a|\rho \leq \rho \) and hence (15.10) holds. (2). \( F(x) = ax \) for \( |x| = \rho \) implies \( (\alpha - 1)^2\rho^2 \geq (\alpha^2 - 1)\rho^2 \) and hence \( \alpha \leq 0 \). (3). Special case of (2) since \( |F(x) - x|^2 = |F(x)|^2 - 2 \langle F(x), x \rangle + |x|^2 \). \( \square \)

15.5. Applications to integral and differential equations

In this section we want to show how our results can be applied to integral and differential equations. To be able to apply our results we will need to know that certain integral operators are compact.

**Lemma 15.10.** Suppose \( I = [a, b] \subseteq \mathbb{R} \) and \( f \in C(I \times I \times \mathbb{R}^n, \mathbb{R}^n) \), \( \tau \in C(I, I) \), then

\[
F : C(I, \mathbb{R}^n) \to C(I, \mathbb{R}^n) \quad \quad \quad (15.11)
\]

\[
x(t) \mapsto F(x)(t) = \int_a^{\tau(t)} f(t, s, x(s))ds
\]

is compact.

**Proof.** We first need to prove that \( F \) is continuous. Fix \( x_0 \in C(I, \mathbb{R}^n) \) and \( \varepsilon > 0 \). Set \( \rho = |x_0| + 1 \) and abbreviate \( \overline{B} = B_\rho(0) \subset \mathbb{R}^n \). The function \( f \) is uniformly continuous on \( Q = I \times I \times \overline{B} \) since \( Q \) is compact. Hence for \( \varepsilon_1 = \varepsilon/(b - a) \) we can find a \( \delta \in (0, 1] \) such that \( |f(t, s, x) - f(t, s, y)| \leq \varepsilon_1 \) for \( |x - y| < \delta \). But this implies

\[
|F(x) - F(x_0)| = \sup_{t \in I} \left| \int_a^{\tau(t)} f(t, s, x(s)) - f(t, s, x_0(s))ds \right|
\]

\[
\leq \sup_{t \in I} \int_a^{\tau(t)} |f(t, s, x(s)) - f(t, s, x_0(s))|ds
\]

\[
\leq \sup_{t \in I} (b - a)\varepsilon_1 = \varepsilon,
\]

(15.12)

for \( |x - x_0| < \delta \). In other words, \( F \) is continuous. Next we note that if \( U \subset C(I, \mathbb{R}^n) \) is bounded, say \( |U| < \rho \), then

\[
|F(U)| \leq \sup_{x \in U} \left| \int_a^{\tau(t)} f(t, s, x(s))ds \right| \leq (b - a)M,
\]

(15.13)

where \( M = \max |f(I, I, B_\rho(0))| \). Moreover, the family \( F(U) \) is equicontinuous. Fix \( \varepsilon \) and \( \varepsilon_1 = \varepsilon/(2(b - a)) \), \( \varepsilon_2 = \varepsilon/(2M) \). Since \( f \) and \( \tau \) are uniformly continuous on \( I \times I \times \overline{B_\rho(0)} \) and \( I \), respectively, we can find a \( \delta > 0 \) such
that $|f(t, s, x) - f(t_0, s, x)| \leq \varepsilon_1$ and $|\tau(t) - \tau(t_0)| \leq \varepsilon_2$ for $|t - t_0| < \delta$. Hence we infer for $|t - t_0| < \delta$

$$|F(x)(t) - F(x)(t_0)| = \left| \int_{t_0}^{\tau(t)} f(t, s, x(s))ds - \int_{t_0}^{\tau(t_0)} f(t_0, s, x(s))ds \right|$$

$$\leq \int_{t_0}^{\tau(t_0)} |f(t, s, x(s)) - f(t_0, s, x(s))|ds + \left| \int_{\tau(t_0)}^{\tau(t)} f(t, s, x(s))|ds \right|$$

$$\leq (b - a)\varepsilon_1 + \varepsilon_2 \rho = \varepsilon.$$ (15.14)

This implies that $F(U)$ is relatively compact by the Arzelà–Ascoli theorem (Theorem 3.4). Thus $F$ is compact. 

As a first application we use this result to show existence of solutions to integral equations.

**Theorem 15.11.** Let $F$ be as in the previous lemma. Then the integral equation

$$x - \lambda F(x) = y, \quad \lambda \in \mathbb{R}, y \in C(I, \mathbb{R}^n)$$ (15.15)

has at least one solution $x \in C(I, \mathbb{R}^n)$ if $|\lambda| \leq \rho/M(\rho)$, where $M(\rho) = (b - a) \max_{(s,t,u) \in I \times I \times \mathbb{R}^n} |f(s, t, x - y(s))|$ and $\rho > 0$ is arbitrary.

**Proof.** Note that, by our assumption on $\lambda$, $\lambda F + y$ maps $B_\rho(y)$ into itself. Now apply the Schauder fixed-point theorem. 

This result immediately gives the Peano theorem for ordinary differential equations.

**Theorem 15.12** (Peano). Consider the initial value problem

$$\dot{x} = f(t, x), \quad x(t_0) = x_0,$$ (15.16)

where $f \in C(I \times U, \mathbb{R}^n)$ and $I \subseteq \mathbb{R}$ is an interval containing $t_0$. Then (15.16) has at least one local solution $x \in C^1([t_0 - \varepsilon, t_0 + \varepsilon], \mathbb{R}^n)$, $\varepsilon > 0$. For example, any $\varepsilon$ satisfying $\varepsilon M(\varepsilon, \rho) \leq \rho$, $\rho > 0$ with $M(\varepsilon, \rho) = \max |f([t_0 - \varepsilon, t_0 + \varepsilon] \times B_\rho(x_0))|$ works. In addition, if $M(\varepsilon, \rho) \leq M(\varepsilon)(1 + \rho)$, then there exists a global solution.

**Proof.** For notational simplicity we make the shift $t \to t - t_0$, $x \to x - x_0$, $f(t, x) \to f(t + t_0, x + t_0)$ and assume $t_0 = 0, x_0 = 0$. In addition, it suffices to consider $t \geq 0$ since $t \to -t$ amounts to $f \to -f$.

Now observe, that (15.16) is equivalent to

$$x(t) - \int_0^t f(s, x(s))ds, \quad x \in C([-\varepsilon, \varepsilon], \mathbb{R}^n)$$ (15.17)
and the first part follows from our previous theorem. To show the second, fix $\varepsilon > 0$ and assume $M(\varepsilon, \rho) \leq \tilde{M}(\varepsilon)(1 + \rho)$. Then

$$|x(t)| \leq \int_0^t |f(s, x(s))| ds \leq \tilde{M}(\varepsilon) \int_0^t (1 + |x(s)|) ds$$

(15.18)

implies $|x(t)| \leq \exp(\tilde{M}(\varepsilon)\varepsilon)$ by Gronwall’s inequality. Hence we have an a priori bound which implies existence by the Leary–Schauder principle. Since $\varepsilon$ was arbitrary we are done. \qed
Chapter 16

The stationary Navier–Stokes equation

16.1. Introduction and motivation

In this chapter we turn to partial differential equations. In fact, we will only consider one example, namely the stationary Navier–Stokes equation. Our goal is to use the Leray–Schauder principle to prove an existence and uniqueness result for solutions.

Let $U (\neq \emptyset)$ be an open, bounded, and connected subset of $\mathbb{R}^3$. We assume that $U$ is filled with an incompressible fluid described by its velocity field $v_j(t, x)$ and its pressure $p(t, x)$, $(t, x) \in \mathbb{R} \times U$. The requirement that our fluid is incompressible implies $\partial_j v_j = 0$ (we sum over two equal indices from 1 to 3), which follows from the Gauss theorem since the flux through any closed surface must be zero.

Rather than just writing down the equation, let me give a short physical motivation. To obtain the equation which governs such a fluid we consider the forces acting on a small cube spanned by the points $(x_1, x_2, x_3)$ and $(x_1 + \Delta x_1, x_2 + \Delta x_2, x_3 + \Delta x_3)$. We have three contributions from outer forces, pressure differences, and viscosity.

The outer force density (force per volume) will be denoted by $K_j$ and we assume that it is known (e.g. gravity).

The force from pressure acting on the surface through $(x_1, x_2, x_3)$ normal to the $x_1$-direction is $p \Delta x_2 \Delta x_3 \delta_{1j}$. The force from pressure acting on the opposite surface is $-(p + \partial_1 p \Delta x_1) \Delta x_2 \Delta x_3 \delta_{1j}$. In summary, we obtain

$$-(\partial_j p) \Delta V, \quad (16.1)$$
where $\Delta V = \Delta x_1 \Delta x_2 \Delta x_3$.

The viscosity acting on the surface through $(x_1, x_2, x_3)$ normal to the $x_1$-direction is $-\eta \Delta x_2 \Delta x_3 \partial_1 v_j$ by some physical law. Here $\eta > 0$ is the viscosity constant of the fluid. On the opposite surface we have $\eta \Delta x_2 \Delta x_3 \partial_1 (v_j + \partial_1 v_j \Delta x_1)$. Adding up the contributions of all surface we end up with

$$\eta \Delta V \partial_i \partial_i v_j.$$ \hfill (16.2)

Putting it all together we obtain from Newton’s law

$$\rho \Delta V \frac{d}{dt} v_j(t, x(t)) = \eta \Delta V \partial_i \partial_i v_j(t, x(t)) - (\partial_j p(t, x(t)) + \Delta V K_j(t, x(t))),$$ \hfill (16.3)

where $\rho > 0$ is the density of the fluid. Dividing by $\Delta V$ and using the chain rule yields the Navier–Stokes equation

$$\rho \partial_t v_j = \eta \partial_i \partial_i v_j - \rho (v_i \partial_i) v_j - \partial_j p + K_j.$$ \hfill (16.4)

Note that it is no restriction to assume $\rho = 1$.

In what follows we will only consider the stationary Navier–Stokes equation

$$0 = \eta \partial_i \partial_i v_j - (v_i \partial_i) v_j - \partial_j p + K_j.$$ \hfill (16.5)

In addition to the incompressibility condition $\partial_j v_j = 0$ we also require the boundary condition $v|_{\partial U} = 0$, which follows from experimental observations.

In summary, we consider the problem (16.5) for $v$ in (e.g.) $X = \{ v \in C^2(\overline{U}, \mathbb{R}^3) | \partial_j v_j = 0 \text{ and } v|_{\partial U} = 0 \}$.

Our strategy is to rewrite the stationary Navier–Stokes equation in integral form, which is more suitable for our further analysis. For this purpose we need to introduce some function spaces first.

### 16.2. An insert on Sobolev spaces

Let $U$ be a bounded open subset of $\mathbb{R}^n$ and let $L^p(U, \mathbb{R})$ denote the Lebesgue spaces of $p$ integrable functions with norm

$$\|u\|_p = \left( \int_U |u(x)|^p dx \right)^{1/p}.$$ \hfill (16.6)

In the case $p = 2$ we even have a scalar product

$$\langle u, v \rangle_2 = \int_U u(x)v(x)dx$$ \hfill (16.7)

and our aim is to extend this case to include derivatives.

Given the set $C^1(U, \mathbb{R})$ we can consider the scalar product

$$\langle u, v \rangle_{2,1} = \int_U u(x)v(x)dx + \int_U (\partial_j u)(x)(\partial_j v)(x)dx.$$ \hfill (16.8)
16.2. An insert on Sobolev spaces

Taking the completion with respect to the associated norm we obtain the Sobolev space $H^1(U, \mathbb{R})$. Similarly, taking the completion of $C_c^1(U, \mathbb{R})$ with respect to the same norm, we obtain the Sobolev space $H^1_0(U, \mathbb{R})$. Here $C_c^r(U, Y)$ denotes the set of functions in $C^r(U, Y)$ with compact support. This construction of $H^1(U, \mathbb{R})$ implies that a sequence $u_k$ in $C^1_c(U, \mathbb{R})$ converges to $u \in H^1_0(U, \mathbb{R})$ if and only if $u_k$ and all its first order derivatives $\partial_j u_k$ converge in $L^2(U, \mathbb{R})$. Hence we can assign each $u \in H^1(U, \mathbb{R})$ its first order derivatives $\partial_j u$ by taking the limits from above. In order to show that this is a useful generalization of the ordinary derivative, we need to show that the derivative depends only on the limiting function $u \in L^2(U, \mathbb{R})$. To see this we need the following lemma.

**Lemma 16.1 (Integration by parts).** Suppose $u \in H^1_0(U, \mathbb{R})$ and $v \in H^1(U, \mathbb{R})$, then

\[
\int_U u(\partial_j v) \, dx = -\int_U (\partial_j u)v \, dx. \tag{16.9}
\]

**Proof.** By continuity it is no restriction to assume $u \in C^1_0(U, \mathbb{R})$ and $v \in C^1(U, \mathbb{R})$. Moreover, we can find a function $\phi \in C^1_0(U, \mathbb{R})$ which is 1 on the support of $u$. Hence by considering $\phi v$ we can even assume $v \in C^1_0(U, \mathbb{R})$.

Moreover, we can replace $U$ by a rectangle $K$ containing $U$ and extend $u$, $v$ to $K$ by setting it 0 outside $U$. Now use integration by parts with respect to the $j$-th coordinate. \qed

In particular, this lemma says that if $u \in H^1(U, \mathbb{R})$, then

\[
\int_U (\partial_j u)\phi \, dx = -\int_U u(\partial_j \phi) \, dx, \quad \phi \in C^\infty_0(U, \mathbb{R}). \tag{16.10}
\]

And since $C^\infty_0(U, \mathbb{R})$ is dense in $L^2(U, \mathbb{R})$, the derivatives are uniquely determined by $u \in L^2(U, \mathbb{R})$ alone. Moreover, if $u \in C^1(U, \mathbb{R})$, then the derivative in the Sobolev space corresponds to the usual derivative. In summary, $H^1(U, \mathbb{R})$ is the space of all functions $u \in L^2(U, \mathbb{R})$ which have first order derivatives (in the sense of distributions, i.e., (16.10)) in $L^2(U, \mathbb{R})$.

Next, we want to consider some additional properties which will be used later on. First of all, the Poincaré-Friedrichs inequality.

**Lemma 16.2 (Poincaré-Friedrichs inequality).** Suppose $u \in H^1_0(U, \mathbb{R})$, then

\[
\int_U u^2 \, dx \leq d_j^2 \int_U (\partial_j u)^2 \, dx, \tag{16.11}
\]

where $d_j = \sup \{|x_j - y_j| \mid (x_1, \ldots, x_n), (y_1, \ldots, y_n) \in U\}$.

**Proof.** Again we can assume $u \in C^1_c(U, \mathbb{R})$ and we assume $j = 1$ for notational convenience. Replace $U$ by a set $K = [a, b] \times \tilde{K}$ containing $U$ and
extend $u$ to $K$ by setting it 0 outside $U$. Then we have
\[
  u(x_1, x_2, \ldots, x_n)^2 = \left( \int_a^{x_1} 1 \cdot (\partial_1 u)(\xi, x_2, \ldots, x_n)d\xi \right)^2 
  \leq (b - a) \int_a^b (\partial_1 u)^2(\xi, x_2, \ldots, x_n)d\xi, \tag{16.12}
\]
where we have used the Cauchy-Schwarz inequality. Integrating this result over $[a, b]$ gives
\[
  \int_a^b u^2(\xi, x_2, \ldots, x_n)d\xi \leq (b - a)^2 \int_a^b (\partial_1 u)^2(\xi, x_2, \ldots, x_n)d\xi \tag{16.13}
\]
and integrating over $\tilde{K}$ finishes the proof. \(\blacksquare\)

Hence, from the viewpoint of Banach spaces, we could also equip $H^1_0(U, \mathbb{R})$ with the scalar product
\[
  \langle u, v \rangle = \int_U (\partial_j u)(x)(\partial_j v)(x)dx. \tag{16.14}
\]
This scalar product will be more convenient for our purpose and hence we will use it from now on. (However, all results stated will hold in either case.) The norm corresponding to this scalar product will be denoted by $|.|$.

Next, we want to consider the embedding $H^1_0(U, \mathbb{R}) \hookrightarrow L^2(U, \mathbb{R})$ a little closer. This embedding is clearly continuous since by the Poincaré-Friedrichs inequality we have
\[
  \|u\|_2 \leq \frac{d(U)}{\sqrt{n}}\|u\|, \quad d(U) = \sup\{|x - y|; x, y \in U\}. \tag{16.15}
\]
Moreover, by a famous result of Rellich, it is even compact. To see this we first prove the following inequality.

**Lemma 16.3 (Poincaré inequality).** Let $Q \subset \mathbb{R}^n$ be a cube with edge length $\rho$. Then
\[
  \int_Q u^2 dx \leq \frac{1}{\rho^n} \left( \int_Q u dx \right)^2 + \frac{n\rho^2}{2} \int_Q (\partial_k u)(\partial_k u)dx \tag{16.16}
\]
for all $u \in H^1(Q, \mathbb{R})$.

**Proof.** After a scaling we can assume $Q = (0, 1)^n$. Moreover, it suffices to consider $u \in C^1(Q, \mathbb{R})$.

Now observe
\[
  u(x) - u(\tilde{x}) = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} (\partial_i u)dx_i, \tag{16.17}
\]
where \( x^i = (\tilde{x}_1, \ldots, \tilde{x}_i, x_{i+1}, \ldots, x_n) \). Squaring this equation and using Cauchy–Schwarz on the right hand side we obtain
\[
u(x)^2 - 2u(x)u(\tilde{x}) + u(\tilde{x})^2 \leq \left( \sum_{i=1}^{n} \int_0^1 |\partial_i u| dx_i \right)^2 \leq n \sum_{i=1}^{n} \left( \int_0^1 |\partial_i u| dx_i \right)^2 \leq n \sum_{i=1}^{n} \int_0^1 (\partial_i u)^2 dx_i. \tag{16.18}\]

Now we integrate over \( x \) and \( \tilde{x} \), which gives
\[
2 \int_Q u^2 dx - 2 \left( \int_Q u \, dx \right)^2 \leq n \int_Q (\partial_i u)(\partial_i u) dx \tag{16.19}
\]
and finishes the proof. \( \square \)

Now we are ready to show Rellich’s compactness theorem.

**Theorem 16.4** (Rellich’s compactness theorem). Let \( U \) be a bounded open subset of \( \mathbb{R}^n \). Then the embedding
\[
H^1_0(U, \mathbb{R}) \hookrightarrow L^2(U, \mathbb{R}) \tag{16.20}
\]
is compact.

**Proof.** Pick a cube \( Q \) (with edge length \( \rho \)) containing \( U \) and a bounded sequence \( u^k \in H^1_0(U, \mathbb{R}) \). Since bounded sets are weakly compact, it is no restriction to assume that \( u^k \) is weakly convergent in \( L^2(U, \mathbb{R}) \). By setting \( u^k(x) = 0 \) for \( x \notin U \) we can also assume \( u^k \in H^1(Q, \mathbb{R}) \) (show this). Next, subdivide \( Q \) into \( N \) subcubes \( Q_i \) with edge lengths \( \rho/N \). On each subcube \( (16.16) \) holds and hence
\[
\int_Q u^2 dx = \int_Q |u|^2 dx = \sum_{i=1}^{N^n} \frac{N}{\rho} \left( \int_{Q_i} |u| dx \right)^2 + \frac{n\rho^2}{2N^2} \int_U (\partial_k u)(\partial_k u) dx \tag{16.21}
\]
for all \( u \in H^1_0(U, \mathbb{R}) \). Hence we infer
\[
\|u^k - u^\ell\|_2^2 \leq \sum_{i=1}^{N^n} \frac{N}{\rho} \left( \int_{Q_i} (u^k - u^\ell) dx \right)^2 + \frac{n\rho^2}{2N^2} \|u^k - u^\ell\|^2. \tag{16.22}
\]
The last term can be made arbitrarily small by picking \( N \) large. The first term converges to 0 since \( u^k \) converges weakly and each summand contains the \( L^2 \) scalar product of \( u^k - u^\ell \) and \( \chi_{Q_i} \) (the characteristic function of \( Q_i \)). \( \square \)

In addition to this result we will also need the following interpolation inequality.
Lemma 16.5 (Ladyzhenskaya inequality). Let \( U \subset \mathbb{R}^3 \). For all \( u \in H^1_0(U, \mathbb{R}) \) we have
\[
\|u\|_4 \leq \sqrt[4]{8}\|u\|_2^{1/4}\|u\|^{3/4}.
\] (16.23)

Proof. We first prove the case where \( u \in C^1_c(U, \mathbb{R}) \). The key idea is to start with \( U \subset \mathbb{R}^1 \) and then work one way up to \( U \subset \mathbb{R}^2 \) and \( U \subset \mathbb{R}^3 \).

If \( U \subset \mathbb{R}^1 \) we have
\[
\begin{align*}
\int_0^1 \partial_1 u^2(x_1) dx_1 &\leq 2 \int |u \partial_1 u| dx_1 \\
\end{align*}
\] (16.24)
and hence
\[
\max_{x \in U} u(x)^2 \leq 2 \int |u \partial_1 u| dx_1.
\] (16.25)
Here, if an integration limit is missing, it means that the integral is taken over the whole support of the function.

If \( U \subset \mathbb{R}^2 \) we have
\[
\begin{align*}
\int \int u^4 dx_1 dx_2 &\leq \int \int \max_x u(x, x_2)^2 dx_2 \int \max_y u(x_1, y)^2 dx_1 \\
\end{align*}
\] (16.26)
\[
\begin{align*}
\leq 4 \int \int |u \partial_1 u| dx_1 dx_2 \int \int |u \partial_2 u| dx_1 dx_2 \\
\leq 4 \left( \int \int u^2 dx_1 dx_2 \right)^{2/2} \left( \int \int (\partial_1 u)^2 dx_1 dx_2 \right)^{1/2} \left( \int \int (\partial_2 u)^2 dx_1 dx_2 \right)^{1/2} \\
\leq 4 \int \int u^2 dx_1 dx_2 \int \int ((\partial_1 u)^2 + (\partial_2 u)^2) dx_1 dx_2
\end{align*}
\]
Now let \( U \subset \mathbb{R}^3 \), then
\[
\begin{align*}
\int \int \int u^4 dx_1 dx_2 dx_3 &\leq \int \int \int u^2 dx_1 dx_2 dx_3 \int \int ((\partial_1 u)^2 + (\partial_2 u)^2) dx_1 dx_2 dx_3 \\
\end{align*}
\] (16.27)
and applying Cauchy–Schwarz finishes the proof for \( u \in C^1_c(U, \mathbb{R}) \).

If \( u \in H^1_0(U, \mathbb{R}) \) pick a sequence \( u_k \) in \( C^1_c(U, \mathbb{R}) \) which converges to \( u \) in \( H^1_0(U, \mathbb{R}) \) and hence in \( L^2(U, \mathbb{R}) \). By our inequality, this sequence is Cauchy in \( L^4(U, \mathbb{R}) \) and converges to a limit \( v \in L^4(U, \mathbb{R}) \). Since \( \|u\|_2 \leq \sqrt[4]{8\|u\|_4} \) \( \left( \int \int \int 1 \cdot u^2 dx \leq \sqrt{\int \int \int u^4 dx} \right) \), \( u_k \) converges to \( v \) in \( L^2(U, \mathbb{R}) \) as well and hence \( u = v \). Now take the limit in the inequality for \( u_k \).

As a consequence we obtain
\[
\|u\|_4 \leq \left( \frac{8d(U)}{\sqrt{3}} \right)^{1/4} \|u\|, \quad U \subset \mathbb{R}^3,
\] (16.28)
and

**Corollary 16.6.** The embedding

\[ H^1_0(U, \mathbb{R}) \hookrightarrow L^4(U, \mathbb{R}), \quad U \subset \mathbb{R}^3, \tag{16.29} \]

is compact.

**Proof.** Let \( u_k \) be a bounded sequence in \( H^1_0(U, \mathbb{R}) \). By Rellich’s theorem there is a subsequence converging in \( L^2(U, \mathbb{R}) \). By the Ladyzhenskaya inequality this subsequence converges in \( L^4(U, \mathbb{R}) \). \( \square \)

Our analysis clearly extends to functions with values in \( \mathbb{R}^n \) since we have \( H^1_0(U, \mathbb{R}^n) = \bigoplus_{j=1}^n H^1_0(U, \mathbb{R}) \).

### 16.3. Existence and uniqueness of solutions

Now we come to the reformulation of our original problem (16.5). We pick as underlying Hilbert space \( H^1_0(U, \mathbb{R}^3) \) with scalar product

\[ \langle u, v \rangle = \int_U (\partial_j u_i)(\partial_j v_i) dx. \tag{16.30} \]

Let \( \mathcal{X} \) be the closure of \( X \) in \( H^1_0(U, \mathbb{R}^3) \), that is,

\[ \mathcal{X} = \{ v \in C^2(\overline{U}, \mathbb{R}^3) | \partial_j v_j = 0 \text{ and } v|_{\partial U} = 0 \} = \{ v \in H^1_0(U, \mathbb{R}^3) | \partial_j v_j = 0 \}. \tag{16.31} \]

Now we multiply (16.5) by \( w \in \mathcal{X} \) and integrate over \( U \)

\[ \int_U \left( \eta \partial_k \partial_k v_j - (v_k \partial_k)v_j + K_j \right) w_j dx = \int_U (\partial_j p)w_j dx = 0. \tag{16.32} \]

Using integration by parts this can be rewritten as

\[ \int_U \left( \eta (\partial_k v_j)(\partial_k w_j) - v_k v_j(\partial_k w_j) - K_j w_j \right) dx = 0. \tag{16.33} \]

Hence if \( v \) is a solution of the Navier-Stokes equation, then it is also a solution of

\[ \eta \langle v, w \rangle - a(v, v, w) - \int_U K w dx = 0, \quad \text{for all } w \in \mathcal{X}, \tag{16.34} \]

where

\[ a(u, v, w) = \int_U u_k v_j(\partial_k w_j) dx. \tag{16.35} \]

In other words, (16.34) represents a necessary solubility condition for the Navier-Stokes equations. A solution of (16.34) will also be called a **weak solution** of the Navier-Stokes equations. If we can show that a weak solution is in \( C^2 \), then we can read our argument backwards and it will be also a classical solution. However, in general this might not be true and it will
only solve the Navier-Stokes equations in the sense of distributions. But let us try to show existence of solutions for (16.34) first.

For later use we note

\[
a(v, v, v) = \int_U v_k v_j (\partial_k v_j) dx = \frac{1}{2} \int_U v_k \partial_k (v_j v_j) dx
\]

\[= -\frac{1}{2} \int_U (v_j v_j) \partial_k v_k dx = 0, \quad v \in \mathcal{X}.
\]

(16.36)

We proceed by studying (16.34). Let \( K \in L^2(U, \mathbb{R}^3) \), then \( \int_U K w dx \) is a linear functional on \( \mathcal{X} \) and hence there is \( \tilde{K} \in \mathcal{X} \) such that

\[
\int_U K w dx = \langle \tilde{K}, w \rangle, \quad w \in \mathcal{X}.
\]

(16.37)

Moreover, the same is true for the map \( a(u, v, w) \), \( u, v \in \mathcal{X} \), and hence there is an element \( B(u, v) \in \mathcal{X} \) such that

\[
a(u, v, w) = \langle B(u, v), w \rangle, \quad w \in \mathcal{X}.
\]

(16.38)

In addition, the map \( B : \mathcal{X}^2 \to \mathcal{X} \) is bilinear. In summary we obtain

\[
\langle \eta v - B(v, v) - \tilde{K}, w \rangle = 0, \quad w \in \mathcal{X},
\]

(16.39)

and hence

\[
\eta v - B(v, v) = \tilde{K}.
\]

(16.40)

So in order to apply the theory from our previous chapter, we need a Banach space \( Y \) such that \( \mathcal{X} \hookrightarrow Y \) is compact.

Let us pick \( Y = L^4(U, \mathbb{R}^3) \). Then, applying the Cauchy-Schwarz inequality twice to each summand in \( a(u, v, w) \) we see

\[
|a(u, v, w)| \leq \sum_{j,k} \left( \int_U (u_k v_j)^2 dx \right)^{1/2} \left( \int_U (\partial_k w_j)^2 dx \right)^{1/2}
\]

\[\leq \|w\| \sum_{j,k} \left( \int_U (u_k)^4 dx \right)^{1/4} \left( \int_U (v_j)^4 dx \right)^{1/4} = \|u\|_4 \|v\|_4 \|w\|.
\]

(16.41)

Moreover, by Corollary 16.6 the embedding \( \mathcal{X} \hookrightarrow Y \) is compact as required.

Motivated by this analysis we formulate the following theorem.

**Theorem 16.7.** Let \( \mathcal{X} \) be a Hilbert space, \( Y \) a Banach space, and suppose there is a compact embedding \( \mathcal{X} \hookrightarrow Y \). In particular, \( \|u\|_Y \leq \beta \|u\| \). Let \( a : \mathcal{X}^3 \to \mathbb{R} \) be a multilinear form such that

\[
|a(u, v, w)| \leq \alpha \|u\|_Y \|v\|_Y \|w\|
\]

(16.42)

and \( a(v, v, v) = 0 \). Then for any \( \tilde{K} \in \mathcal{X} \), \( \eta > 0 \) we have a solution \( v \in \mathcal{X} \) to the problem

\[
\eta \langle v, w \rangle - a(v, v, w) = \langle \tilde{K}, w \rangle, \quad w \in \mathcal{X}.
\]

(16.43)
Moreover, if $2\alpha\beta|\tilde{K}| < \eta^2$ this solution is unique.

**Proof.** It is no loss to set $\eta = 1$. Arguing as before we see that our equation is equivalent to

$$v - B(v, v) + \tilde{K} = 0,$$

(16.44)

where our assumption (16.42) implies

$$\|B(u, v)\| \leq \alpha\|u\|_Y\|v\|_Y \leq \alpha\beta^2\|u\|\|v\|$$

(16.45)

Here the second equality follows since the embedding $X \hookrightarrow Y$ is continuous.

Abbreviate $F(v) = B(v, v)$. Observe that $F$ is locally Lipschitz continuous since if $\|u\|, \|v\| \leq \rho$ we have

$$\|F(u) - F(v)\| = \|B(u - v, u) - B(v, u - v)\| \leq 2\alpha\rho\|u - v\|_Y$$

$$\leq 2\alpha\beta^2\rho\|u - v\|.$$  

(16.46)

Moreover, let $v_n$ be a bounded sequence in $X$. After passing to a subsequence we can assume that $v_n$ is Cauchy in $Y$ and hence $F(v_n)$ is Cauchy in $X$ by $\|F(u) - F(v)\| \leq 2\alpha\rho\|u - v\|_Y$. Thus $F : X \rightarrow X$ is compact.

Hence all we need to apply the Leray-Schauder principle is an a priori estimate. Suppose $v$ solves $v = tF(v) + tK$, $t \in [0, 1]$, then

$$\langle v, v \rangle = ta(v, v, v) + t\langle K, v \rangle = t\langle K, v \rangle.$$  

(16.47)

Hence $\|v\| \leq \|\tilde{K}\|$ is the desired estimate and the Leray-Schauder principle yields existence of a solution.

Now suppose there are two solutions $v_i$, $i = 1, 2$. By our estimate they satisfy $\|v_i\| \leq \|\tilde{K}\|$ and hence $\|v_1 - v_2\| = \|F(v_1) - F(v_2)\| \leq 2\alpha\beta^2\|\tilde{K}\|\|v_1 - v_2\|$ which is a contradiction if $2\alpha\beta^2\|\tilde{K}\| < 1$. \qed

Hence we have found a solution $v$ to the generalized problem (16.34). This solution is unique if $2\left(\frac{2d(U)}{\sqrt{3}}\right)^{3/2}\|K\|_2 < \eta^2$. Under suitable additional conditions on the outer forces and the domain, it can be shown that weak solutions are $C^2$ and thus also classical solutions. However, this is beyond the scope of this introductory text.
Chapter 17

Monotone maps

17.1. Monotone maps

The Leray–Schauder theory can only be applied to compact perturbations of the identity. If $F$ is not compact, we need different tools. In this section we briefly present another class of maps, namely monotone ones, which allow some progress.

If $F : \mathbb{R} \to \mathbb{R}$ is continuous and we want $F(x) = y$ to have a unique solution for every $y \in \mathbb{R}$, then $f$ should clearly be strictly monotone increasing (or decreasing) and satisfy $\lim_{x \to \pm \infty} F(x) = \pm \infty$. Rewriting these conditions slightly such that they make sense for vector valued functions the analogous result holds.

Lemma 17.1. Suppose $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuous and satisfies

$$\lim_{|x| \to \infty} \frac{F(x)x}{|x|} = \infty. \quad (17.1)$$

Then the equation

$$F(x) = y \quad (17.2)$$

has a solution for every $y \in \mathbb{R}^n$. If $F$ is strictly monotone

$$(F(x) - F(y))(x - y) > 0, \quad x \neq y, \quad (17.3)$$

then this solution is unique.

Proof. Our first assumption implies that $G(x) = F(x) - y$ satisfies $G(x)x = F(x)x - yx > 0$ for $|x|$ sufficiently large. Hence the first claim follows from Theorem 14.13. The second claim is trivial. \qed
Now we want to generalize this result to infinite dimensional spaces. Throughout this chapter, $H$ will be a Hilbert space with scalar product $\langle ., . \rangle$. A map $F : H \to H$ is called **monotone** if
\[
\langle F(x) - F(y), x - y \rangle \geq 0, \quad x, y \in H,
\]
strictly monotone if
\[
\langle F(x) - F(y), x - y \rangle > 0, \quad x \neq y \in H,
\]
and finally **strongly monotone** if there is a constant $C > 0$ such that
\[
\langle F(x) - F(y), x - y \rangle \geq C \| x - y \|^2, \quad x, y \in H.
\]
Note that the same definitions can be made for a Banach space $X$ and mappings $F : X \to X^*$.

Observe that if $F$ is strongly monotone, then it automatically satisfies
\[
\lim_{|x| \to \infty} \frac{\langle F(x), x \rangle}{\| x \|} = \infty.
\]
(Just take $y = 0$ in the definition of strong monotonicity.) Hence the following result is not surprising.

**Theorem 17.2** (Zarantonello). **Suppose** $F \in C(H, H)$ **is (globally) Lipschitz continuous and strongly monotone. Then, for each $y \in H$ the equation**
\[
F(x) = y
\]
**has a unique solution** $x(y) \in H$ **which depends continuously on** $y$.

**Proof.** Set
\[
G(x) = x - t(F(x) - y), \quad t > 0,
\]
then $F(x) = y$ is equivalent to the fixed point equation
\[
G(x) = x.
\]
It remains to show that $G$ is a contraction. We compute
\[
\| G(x) - G(\tilde{x}) \|^2 = \| x - \tilde{x} \|^2 - 2t \langle F(x) - F(\tilde{x}), x - \tilde{x} \rangle + t^2 \| F(x) - F(\tilde{x}) \|^2
\leq (1 - 2 \frac{C}{L}(Lt) + (Lt)^2) \| x - \tilde{x} \|^2,
\]
where $L$ is a Lipschitz constant for $F$ (i.e., $\| F(x) - F(\tilde{x}) \| \leq L \| x - \tilde{x} \|$). Thus, if $t \in (0, \frac{2C}{L})$, $G$ is a uniform contraction and the rest follows from the uniform contraction principle. □

Again observe that our proof is constructive. In fact, the best choice for $t$ is clearly $t = \frac{C}{L^2}$ such that the contraction constant $\theta = 1 - (\frac{C}{L^2})^2$ is minimal. Then the sequence
\[
x_{n+1} = x_n - \frac{C}{L^2}(F(x_n) - y), \quad x_0 = y,
\]

\[
17. Monotone maps
\]
17.2. The nonlinear Lax–Milgram theorem

As a consequence of the last theorem we obtain a nonlinear version of the Lax–Milgram theorem. We want to investigate the following problem:

\[ a(x, y) = b(y), \quad \text{for all } y \in \mathcal{H}, \quad (17.12) \]

where \( a : \mathcal{H}^2 \to \mathbb{R} \) and \( b : \mathcal{H} \to \mathbb{R} \). For this equation the following result holds.

**Theorem 17.3** (Nonlinear Lax–Milgram theorem). Suppose \( b \in \mathcal{L}(\mathcal{H}, \mathbb{R}) \) and \( a(x, \cdot) \in \mathcal{L}(\mathcal{H}, \mathbb{R}) \), \( x \in \mathcal{H} \), are linear functionals such that there are positive constants \( L \) and \( C \) such that for all \( x, y, z \in \mathcal{H} \) we have

\[ a(x, x - y) - a(y, x - y) \geq C|x - y|^2 \quad (17.13) \]

and

\[ |a(x, z) - a(y, z)| \leq L|z||x - y|. \quad (17.14) \]

Then there is a unique \( x \in \mathcal{H} \) such that (17.12) holds.

**Proof.** By the Riez lemma (Theorem 2.10) there are elements \( F(x) \in \mathcal{H} \) and \( z \in \mathcal{H} \) such that \( a(x, y) = b(y) \) is equivalent to \( \langle F(x) - z, y \rangle = 0 \), \( y \in \mathcal{H} \), and hence to

\[ F(x) = z. \quad (17.15) \]

By (17.13) the map \( F \) is strongly monotone. Moreover, by (17.14) we infer

\[ \|F(x) - F(y)\| = \sup_{\tilde{x} \in \mathcal{H}, \|	ilde{x}\|=1} |\langle F(x) - F(y), \tilde{x} \rangle| \leq L\|x - y\| \quad (17.16) \]

that \( F \) is Lipschitz continuous. Now apply Theorem 17.2. \( \square \)

The special case where \( a \in \mathcal{L}^2(\mathcal{H}, \mathbb{R}) \) is a bounded bilinear form which is strongly coercive, that is,

\[ a(x, x) \geq C\|x\|^2, \quad x \in \mathcal{H}, \quad (17.17) \]

is usually known as (linear) Lax–Milgram theorem (Theorem 2.15).

The typical application of this theorem is the existence of a unique weak solution of the Dirichlet problem for **elliptic equations**

\[ -\partial_i A_{ij}(x)\partial_j u(x) + b_j(x)\partial_j u(x) + c(x)u(x) = f(x), \quad x \in U, \]

\[ u(x) = 0, \quad x \in \partial U, \quad (17.18) \]

where \( U \) is a bounded open subset of \( \mathbb{R}^n \). By elliptic we mean that all coefficients \( A, b, c \) plus the right hand side \( f \) are bounded and \( a_0 > 0 \), where

\[ a_0 = \inf_{e \in S^n, x \in U} e_i A_{ij}(x)e_j, \quad b_0 = \sup_{x \in U} |b(x)|, \quad c_0 = \inf_{x \in U} c(x). \quad (17.19) \]
As in Section 16.3 we pick $H^1_0(U, \mathbb{R})$ with scalar product
\[ \langle u, v \rangle = \int_U (\partial_j u)(\partial_j v)dx \] (17.20)
as underlying Hilbert space. Next we multiply (17.18) by $v \in H^1_0$ and integrate over $U$
\[ \int_U \left( -\partial_i A_{ij}(x)\partial_j u(x) + b_j(x)\partial_j u(x) + c(x)u(x) \right)v(x) dx = \int_U f(x)v(x) dx. \] (17.21)
After integration by parts we can write this equation as
\[ a(v, u) = f(v), \quad v \in H^1_0, \] (17.22)
where
\[ a(v, u) = \int_U \left( \partial_j v(x)A_{ij}(x)\partial_j u(x) + b_j(x)v(x)\partial_j u(x) + c(x)v(x)u(x) \right)dx \]
\[ f(v) = \int_U f(x)v(x) dx, \] (17.23)
We call a solution of (17.22) a \textbf{weak solution} of the elliptic Dirichlet problem (17.18).

By a simple use of the Cauchy-Schwarz and Poincaré-Friedrichs inequalities we see that the bilinear form $a(u, v)$ is bounded. To be able to apply the (linear) Lax–Milgram theorem we need to show that it satisfies
\[ a(u, u) \geq C \int |\partial_j u|^2 dx. \]
Using (17.19) we have
\[ a(u, u) \geq \int_U \left( a_0|\partial_j u|^2 - b_0|u|\|\partial_j u\| + c_0|u|^2 \right), \] (17.24)
and we need to control the middle term. If $b_0 = 0$ there is nothing to do and it suffices to require $c_0 \geq 0$.

If $b_0 > 0$ we distribute the middle term by means of the elementary inequality
\[ |u|\|\partial_j u\| \leq \frac{\varepsilon}{2}|u|^2 + \frac{1}{2\varepsilon}|\partial_j u|^2 \] (17.25)
which gives
\[ a(u, u) \geq \int_U \left( (a_0 - \frac{b_0}{2\varepsilon})|\partial_j u|^2 + (c_0 - \frac{\varepsilon b_0}{2})|u|^2 \right). \] (17.26)
Since we need $a_0 - \frac{b_0}{2\varepsilon} > 0$ and $c_0 - \frac{\varepsilon b_0}{2} \geq 0$, or equivalently $\frac{2c_0}{b_0} \geq \varepsilon > \frac{b_0}{2a_0}$, we see that we can apply the Lax–Milgram theorem if $4a_0c_0 > b_0^2$. In summary, we have proven

\textbf{Theorem 17.4.} \textit{The elliptic Dirichlet problem (17.18) has a unique weak solution} $u \in H^1_0(U, \mathbb{R})$ \textit{if} $a_0 > 0$, $b_0 = 0$, $c_0 \geq 0$ \textit{or} $a_0 > 0$, $4a_0c_0 > b_0^2$.  

17.3. The main theorem of monotone maps

Now we return to the investigation of \( F(x) = y \) and weaken the conditions of Theorem 17.2. We will assume that \( \mathcal{H} \) is a separable Hilbert space and that \( F : \mathcal{H} \rightarrow \mathcal{H} \) is a continuous monotone map satisfying

\[
\lim_{|x| \to \infty} \frac{\langle F(x), x \rangle}{\|x\|} = \infty.
\]  

(17.27)

In fact, if suffices to assume that \( F \) is weakly continuous

\[
\lim_{n \to \infty} \langle F(x_n), y \rangle = \langle F(x), y \rangle, \quad \text{for all } y \in \mathcal{H}
\]  

(17.28)

whenever \( x_n \to x \).

The idea is as follows: Start with a finite dimensional subspace \( \mathcal{H}_n \subset \mathcal{H} \) and project the equation \( F(x) = y \) to \( \mathcal{H}_n \) resulting in an equation

\[
F_n(x_n) = y_n, \quad x_n, y_n \in \mathcal{H}_n.
\]  

(17.29)

More precisely, let \( P_n \) be the (linear) projection onto \( \mathcal{H}_n \) and set \( F_n(x_n) = P_n F(x_n), \ y_n = P_n y \) (verify that \( F_n \) is continuous and monotone!).

Now Lemma 17.1 ensures that there exists a solution \( u_n \). Now chose the subspaces \( \mathcal{H}_n \) such that \( \mathcal{H}_n \to \mathcal{H} \) (i.e., \( \mathcal{H}_n \subset \mathcal{H}_{n+1} \) and \( \bigcup_{n=1}^{\infty} \mathcal{H}_n \) is dense). Then our hope is that \( u_n \) converges to a solution \( u \).

This approach is quite common when solving equations in infinite dimensional spaces and is known as Galerkin approximation. It can often be used for numerical computations and the right choice of the spaces \( \mathcal{H}_n \) will have a significant impact on the quality of the approximation.

So how should we show that \( x_n \) converges? First of all observe that our construction of \( x_n \) shows that \( x_n \) lies in some ball with radius \( R_n \), which is chosen such that

\[
\langle F_n(x), x \rangle > \|y_n\|\|x\|, \quad \|x\| \geq R_n, \ x \in \mathcal{H}_n.
\]  

(17.30)

Since \( \langle F_n(x), x \rangle = \langle P_n F(x), x \rangle = \langle F(x), P_n x \rangle = \langle F(x), x \rangle \) for \( x \in \mathcal{H}_n \) we can drop all \( n \)'s to obtain a constant \( R \) which works for all \( n \). So the sequence \( x_n \) is uniformly bounded

\[
\|x_n\| \leq R.
\]  

(17.31)

Now by a well-known result there exists a weakly convergent subsequence. That is, after dropping some terms, we can assume that there is some \( x \) such that \( x_n \to x \), that is,

\[
\langle x_n, z \rangle \to \langle x, z \rangle, \quad \text{for every } z \in \mathcal{H}.
\]  

(17.32)

And it remains to show that \( x \) is indeed a solution. This follows from
Lemma 17.5. Suppose $F : \mathcal{H} \to \mathcal{H}$ is weakly continuous and monotone, then
\[ \langle y - F(z), x - z \rangle \geq 0 \quad \text{for every } z \in \mathcal{H} \] (17.33)
implies $F(x) = y$.

Proof. Choose $z = x \pm tw$, then $\mp \langle y - F(x \pm tw), w \rangle \geq 0$ and by continuity $\mp \langle y - F(x), w \rangle \geq 0$. Thus $\langle y - F(x), w \rangle = 0$ for every $w$ implying $y - F(x) = 0$. \hfill \square

Now we can show

Theorem 17.6 (Browder, Minty). Suppose $F : \mathcal{H} \to \mathcal{H}$ is weakly continuous, monotone, and satisfies
\[ \lim_{|x| \to \infty} \frac{\langle F(x), x \rangle}{\|x\|} = \infty. \] (17.34)
Then the equation
\[ F(x) = y \] (17.35)
has a solution for every $y \in \mathcal{H}$. If $F$ is strictly monotone then this solution is unique.

Proof. Abbreviate $y_n = F(x_n)$, then we have $\langle y - F(z), x_n - z \rangle = \langle y_n - F_n(z), x_n - z \rangle \geq 0$ for $z \in \mathcal{H}_n$. Taking the limit implies $\langle y - F(z), x - z \rangle \geq 0$ for every $z \in \mathcal{H}_\infty = \bigcup_{n=1}^{\infty} \mathcal{H}_n$. Since $\mathcal{H}_\infty$ is dense, $\langle y - F(z), x - z \rangle \geq 0$ for every $z \in \mathcal{H}$ by continuity and hence $F(x) = y$ by our lemma. \hfill \square

Note that in the infinite dimensional case we need monotonicity even to show existence. Moreover, this result can be further generalized in two more ways. First of all, the Hilbert space $\mathcal{H}$ can be replaced by a reflexive Banach space $X$ if $F : X \to X^*$. The proof is almost identical. Secondly, it suffices if
\[ t \mapsto \langle F(x + ty), z \rangle \] (17.36)
is continuous for $t \in [0, 1]$ and all $x, y, z \in \mathcal{H}$, since this condition together with monotonicity can be shown to imply weak continuity.
Some set theory

At the beginning of the 20th century Russell showed with his famous paradox that naive set theory can lead into contradictions. Hence it was replaced by axiomatic set theory, more specific we will take the Zermelo–Fraenkel set theory, which assumes existence of some sets (like the empty set and the integers) and defines what operations are allowed. Somewhat informally (i.e. without writing them using the symbolism of first order logic) they can be stated as follows:

- **Axiom of empty set.** There is a set ∅ which contains no elements.
- **Axiom of extensionality.** Two sets $A$ and $B$ are equal $A = B$ if they contain the same elements. If a set $A$ contains all elements from a set $B$, it is called a subset $A \subseteq B$. In particular $A \subseteq B$ and $B \subseteq A$ if and only if $A = B$.
- **Axiom of pairing.** If $A$ and $B$ are sets, then there exists a set $\{A, B\}$ which contains $A$ and $B$ as elements. One writes $\{A, A\} = \{A\}$.
- **Axiom of union.** Given a set $\mathcal{F}$ whose elements are again sets, there is a set $A = \bigcup \mathcal{F}$ containing every element that is a member of some member of $\mathcal{F}$. In particular, given two sets $A, B$ there exists a set $A \cup B = \bigcup \{A, B\}$ consisting of the elements of both sets.
- **Axiom schema of specification.** Given a set $A$ and a given a logical statement $\phi(x)$ depending on $x \in A$ we can form the set $B = \{x \in A | \phi(x)\}$ of all elements from $A$ obeying $\phi$. For example, given two sets $A$ and $B$ we can define their intersection
as \( A \cap B = \{ x \in A \cup B | (x \in A) \land (x \in B) \} \) and the complement as \( A \setminus B = \{ x \in A | x \notin B \} \). Or the intersection of a family of sets \( \mathcal{F} \) as \( \bigcap \mathcal{F} = \{ x \in \bigcup \mathcal{F} | \forall F \in \mathcal{F} : x \in F \} \).

- **Axiom of power set.** For any set \( A \), there is a set \( \mathcal{P}(A) \) that contains every subset of \( A \).

From these axioms one can define ordered pairs as \((x, y) = \{\{x\}, \{x, y\}\}\) and the Cartesian product as \( A \times B = \{ z \in \mathcal{P}(A \cup \mathcal{P}(A \cup B)) | \exists x \in A, y \in B : z = (x, y) \} \). Functions \( f : A \rightarrow B \) are defined as single valued relations, that is \( f \subseteq A \times B \) such that \((x, y) \in f \) and \((x, \tilde{y}) \in f \) implies \( y = \tilde{y} \).

- **Axiom schema of replacement.** For every function \( f \) the image of a set \( A \) is again a set \( B = \{ f(x) | x \in A \} \).

- **Axiom of infinity.** There exists a set \( A \) which contains the empty set and for every element \( x \in A \) we also have \( x \cup \{x\} \in A \). The smallest such set \( \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \ldots \} \) can be identified with the integers via \( 0 = \emptyset, 1 = \{\emptyset\}, 2 = \{\emptyset, \{\emptyset\}\}, \ldots \)

- **Axiom of Regularity.** Every nonempty set \( A \) contains an element \( x \) with \( x \cap A = \emptyset \). This excludes for example the possibility that a set contains itself as an element (apply the axiom to \( \{A\} \)). Similarly, we can only have \( A \in B \) or \( B \in A \) but not both (apply it to the set \( \{A, B\} \)).

Hence a set is something which can be constructed from the above axioms. Of course this raises the question if these axioms are consistent but as has been shown by Gödel this question cannot be answered: If ZF contains a statement of its own consistency then ZF is inconsistent. In fact, the same holds for any other sufficiently rich (such that one can do basic math) system of axioms. In particular, it also holds for ZFC defined below. So we have to live with the fact that someday someone might come and prove that ZFC is inconsistent.

Starting from ZF one can develop basic analysis (including the construction of the real numbers). However, it turns out that several fundamental results require yet another construction for their proof:

Given an index set \( A \) and for every \( \alpha \in A \) some set \( M_{\alpha} \) the product \( \prod_{\alpha \in A} M_{\alpha} \) is defined to be the set of all \( \varphi : A \rightarrow \bigcup_{\alpha \in A} M_{\alpha} \) which assigns each element \( \alpha \in A \) some element \( m_{\alpha} \in M_{\alpha} \). If all sets \( M_{\alpha} \) are nonempty it seems quite reasonable that there should be such a choice function which chooses an element from \( M_{\alpha} \) for every \( \alpha \in A \). However, no matter how obvious this might seem, it cannot be deduced from the ZF axioms alone and hence has to be added:
A. Some set theory

- **Axiom of Choice**: Given an index set $A$ and nonempty sets \{${M}_\alpha$\}$_{\alpha \in A}$ their product \(\prod_{\alpha \in A} M_\alpha\) is nonempty.

ZF augmented by the axiom of choice is known as **ZFC** and we accept it as the fundament upon which our functional analytic house is built.

Note that the axiom of choice is not only used to ensure that infinite products are nonempty but also in many proofs! For example, suppose you start with a set $M_1$ and recursively construct some sets $M_n$ such that in every step you have a nonempty set. Then the axiom of choice guarantees the existence of a sequence $x = (x_n)_{n \in \mathbb{N}}$ with $x_n \in M_n$.

The axiom of choice has many important consequences (in fact, many of them are in fact equivalent to the axiom of choice and it is hence a matter of taste which one choose as axiom).

A **partial order** is a binary relation "\(\preceq\)" over a set $P$ such that for all $A, B, C \in P$:

- $A \preceq A$ (reflexivity),
- if $A \preceq B$ and $B \preceq A$ then $A = B$ (antisymmetry),
- if $A \preceq B$ and $B \preceq C$ then $A \preceq C$ (transitivity).

It is custom to write $A \prec B$ if $A \preceq B$ and $A \neq B$.

**Example.** Let $\mathcal{P}(X)$ be the collections of all subsets of a set $X$. Then $\mathcal{P}$ is partially ordered by inclusion $\subseteq$.

It is important to emphasize that two elements of $\mathcal{P}$ need not be comparable, that is, in general neither $A \preceq B$ nor $B \preceq A$ might hold. However, if any two elements are comparable, $\mathcal{P}$ will be called **totally ordered**. A set with a total order is called **well-ordered** if every nonempty subset has a **least element**, that is some $A \in \mathcal{P}$ with $A \preceq B$ for every $B \in \mathcal{P}$. Note that the least element is unique by antisymmetry.

**Example.** $\mathbb{R}$ with $\leq$ is totally ordered and $\mathbb{N}$ with $\leq$ is well-ordered.

On every well-ordered set we have the

**Theorem A.1** (Induction principle). Let $K$ be well ordered and let $S(k)$ be a statement for arbitrary $k \in K$. Then, if $A(l)$ true for all $l \prec k$ implies $A(k)$ true, then $A(k)$ is true for all $k \in K$.

**Proof.** Otherwise the set of all $k$ for which $A(k)$ is false had a least element $k_0$. But by our choice of $k_0$, $A(l)$ holds for all $l \prec k_0$ and thus for $k_0$ contradicting our assumption.

The induction principle also shows that in a well-ordered set functions $f$ can be defined recursively, that is, by a function $\varphi$ which computes the
value of $f(k)$ from the values $f(l)$ for all $l < k$. Indeed, the induction principle implies that on the set $M_k = \{ l \in K | l < k \}$ there is at most one such function $f_k$. Since $k$ is arbitrary, $f$ is unique. In case of the integers existence of $f_k$ is also clear provided $f(1)$ is given. In general, one can prove existence provided $f_k$ is given for some $k$ but we will not need this.

If $\mathcal{P}$ is partially ordered, then every totally ordered subset is also called a **chain**. If $\mathcal{Q} \subseteq \mathcal{P}$, then an element $M \in \mathcal{P}$ satisfying $A \lesssim M$ for all $A \in \mathcal{Q}$ is called an **upper bound**.

**Example.** Let $\mathcal{P}(X)$ as before. Then a collection of subsets $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{P}(X)$ satisfying $A_n \subseteq A_{n+1}$ is a chain. The set $\bigcup_n A_n$ is an upper bound. \hfill \diamond

An element $M \in \mathcal{P}$ for which $M \lesssim A$ for some $A \in \mathcal{P}$ is only possible if $M = A$ is called a **maximal element**.

**Theorem A.2** (Zorn’s lemma). *Every partially ordered set in which every chain has an upper bound contains at least one maximal element.*

**Proof.** Suppose it were false. Then to every chain $C$ we can assign an element $m(C)$ such that $m(C) \succ x$ for all $x \in C$ (here we use the axiom of choice). We call a chain $C$ distinguished if it is well-ordered and if for every segment $C_x = \{ y \in C | y < x \}$ we have $m(C_x) = x$. We will also regard $C$ as a segment of itself.

Then (since for the least element of $C$ we have $C_x = \emptyset$) every distinguished chain must start like $m(\emptyset) \prec m(m(\emptyset)) \prec \cdots$ and given two segments $C, D$ we expect that always one must be a segment of the other.

So let us first prove this claim. Suppose $D$ is not a segment of $C$. Then we need to show $C = D$ for some $z$. We start by showing that $x \in C$ implies $x \in D$ and $C_x = D_x$. To see this suppose it were wrong and let $x$ be the least $x \in C$ for which it fails. Then $y \in K_x$ implies $y \in L$ and hence $K_x \subset L$. Then, since $C_x \neq D$ by assumption, we can find a least $z \in D \setminus C_x$. In fact we must even have $z \succ C_x$ since otherwise we could find a $y \in C_x$ such that $x \succ y \succ z$. But then, using that it holds for $y, y \in D$ and $C_y = D_y$ so we get the contradiction $z \in D_y = C_y \subset C_x$. So $z \succ C_x$ and thus also $C_x = D_z$ which in turn shows $x = m(C_x) = m(D_z) = z$ and proves that $x \in C$ implies $x \in D$ and $C_x = D_x$. In particular $C \subset D$ and as before $C = D_z$ for the least $z \in D \setminus C$. This proves the claim.

Now using this claim we see that we can take the union over all distinguished chains to get a maximal distinguished chain $C_{\text{max}}$. But then we could add $m(C_{\text{max}}) \notin C_{\text{max}}$ to $C_{\text{max}}$ to get a larger distinguished chain $C_{\text{max}} \cup \{ m(C_{\text{max}}) \}$ contradiction maximality. \hfill \Box

We will also frequently use the **cardinality** of sets: Two sets $A$ and $B$ have the same cardinality, written as $|A| = |B|$, if there is a bijection
\(\varphi : A \to B\). We write \(|A| \leq |B|\) if there is an injective map \(\varphi : A \to B\). Note that \(|A| \leq |B|\) and \(|B| \leq |C|\) implies \(|A| \leq |C|\). A set \(A\) is called infinite if \(|A| \geq |\mathbb{N}|\), countable if \(|A| \leq |\mathbb{N}|\), and countably infinite if \(|A| = |\mathbb{N}|\).

**Theorem A.3** (Schröder–Bernstein). \(|A| \leq |B|\) and \(|B| \leq |A|\) implies \(|A| = |B|\).

**Proof.** Suppose \(\varphi : A \to B\) and \(\psi : B \to A\) are two injective maps. Now consider sequences \(x_n\) defined recursively via \(x_{2n+1} = \varphi(x_{2n})\) and \(x_{2n+1} = \psi(x_{2n})\). Given a start value \(x_0 \in A\) the sequence is uniquely defined but might terminate at a negative integer since our maps are not surjective. In any case, if an element appears in two sequences, the elements to the left and to the right must also be equal (use induction) and hence the two sequences differ only by an index shift. So the ranges of such sequences form a partition for \(A \sqcup B\) and it suffices to find a bijection between elements in one partition. If the sequence stops at an element in \(A\) we can take \(\varphi\). If the sequence stops at an element in \(B\) we can take \(\psi^{-1}\). If the sequence is doubly infinite either of the previous choices will do. \(\square\)

**Theorem A.4** (Zerlemo). Either \(|A| \leq |B|\) or \(|B| \leq |A|\).

**Proof.** Consider the set of all bijective functions \(\varphi_\alpha : A_\alpha \to B\) with \(A_\alpha \subseteq A\). Then we can define a partial ordering via \(\varphi_\alpha \preceq \varphi_\beta\) if \(A_\alpha \subseteq A_\beta\) and \(\varphi_\beta|_{A_\alpha} = \varphi_\alpha\). Then every chain has an upper bound (the unique function defined on the union of all domains) and by Zorn’s lemma there is a maximal element \(\varphi_m\). For \(\varphi_m\) we have either \(A_m = A\) or \(\varphi_m(A_m) = B\) since otherwise there is some \(x \in A \setminus A_m\) and some \(y \in B \setminus f(A_m)\) which could be used to extend \(\varphi_m\) to \(A_m \cup \{x\}\) by setting \(\varphi(x) = y\). But if \(A_m = A\) we have \(|A| \leq |B|\) and if \(\varphi_m(A_m) = B\) we have \(|B| \leq |A|\). \(\square\)

The powerset \(\mathcal{P}(A)\) of a set \(A\) is the set of all its subsets. Its cardinality is strictly larger than the cardinality of the original set.

**Theorem A.5** (Cantor). \(|A| < |\mathcal{P}(A)|\).

**Proof.** Suppose there were a bijection \(\varphi : A \to \mathcal{P}(A)\). Then, for \(B = \{x \in A| x \notin \varphi(x)\}\) there must be some \(y\) such that \(B = \varphi(y)\). But \(y \in B\) if and only if \(y \notin \varphi(y) = B\), a contradiction. \(\square\)

**Lemma A.6.** Any infinite set can be written as a disjoint union of countably infinite sets.

**Proof.** Consider collections of disjoint countably infinite subsets. Such collections can be partially ordered by inclusion and hence there is a maximal collection by Zorn’s lemma. If the union of such a maximal collection falls
Note that for example.

**Theorem A.9.** Suppose \( A \) is infinite. Then \(|A \cup B| = \max\{|A|, |B|\}\).

**Proof.** Without loss of generality we can assume \(|B| \leq |A|\) (otherwise exchange both sets). Then \(|A| \leq |A \cup B| \leq |A \cup B| \leq |A \cup B| = |A|\) by the previous corollary. Here \(\cup\) denotes the disjoint union.

A standard theorem proven in every introductory course is that \(\mathbb{N} \times \mathbb{N}\) is countable. The generalization of this result is also true.

**Theorem A.9.** Suppose \( A \) is infinite and \( B \neq \emptyset \). Then \(|A \times B| = \max\{|A|, |B|\}\).

**Proof.** Without loss of generality we can assume \(|B| \leq |A|\) (otherwise exchange both sets). Then \(|A| \leq |A \times B| \leq |A \times A|\) and it suffices to show \(|A \times A| = |A|\).

We proceed as before and consider the set of all bijective functions \(\varphi_{A} : A_{\alpha} \to A_{\alpha} \times A_{\alpha}\) with \(A_{\alpha} \subseteq A\) with the same partial ordering as before. By Zorn’s lemma there is a maximal element \(\varphi_{m}\). Let \(A_{m}\) be its domain and let \(A'_{m} = A_{\alpha} \setminus A_{m}\). We claim that \(|A'_{m}| < |A_{m}|\). If not, \(A'_{m}\) had a subset \(A''_{m}\) with the same cardinality of \(A_{m}\) and hence we had a bijection from \(A'_{m} \to A''_{m} \times A''_{m}\) which could be used to extend \(\varphi\). So \(|A'_{m}| < |A_{m}|\) and thus \(|A| = |A_{m} \cup A'_{m}| = |A_{m}|\). Since we have shown \(|A_{m} \times A_{m}| = |A_{m}|\) the claim follows.

**Example.** Note that for \( A = \mathbb{N} \) we have \(|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|\). Indeed, since \(|\mathbb{R}| = |\mathbb{Z} \times [0, 1)| = |(0, 1)|\) it suffices to show \(|\mathcal{P}(\mathbb{N})| = |(0, 1)|\). To this end note that \(\mathcal{P}(\mathbb{N})\) can be identified with the set of all sequences with values in \(\{0, 1\}\) (the value of the sequence at a point tells us whether it is in the corresponding subset). Now every point in \([0, 1]\) can be mapped to such a sequence via its binary expansion. This map is injective but not surjective since a point can have different binary expansions: \([0, 1] \leq |\mathcal{P}(\mathbb{N})|\). Conversely, given a sequence \(a_{n} \in \{0, 1\}\) we can map it to the number \(\sum_{n=1}^{\infty} a_{n}4^{-n}\). Since this map is again injective (note that we avoid expansions which are eventually 1) we get \(|\mathcal{P}(\mathbb{N})| \leq |[0, 1]|\).
Hence we have
\[ |\mathbb{N}| < |\mathcal{P}(\mathbb{N})| = |\mathbb{R}| \]  
(A.1)
and the continuum hypothesis states that there are no sets whose cardinality lies in between. It was shown by Gödel and Cohen that it, as well as its negation, is consistent with ZFC and hence cannot be decided within this framework.

**Problem A.1.** *Show that Zorn’s lemma implies the axiom of choice.* (Hint: Consider the set of all partial choice functions defined on a subset.)
Bibliography


Glossary of notation
AC$[a,b]$ \ldots absolutely continuous functions, 230
arg(z) \ldots argument of $z \in \mathbb{C}$; $\arg(z) \in (-\pi, \pi]$, $\arg(0) = 0$
$B_r(x)$ \ldots open ball of radius $r$ around $x$, 9
$B(X)$ \ldots Banach space of bounded measurable functions
$BV[a,b]$ \ldots functions of bounded variation, 228
$\mathcal{B}$ = $\mathcal{B}^1$
$\mathcal{B}^n$ \ldots Borel $\sigma$-field of $\mathbb{R}^n$, 156
$\mathbb{C}$ \ldots the set of complex numbers
$C(U)$ \ldots set of continuous functions from $U$ to $\mathbb{C}$
$C(U,V)$ \ldots set of continuous functions from $U$ to $V$
$C_0(U)$ \ldots set of continuous functions vanishing on the boundary $\partial U$, 45
$C_c(U)$ \ldots set of compactly supported continuous functions
$C^k(U)$ \ldots set of $k$ times continuously differentiable functions
$C^\infty_{\text{loc}}(U)$ \ldots set of smooth functions with at most polynomial grow, 268
$C^\infty_c(U)$ \ldots set of compactly supported smooth functions
$C(U,Y)$ \ldots set of continuous functions from $U$ to $Y$, 287
$C^r(U,Y)$ \ldots set of $r$ times continuously differentiable functions, 288
$C^r_c(U,Y)$ \ldots bounded functions in $C^r$, 288
$C^r_c(U,Y)$ \ldots functions in $C^r$ with compact support, 329
c_0(\mathbb{N}) \ldots set of sequences converging to zero, 26
$\mathcal{C}(\delta)$ \ldots set of compact operators, 69
$C(U,Y)$ \ldots set of compact maps from $U$ to $Y$, 320
$CP(f)$ \ldots critical points of $f$, 299
$CS(K)$ \ldots nonempty convex subsets of $K$, 312
$CV(f)$ \ldots critical values of $f$, 299
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi_\Omega(.)$</td>
<td>characteristic function of the set $\Omega$</td>
</tr>
<tr>
<td>$\mathcal{D}(.)$</td>
<td>domain of an operator</td>
</tr>
<tr>
<td>$\deg(D, f, y)$</td>
<td>mapping degree, 299, 307</td>
</tr>
<tr>
<td>$\det$</td>
<td>determinant</td>
</tr>
<tr>
<td>$\dim$</td>
<td>dimension of a linear space</td>
</tr>
<tr>
<td>$\div$</td>
<td>divergence</td>
</tr>
<tr>
<td>$\dist(U,V)$</td>
<td>distance of two sets</td>
</tr>
<tr>
<td>$\mathcal{D}_y(U,Y)$</td>
<td>functions in $\mathcal{C}^r(U,Y)$ which do not attain $y$ on the boundary, 321</td>
</tr>
<tr>
<td>$\mathcal{E}$</td>
<td>exponential function, $e^z = \exp(z)$</td>
</tr>
<tr>
<td>$dF$</td>
<td>derivative of $F$, 287</td>
</tr>
<tr>
<td>$\mathcal{F}(X,Y)$</td>
<td>set of compact finite dimensional functions, 320</td>
</tr>
<tr>
<td>$\GL(n)$</td>
<td>general linear group in $n$ dimensions</td>
</tr>
<tr>
<td>$\Gamma(z)$</td>
<td>gamma function, 188</td>
</tr>
<tr>
<td>$\mathcal{H}(U)$</td>
<td>set of holomorphic functions on a domain $U \subseteq \mathbb{C}$, 297</td>
</tr>
<tr>
<td>$H^1(U,\mathbb{R}^n)$</td>
<td>Sobolev space, 329</td>
</tr>
<tr>
<td>$H^1_0(U,\mathbb{R}^n)$</td>
<td>Sobolev space, 329</td>
</tr>
<tr>
<td>$i$</td>
<td>complex unity, $i^2 = -1$</td>
</tr>
<tr>
<td>$\text{Im}(.)$</td>
<td>imaginary part of a complex number</td>
</tr>
<tr>
<td>$\inf$</td>
<td>infimum</td>
</tr>
<tr>
<td>$J_f(x)$</td>
<td>Jacobi determinant of $f$ at $x$, 299</td>
</tr>
<tr>
<td>$\text{Ker}(A)$</td>
<td>kernel of an operator $A$, 39</td>
</tr>
<tr>
<td>$\lambda^n$</td>
<td>Lebesgue measure in $\mathbb{R}^n$, 182</td>
</tr>
<tr>
<td>$\mathcal{L}(X,Y)$</td>
<td>set of all bounded linear operators from $X$ to $Y$, 41</td>
</tr>
<tr>
<td>$\mathcal{L}(X)$</td>
<td>$= \mathcal{L}(X,X)$</td>
</tr>
<tr>
<td>$L^p(X,d\mu)$</td>
<td>Lebesgue space of $p$ integrable functions, 198</td>
</tr>
<tr>
<td>$L^\infty(X,d\mu)$</td>
<td>Lebesgue space of bounded functions, 198</td>
</tr>
<tr>
<td>$L^p_{\text{loc}}(X,d\mu)$</td>
<td>locally $p$ integrable functions, 199</td>
</tr>
<tr>
<td>$L^1(X)$</td>
<td>space of integrable functions, 175</td>
</tr>
<tr>
<td>$L^2_{\text{cont}}(I)$</td>
<td>space of continuous square integrable functions, 35</td>
</tr>
<tr>
<td>$\ell^p(\mathbb{N})$</td>
<td>Banach space of $p$ summable sequences, 26</td>
</tr>
<tr>
<td>$\ell^2(\mathbb{N})$</td>
<td>Hilbert space of square summable sequences, 32</td>
</tr>
<tr>
<td>$\ell^\infty(\mathbb{N})$</td>
<td>Banach space of bounded summable sequences, 26</td>
</tr>
<tr>
<td>$\max$</td>
<td>maximum</td>
</tr>
<tr>
<td>$\mathbb{N}$</td>
<td>the set of positive integers</td>
</tr>
<tr>
<td>$\mathbb{N}_0$</td>
<td>$= \mathbb{N} \cup {0}$</td>
</tr>
<tr>
<td>$n(\gamma,z_0)$</td>
<td>winding number</td>
</tr>
<tr>
<td>$O(.)$</td>
<td>Landau symbol, $f = O(g)$ iff $\limsup_{x \to x_0}</td>
</tr>
<tr>
<td>$o(.)$</td>
<td>Landau symbol, $f = o(g)$ iff $\lim_{x \to x_0}</td>
</tr>
<tr>
<td>$\mathbb{Q}$</td>
<td>the set of rational numbers</td>
</tr>
<tr>
<td>$\mathbb{R}$</td>
<td>the set of real numbers</td>
</tr>
</tbody>
</table>
Glossary of notation

RV($f$) . . . regular values of $f$, 290
Ran($A$) . . . range of an operator $A$, 39
Re(.) . . . real part of a complex number
$R(I,X)$ . . . set of regulated functions, 290
$\sigma^{n-1}$ . . . surface measure on $S^{n-1}$, 186
$S^{n-1}$ = \{ $x \in \mathbb{R}^n$| $|x| = 1$ \} unit sphere in $\mathbb{R}^n$
$S_n$ = $n\pi^{n/2}/\Gamma(n/2 + 1)$, surface area of the unit sphere in $\mathbb{R}^n$, 186
sign($z$) = $z/|z|$ for $z \neq 0$ and 1 for $z = 0$; complex sign function
$S(I,X)$ . . . simple functions $f : I \rightarrow X$, 290
sup . . . supremum
supp($f$) . . . support of a function $f$, 15
supp($\mu$) . . . support of a measure $\mu$, 161
span($M$) . . . set of finite linear combinations from $M$, 27
$V_n$ = $\pi^{n/2}/\Gamma(n/2 + 1)$, volume of the unit ball in $\mathbb{R}^n$, 187
$\mathbb{Z}$ . . . the set of integers
$I$ . . . identity operator
$\sqrt{z}$ . . . square root of $z$ with branch cut along $(-\infty, 0)$
$z^*$ . . . complex conjugation
$A^*$ . . . adjoint of $A$, 58
$\overline{A}$ . . . closure of $A$, 90
$\hat{f}$ = $\mathcal{F}f$, Fourier transform of $f$, 241
$\check{f}$ = $\mathcal{F}^{-1}f$, inverse Fourier transform of $f$, 244
$|x|$ = $\sqrt{\sum_{j=1}^{n}|x_j|^2}$ Euclidean norm in $\mathbb{R}^n$ or $\mathbb{C}^n$
$|\Omega|$ . . . Lebesgue measure of a Borel set $\Omega$
$\|\cdot\|$ . . . norm, 32
$\|\cdot\|_p$ . . . norm in the Banach space $\ell^p$ and $L^p$, 26, 197
$\langle . , . \rangle$ . . . scalar product in $\mathcal{F}$, 32
$\oplus$ . . . orthogonal sum of vector spaces or operators, 61
$\otimes$ . . . tensor product, 62
$[x]$ = max\{$n \in \mathbb{Z}|n \leq x$\} floor function
$\partial$ . . . gradient
$\partial U$ . . . boundary of the set $U$
$\partial_\alpha$ . . . partial derivative
$\partial_x F(x,y)$ . . . partial derivative with respect to $x$, 287
$M^\perp$ . . . orthogonal complement, 55
$(\lambda_1, \lambda_2)$ = \{ $\lambda \in \mathbb{R} | \lambda_1 < \lambda < \lambda_2$ \}, open interval
$[\lambda_1, \lambda_2]$ = \{ $\lambda \in \mathbb{R} | \lambda_1 \leq \lambda \leq \lambda_2$ \}, closed interval
$x_n \to x$ . . . norm convergence, 24
$x_n \rightharpoonup x$ . . . weak convergence, 106
$A_n \rightarrow A$ . . . norm convergence
$A_n \overset{s}{\rightharpoonup} A$ . . . strong convergence, 109
$A_n \rightharpoonup A$ . . . weak convergence, 109
Index

a.e., see almost everywhere
absolute convergence, 30
absolutely continuous
  function, 230
  measure, 213
absolutely convex, 118
absorbing set, 102
accumulation point, 10
adjoint operator, 99
algebra, 155
almost everywhere, 161
annihilator, 101
approximate identity, 207
Axiom of Choice, 345

B.L.T. theorem, 41
Baire category theorem, 85
balanced set, 122
ball
  closed, 13
  open, 9
Banach algebra, 42, 137
Banach space, 25
Banach–Steinhaus theorem, 86
base, 11
basis, 28
  orthonormal, 51
Bessel inequality, 50
Bessel potential, 252
best reply, 314
bidual space, 96
bijective, 14
Bolzano–Weierstraß theorem, 21
Borel
  function, 168
  measure, 158
regular, 158
set, 156
σ-algebra, 156
boundary condition, 6
boundary point, 10
boundary value problem, 6
bounded
  operator, 39
  sesquilinear form, 37
  set, 20
bounded variation, 228
Brezis–Lieb lemma, 202
Brouwer fixed-point theorem, 309

Calculus of variations, 111
Cantor
  function, 221
  measure, 221
  set, 162
Cauchy sequence, 13
Cauchy transform, 210
Cauchy–Schwarz–Bunjakowski inequality, 33
Cayley transform, 145
chain rule, 288
characteristic function, 172
Chebyshev inequality, 221, 227, 279
Chebyshev polynomials, 74
closed
  ball, 13
  set, 12
closure, 12
cluster point, 10
compact, 17
locally, 21
sequentially, 19
<table>
<thead>
<tr>
<th>Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>compact map</td>
<td>320</td>
</tr>
<tr>
<td>complete</td>
<td>13, 25</td>
</tr>
<tr>
<td>completion</td>
<td>38</td>
</tr>
<tr>
<td>content</td>
<td>237</td>
</tr>
<tr>
<td>continuous</td>
<td>14</td>
</tr>
<tr>
<td>contraction principle</td>
<td>291</td>
</tr>
<tr>
<td>uniform</td>
<td>292</td>
</tr>
<tr>
<td>convergence</td>
<td>13</td>
</tr>
<tr>
<td>convex</td>
<td>24, 199</td>
</tr>
<tr>
<td>convolution</td>
<td>206, 246</td>
</tr>
<tr>
<td>cover</td>
<td>16</td>
</tr>
<tr>
<td>locally finite</td>
<td>17</td>
</tr>
<tr>
<td>refinement</td>
<td>17</td>
</tr>
<tr>
<td>critical value</td>
<td>299</td>
</tr>
<tr>
<td>(C^*) algebra</td>
<td>143</td>
</tr>
<tr>
<td>d’Alembert’s formula</td>
<td>263</td>
</tr>
<tr>
<td>dense</td>
<td>14</td>
</tr>
<tr>
<td>derivative</td>
<td>287</td>
</tr>
<tr>
<td>partial</td>
<td>287</td>
</tr>
<tr>
<td>diffeomorphism</td>
<td>288</td>
</tr>
<tr>
<td>differentiable</td>
<td>287</td>
</tr>
<tr>
<td>differential equations</td>
<td>293</td>
</tr>
<tr>
<td>diffusion equation</td>
<td>3</td>
</tr>
<tr>
<td>dimension</td>
<td>53</td>
</tr>
<tr>
<td>Dirac delta distribution</td>
<td>267</td>
</tr>
<tr>
<td>Dirac measure</td>
<td>161, 177</td>
</tr>
<tr>
<td>direct sum</td>
<td>44</td>
</tr>
<tr>
<td>Dirichlet integral</td>
<td>265</td>
</tr>
<tr>
<td>Dirichlet kernel</td>
<td>64</td>
</tr>
<tr>
<td>discrete set</td>
<td>10</td>
</tr>
<tr>
<td>discrete topology</td>
<td>10</td>
</tr>
<tr>
<td>distance</td>
<td>9, 21</td>
</tr>
<tr>
<td>distribution</td>
<td>329</td>
</tr>
<tr>
<td>distribution function</td>
<td>158</td>
</tr>
<tr>
<td>domain</td>
<td>39</td>
</tr>
<tr>
<td>dominated convergence theorem</td>
<td>176</td>
</tr>
<tr>
<td>double dual</td>
<td>96</td>
</tr>
<tr>
<td>Dynkin system</td>
<td>163</td>
</tr>
<tr>
<td>Dynkin’s (\pi)-(\lambda) theorem</td>
<td>163</td>
</tr>
<tr>
<td>eigenspace</td>
<td>73</td>
</tr>
<tr>
<td>eigenvalue</td>
<td>73</td>
</tr>
<tr>
<td>simple</td>
<td>73</td>
</tr>
<tr>
<td>eigenvector</td>
<td>73</td>
</tr>
<tr>
<td>elliptic equation</td>
<td>339</td>
</tr>
<tr>
<td>embedding</td>
<td>331</td>
</tr>
<tr>
<td>equicontinuous</td>
<td>70</td>
</tr>
<tr>
<td>equilibrium</td>
<td></td>
</tr>
<tr>
<td>Nash</td>
<td>314</td>
</tr>
<tr>
<td>equivalent norms</td>
<td>36</td>
</tr>
<tr>
<td>essential supremum</td>
<td>198</td>
</tr>
<tr>
<td>Extreme value theorem</td>
<td>21</td>
</tr>
<tr>
<td>finite dimensional map</td>
<td>320</td>
</tr>
<tr>
<td>finite intersection property</td>
<td>17</td>
</tr>
<tr>
<td>first countable</td>
<td>11</td>
</tr>
<tr>
<td>first resolvent identity</td>
<td>143</td>
</tr>
<tr>
<td>fixed-point theorem</td>
<td></td>
</tr>
<tr>
<td>Altman</td>
<td>324</td>
</tr>
<tr>
<td>Brouwer</td>
<td>309</td>
</tr>
<tr>
<td>contraction principle</td>
<td>291</td>
</tr>
<tr>
<td>Kakutani</td>
<td>312</td>
</tr>
<tr>
<td>Krasnosel’kii</td>
<td>324</td>
</tr>
<tr>
<td>Rothe</td>
<td>324</td>
</tr>
<tr>
<td>Schauder</td>
<td>323</td>
</tr>
<tr>
<td>form</td>
<td></td>
</tr>
<tr>
<td>bounded</td>
<td>37</td>
</tr>
<tr>
<td>Fourier multiplier</td>
<td>252</td>
</tr>
<tr>
<td>Fourier series</td>
<td>52</td>
</tr>
<tr>
<td>cosine</td>
<td>80</td>
</tr>
<tr>
<td>sine</td>
<td>79</td>
</tr>
<tr>
<td>Fourier transform</td>
<td>241</td>
</tr>
<tr>
<td>Fréchet space</td>
<td>120</td>
</tr>
<tr>
<td>Fredholm alternative</td>
<td>136</td>
</tr>
<tr>
<td>Fredholm operator</td>
<td>136</td>
</tr>
<tr>
<td>Frobenius norm</td>
<td>131</td>
</tr>
<tr>
<td>Fubini theorem</td>
<td>180</td>
</tr>
<tr>
<td>function</td>
<td></td>
</tr>
<tr>
<td>open</td>
<td>15</td>
</tr>
<tr>
<td>fundamental solution</td>
<td></td>
</tr>
<tr>
<td>heat equation</td>
<td>260</td>
</tr>
<tr>
<td>Laplace equation</td>
<td>251</td>
</tr>
<tr>
<td>fundamental theorem of calculus</td>
<td>177, 231</td>
</tr>
<tr>
<td>Galerkin approximation</td>
<td>341</td>
</tr>
<tr>
<td>gamma function</td>
<td>188</td>
</tr>
<tr>
<td>gauge</td>
<td>102</td>
</tr>
<tr>
<td>Gaussian</td>
<td>243</td>
</tr>
<tr>
<td>generalized inverse</td>
<td>189</td>
</tr>
<tr>
<td>gradient</td>
<td>243</td>
</tr>
<tr>
<td>Gram determinant</td>
<td>54</td>
</tr>
<tr>
<td>Gram–Schmidt orthogonalization</td>
<td>52</td>
</tr>
<tr>
<td>graph</td>
<td>89</td>
</tr>
<tr>
<td>graph norm</td>
<td>92</td>
</tr>
<tr>
<td>Green function</td>
<td>79</td>
</tr>
<tr>
<td>Gronwall’s inequality</td>
<td>326</td>
</tr>
<tr>
<td>Hahn decomposition</td>
<td>226</td>
</tr>
<tr>
<td>half-space</td>
<td>115</td>
</tr>
<tr>
<td>Hamel basis</td>
<td>31</td>
</tr>
<tr>
<td>Hankel operator</td>
<td>133</td>
</tr>
<tr>
<td>Hardy space</td>
<td>210</td>
</tr>
<tr>
<td>Hardy–Littlewood maximal function</td>
<td>281</td>
</tr>
<tr>
<td>Hardy–Littlewood maximal inequality</td>
<td>281</td>
</tr>
<tr>
<td>Hardy–Littlewood–Sobolev inequality</td>
<td>282</td>
</tr>
<tr>
<td>Hausdorff space</td>
<td>12</td>
</tr>
<tr>
<td>Hausdorff–Young inequality</td>
<td>277</td>
</tr>
<tr>
<td>heat equation</td>
<td>3, 258</td>
</tr>
<tr>
<td>Heine–Borel theorem</td>
<td>20</td>
</tr>
<tr>
<td>Heisenberg uncertainty principle</td>
<td>247</td>
</tr>
<tr>
<td>Helmholtz equation</td>
<td>252</td>
</tr>
<tr>
<td>Hermitian form</td>
<td>32</td>
</tr>
<tr>
<td>Hilbert space</td>
<td>32</td>
</tr>
<tr>
<td>Dimension</td>
<td>Hölder continuous</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------</td>
</tr>
<tr>
<td>, 53</td>
<td>47</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Identity</th>
<th>Implicit function theorem</th>
<th>Index</th>
<th>Induced topology</th>
<th>Induction Principle</th>
<th>Initial topology</th>
<th>Injective</th>
</tr>
</thead>
<tbody>
<tr>
<td>, 42, 137</td>
<td>293</td>
<td>136</td>
<td>11</td>
<td>345</td>
<td>16</td>
<td>14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inner product</th>
<th>Inner product space</th>
<th>Integrable</th>
<th>Rieman</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>32</td>
<td>174</td>
<td>194</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Integral</th>
<th>Integration by parts</th>
<th>Integration by substitution</th>
<th>Interior</th>
<th>Interior point</th>
</tr>
</thead>
<tbody>
<tr>
<td>172, 290</td>
<td>232, 329</td>
<td>191</td>
<td>12</td>
<td>9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inverse function theorem</th>
<th>Involution</th>
<th>Isolated point</th>
<th>Jensen’s inequality</th>
<th>Jordan curve theorem</th>
<th>Jordan decomposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>293</td>
<td>143</td>
<td>10</td>
<td>200</td>
<td>317</td>
<td>224</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Kakutani’s fixed-point theorem</th>
<th>Kernel</th>
<th>Kirchhoff’s formula</th>
<th>Kuratowski closure axioms</th>
</tr>
</thead>
<tbody>
<tr>
<td>312</td>
<td>39</td>
<td>264</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>λ-system</th>
<th>Ladyzhenskaya inequality</th>
<th>Landau symbols</th>
<th>Lax–Milgram theorem</th>
<th>Nonlinear</th>
<th>Lebesgue</th>
</tr>
</thead>
<tbody>
<tr>
<td>163</td>
<td>332</td>
<td>287</td>
<td>60</td>
<td>339</td>
<td>Decomposition</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Point</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>215</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>161</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>218</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>158</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>268</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lemma</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Riemann-Lebesgue</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>243</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Leray–Schauder principle</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>323</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lidskij trace theorem</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>132</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Limit</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Limit point</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Limsup</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lindelöf theorem</th>
<th>Linear functional</th>
<th>Linear operator</th>
<th>Linearly independent</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>41, 56</td>
<td>39</td>
<td>27</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lipschitz continuous</th>
<th>Locally integrable</th>
<th>Lower semicontinuous</th>
<th>Lyapunov inequality</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>198</td>
<td>169</td>
<td>275</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maximum norm</th>
<th>Mean value theorem</th>
<th>Measurable function</th>
<th>Measurable set</th>
<th>Measurable space</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>288</td>
<td>167</td>
<td>156</td>
<td>156</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Absolutely continuous</th>
<th>Complete</th>
<th>Complex</th>
<th>Finite</th>
<th>Lebesgue</th>
</tr>
</thead>
<tbody>
<tr>
<td>213</td>
<td>166</td>
<td>222</td>
<td>156</td>
<td>161</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Minimal support</th>
<th>Mutually singular</th>
<th>Product</th>
<th>Space</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>220</td>
<td>213</td>
<td>179</td>
<td>156</td>
<td>161</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Metric</th>
<th>Metric translation invariant</th>
<th>Minkowski functional</th>
<th>Minkowski inequality</th>
<th>Integral form</th>
<th>Integral</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>120</td>
<td>102</td>
<td>202</td>
<td>202</td>
<td>202</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mollifier</th>
<th>Monotone</th>
<th>Monotone convergence theorem</th>
<th>Morrey inequality</th>
<th>Multi-index</th>
</tr>
</thead>
<tbody>
<tr>
<td>207</td>
<td>338</td>
<td>173</td>
<td>256</td>
<td>46, 243</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Multi-index</th>
<th>Order</th>
<th>Multilinear function</th>
<th>Symmetric</th>
<th>Multiplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>243</td>
<td>46, 243</td>
<td>289</td>
<td>289</td>
<td>Fourier multiplier</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutually singular measures</th>
<th>Nash equilibrium</th>
<th>Nash theorem</th>
<th>Navier-Stokes equation</th>
<th>Stationary</th>
</tr>
</thead>
<tbody>
<tr>
<td>213</td>
<td>314</td>
<td>315</td>
<td>328</td>
<td>328</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neighborhood</th>
<th>Neighborhood base</th>
<th>Neumann series</th>
<th>Noether operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>11</td>
<td>141</td>
<td>136</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Norm</th>
<th>Operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>39</td>
</tr>
</tbody>
</table>
Index

normal, 22, 143
normalized, 33
normed space, 24
nowhere dense, 85
\(n\)-person game, 313
null space, 39
onto, 14
open
  ball, 9
  function, 15
  set, 10
operator
  adjoint, 58
  bounded, 39
  closeable, 90
  closed, 90
  closure, 90
  compact, 69
  domain, 39
  finite rank, 127
  Hilbert–Schmidt, 211
  linear, 39
  nonnegative, 59
  self-adjoint, 73
  strong convergence, 109
  symmetric, 73
  unitary, 54
  weak convergence, 109
order
  partial, 345
  total, 345
  well, 345
orthogonal, 33
  complement, 55
  projection, 55, 147
  sum, 61
outer measure, 164
parallel, 33
parallelogram law, 34
Parseval relation, 52
partial order, 345
partition, 192
partition of unity, 22
payoff, 313
Peano theorem, 325
perpendicular, 33
\(\pi\)-system, 163
Plancherel identity, 245
Poincaré inequality, 330
Poincaré-Friedrichs inequality, 329
Poisson equation, 249
Poisson kernel, 210
Poisson’s formula, 265
polar coordinates, 185
polar decomposition, 126
polarization identity, 34
powerset, 347
premeasure, 157
prisoners dilemma, 314
probability measure, 156
product measure, 179
product rule, 233
product topology, 16
projection-valued measure, 147
proper map, 321
proper metric space, 21
pseudo metric, 9
Pythagorean theorem, 33
quadrangle inequality, 23
quasinorm, 31
quotient space, 44
Radon measure, 171
Radon–Nikodym
  derivative, 215
  theorem, 215
range, 39
rank, 127
Rayleigh–Ritz method, 81
reduction property, 315
refinement, 17
reflexive, 97
regular value, 299
regulated function, 290
relative \(\sigma\)-algebra, 156
relative topology, 11
relatively compact, 17
Rellich’s compactness theorem, 331
resolution of the identity, 148
resolvent, 78, 139
resolvent identity
  first, 143
  resolvent set, 139
Riemann integrable, 194
Riemann integral, 194
  lower, 193
  upper, 193
Riesz lemma, 56
Riesz potential, 250
Ritz method, 81
Rouché’s theorem, 298
Sard’s theorem, 303
scalar product, 32
Schatten \(p\)-class, 129
Schauder basis, 28
Schrödinger equation, 261
Schur criterion, 210
Schur test, 212
Schwartz space, 120, 243
second countable, 11
Index

self-adjoint, 58, 143

semigroup
    generator, 259
    strongly continuous, 258

seminorm, 24

separable, 14, 28

separated seminorms, 117

separation of variables, 4

sequentially closed, 13

sequentially continuous, 15

series
    absolutely convergent, 30

sesquilinear form, 32

    bounded, 37
    parallelogram law, 37
    polarization identity, 37
    $\sigma$-algebra, 155

$\sigma$-finite, 156

signed measure, 224

simple function, 172, 290

sine integral, 264

singular values, 125

Sobolev inequality, 257

Sobolev space, 254

span, 27

spectral measure, 146

spectral projections, 147

spectral radius, 141

spectrum, 139

spherical coordinates, 185

spherically symmetric, 249

$\ast$-subalgebra, 143

Stokes theorem, 305

Stone–Weierstraß theorem, 150

strategy, 313

strong convergence, 109

strong type, 280

Sturm–Liouville problem, 6

subadditive, 280

subcover, 16

subspace topology, 11

support, 15

    distribution, 272
    measure, 161

surjective, 14

tempered distributions, 120, 266


tensor product, 62

theorem

    Altman, 324
    Arzelà–Ascoli, 71
    B.L.T., 41
    Bair, 85
    Banach–Alaoglu, 114
    Banach–Steinhaus, 86
    Bolzano–Weierstraß, 21
    Brezis–Lieb, 202

closed graph, 89

dominated convergence, 176

Dykin’s $\pi$-$\lambda$, 163

Fatou, 174, 176

Fatou–Lebesgue, 176

Fubini, 180

fundamental thm. of calculus, 177, 231

Gelfand–Mazur, 140

Hadamard three-lines, 276

Hahn–Banach, 95

Hahn–Banach, geometric, 104

Heine–Borel, 20

Hellinger–Toeplitz, 92

implicit function, 293

inverse function, 293

Jordan, 229

Jordan–von Neumann, 34

Kolmogorov, 121

Krasnosel’skiĭ, 324

Lax–Milgram, 60, 339

Lebesgue, 176, 194

Lebesgue decomposition, 215

Leray–Schauder, 323

Levi, 173

Lindelöf, 16

Marcinkiewicz, 280

monotone convergence, 173

Nash, 315

open mapping, 88, 89

Peano, 325

Plancherel, 245

Pythagorean, 33

Radon–Nikodym, 215

Rellich, 331

Riesz representation, 238

Riesz–Fischer, 204

Riesz–Thorin, 276

Rothé, 324

Rouché, 298

Sard, 303

Schauder, 100, 323

Schröder–Bernstein, 347

Schur, 210

Sobolev embedding, 256

spectral, 144

spectral mapping, 140

Stone–Weierstraß, 150

Tietze, 22, 310

Tonelli, 181

Tychonoff, 19

Urysohn, 21

Weierstraß, 21, 29

Wiener, 249

Zorn, 346

Tietze extension theorem, 310

Tonelli theorem, 181

topological space, 10
topological vector space, 24
    topology
      base, 11
      product, 16
      relative, 11
    total, 28
    total order, 345
    total variation, 222, 228
    trace, 131
      class, 129
    trace topology, 11
    transport equation, 265
    triangle inequality, 9, 24
      inverse, 9, 24
    trivial topology, 10

    uncertainty principle, 247
    uniform boundedness principle, 86
    uniform contraction principle, 292
    uniformly convex space, 37
    unit sphere, 186
    unit vector, 33
    unitary, 143
    upper semicontinuous, 169
    Urysohn lemma, 21

    vague convergence
      measures, 239
      Vandermonde determinant, 31
    variation, 228
      Vitali set, 162

    wave equation, 5, 262
    weak
      derivative, 255
    weak $L^p$, 279
    weak convergence, 106
    Weak solution, 340
    weak solution, 333
    weak topology, 106, 114
    weak type, 280
    weak-$*$ convergence, 111
    weak-$*$ topology, 114
    Weierstraß approximation, 29
    Weierstraß theorem, 21
    well-order, 345
    Weyl asymptotic, 83
    Wiener covering lemma, 217
    winding number, 297

    Young inequality, 207, 246, 277, 278

    ZFC, 345
    Zorn’s lemma, 346