Theoretical and Mathematical Physics

The series founded in 1975 and formerly (until 2005) entitled *Texts and Monographs in Physics* (TMP) publishes high-level monographs in theoretical and mathematical physics. The change of title to *Theoretical and Mathematical Physics* (TMP) signals that the series is a suitable publication platform for both the mathematical and the theoretical physicist. The wider scope of the series is reflected by the composition of the editorial board, comprising both physicists and mathematicians.

The books, written in a didactic style and containing a certain amount of elementary background materials, bridge the gap between advanced textbooks and research monographs. They can thus serve as basis for advanced studies, not only for lectures and seminars at graduate level, but also for scientists entering a field of research.

Jiří Blank • Pavel Exner • Miloslav Havlíček

Hilbert Space Operators in Quantum Physics

Second Edition

AIP PRESS

Springer

Jiří Blank†
Charles University
Prague
Czechia

Pavel Exner
Doppler Institute
Břehová 7
11519 Prague
and Nuclear Physics Institute
Czech Academy of Sciences
25068 Řež near Prague
Czech Republic
exner@ujf.cas.cz

Miloslav Havlíček
Doppler Institute
and Faculty of Nuclear Sciences
and Physical Engineering
Czech Technical University
Trojanova 13
12000 Prague
Czech Republic
havlicek@fjfi.cvut.cz

ISBN 978-1-4020-8869-8 e-ISBN 978-1-4020-8870-4

Library of Congress Control Number: 2008933703

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

To our wives and daughters

Preface to the second edition

Almost fifteen years later, and there is little change in our motivation. Mathematical physics of quantum systems remains a lively subject of intrinsic interest with numerous applications, both actual and potential.

In the preface to the first edition we have described the origin of this book rooted at the beginning in a course of lectures. With this fact in mind, we were naturally pleased to learn that the volume was used as a course text in many points of the world and we gladly accepted the offer of *Springer Verlag* which inherited the rights from our original publisher, to consider preparation of a second edition.

It was our ambition to bring the reader close to the places where real life dwells, and therefore this edition had to be more than a corrected printing. The field is developing rapidly and since the first edition various new subjects have appeared; as a couple of examples let us mention quantum computing or the major progress in the investigation of random Schrödinger operators. There are, however, good sources in the literature where the reader can learn about these and other new developments.

We decided instead to amend the book with results about new topics which are less well covered, and the same time, closer to the research interests of one of us. The main change here are two new chapters devoted to quantum waveguides and quantum graphs. Following the spirit of this book we have not aspired to full coverage — each of these subjects would deserve a separate monograph — but we have given a detailed enough exposition to allow the interested reader to follow (and enjoy) fresh research results in this area. In connection with this we have updated the list of references, not only in the added chapters but also in other parts of the text in the second part of the book where we found it appropriate.

Naturally we have corrected misprints and minor inconsistencies spotted in the first edition. We thank the colleagues who brought them to our attention, in particular to Jana Stará, who indicated numerous improvements. As with the first edition, we have asked a native speaker to try to remove the foreign "accent" from our writing; we are grateful to Mark Harmer for accepting this role.

Prague, December 2007

Pavel Exner
Miloslav Havlíček

Preface

Relations between mathematics and physics have a long and entangled tradition. In spite of repeated clashes resulting from the different aims and methods of the two disciplines, both sides have always benefitted. The place where contacts are most intensive is usually called mathematical physics, or if you prefer, physical mathematics. These terms express the fact that mathematical methods are needed here more to understand the essence of problems than as a computational tool, and conversely, the investigated properties of physical systems are also inspiring from the mathematical point of view.

In fact, this field does not need any advocacy. When A. Wightman remarked a few years ago that it had become "socially acceptable", it was a pleasant understatement; by all accounts, mathematical physics is flourishing. It has long left the adolescent stage when it cherished only oscillating strings and membranes; nowadays it has built synapses to almost every part of physics. Evidence that the discipline is developing actively is provided by the fruitful oscillation between the investigation of particular systems and synthetizing generalizations, as well as by discoveries of new connections between different branches.

The drawback of this rapid development is that it has become virtually impossible to write a textbook on mathematical physics as a single topic. There are, of course, books which cover a wide range of problems, some of them indeed monumental, but even they are like cities which govern the territory while watching the frontier slowly moving towards the gray distance. This is simply the price we have to pay for the flood of ideas, concepts, tools, and results that our science is producing.

It was not our aim to write a poor man's version of some of the big textbooks. What we want is to give students basic information about the field, by which we mean an amount of knowledge that could constitute the basis of an intensive one–year course for those who already have the necessary training in algebra and analysis, as well as in classical and quantum mechanics. If our exposition should kindle interest in the subject, the student will be able, after taking such a course, to read specialized monographs and research papers, and to discover a research topic to his or her taste. We have mentioned that the span of the contemporary mathematical physics is vast; nevertheless the cornerstone remains where it was laid by J. von Neumann, H. Weyl, and the other founding fathers, namely in regions connected with quantum theory. Apart from its importance for fundamental problems such as the constitution of matter, this claim is supported by the fact that quantum theory is gradually becoming a basis for most branches of *applied* physics, and has in this way entered our everyday life.

The mathematical backbone of quantum physics is provided by the theory of linear operators on Hilbert spaces, which we discuss in the first half of this book. Here we follow a well–trodden path; this is why references in this part aim mostly at standard book sources, even for the few problems which maybe go beyond the standard curriculum. To make the exposition self–contained without burdening the main text, we have collected the necessary information about measure theory, integration, and some algebraic notions in the appendices.

The physical chapters in the second half are not intended to provide a self–contained exposition of quantum theory. As we have remarked, we suppose that the reader has background knowledge up to the level of a standard quantum mechanics course; the present text should rather provide new insights and help to reach a deeper understanding. However, we attempt to describe the mathematical foundations of quantum theory in a sufficiently complete way, so that a student coming from mathematics can start his or her way into this part of physics through our book.

In connection with the intended purpose of the text, the character of referencing changes in the second part. Though the material discussed here is with a few exceptions again standard, we try in the notes to each chapter to explain extensions of the discussed results and their relations to other problems; occasionally we have set traps for the reader's curiosity. The notes are accompanied by a selective but quite broad list of references, which map ways to the areas where real life dwells.

Each chapter is accompanied by a list of problems. Solving at least some of them in full detail is the safest way for the reader to check that he or she has indeed mastered the topic. The problem level ranges from elementary exercises to fairly complicated proofs and computations. We have refrained from marking the more difficult ones with asterisks because such a classification is always subjective, and after all, in real life you also often do not know in advance whether it will take you an hour or half a year to deal with a given problem.

Let us add a few words about the history of the book. It originates from courses of lectures we have given in different forms during the past two decades at Charles University and the Czech Technical University in Prague. In the 1970s we prepared several volumes of lecture notes; ten years later we returned to them and rewrote the material into a textbook, again in Czech. It was prepared for publication in 1989, but the economic turmoil which inevitably accompanied the welcome changes delayed its publication, so that it appeared only recently.

In the meantime we suffered a heavy blow. Our friend and coauthor, Jiří Blank, died in February 1990 at the age of 50. His departure reminded us of the bitter truth that we usually are able to appreciate the real value of our relationships with fellow humans only after we have lost them. He was always a stabilizing element of our triumvirate of authors, and his spirit as a devoted and precise teacher is felt throughout this book; we hope that in this indirect way his classes will continue.

Preparing the English edition was therefore left to the remaining two authors. It has been modified in many places. First of all, we have included two chapters and some other material which was prepared for the Czech version but then left out due to editorial restrictions. Though the aim of the book is not to report on the present state of research, as we have already remarked, the original manuscript was finished four years ago and we felt it was necessary to update the text and references in some places. On the other hand, since the audience addressed by the English text is different — and is equipped with different libraries — we decided to rewrite certain parts from the first half of the book in a more condensed form.

One consequence of these alterations was that we chose to do the translation ourselves. This decision contained an obvious danger. If you write in a language which you did not master during your childhood, the result will necessarily contain some unwanted comical twists reminiscent of the famous character of Leo Rosten. We are indebted to P. Moylan and, in particular, to R. Healey, who have read the text and counteracted our numerous petty attacks against the English language; those clumsy expressions that remain are, of course, our own.

There are many more people who deserve our thanks: coauthors of our research papers, colleagues with whom we have had the pleasure of exchanging ideas, and simply friends who have supported us during difficult times. We should not forget about students in our courses who have helped just by asking questions; some of them have now become our colleagues. In view of the book complex history, the list should be very long. We prefer to thank all of them anonymously. However, since every rule should have an exception, let us name J. Dittrich, who read the manuscript and corrected numerous mistakes. Last but not least we want to thank our wives, whose patience and understanding made the writing of this book possible.

Prague, July 1993

Pavel Exner
Miloslav Havlíček

Contents

Chapter 1

Some notions from functional analysis

1.1 Vector and normed spaces

The notion of a vector space is obtained by axiomatization of the properties of the three–dimensional space of Euclidean geometry, or of configuration spaces of classical mechanics. A **vector** (or *linear*) **space** V is a set $\{x, y, \dots\}$ equipped with the operations of summation, $[x, y] \mapsto x + y \in V$, and multiplication by a complex or real number α, $[\alpha, x] \mapsto \alpha x \in V$, such that

(i) The summation is commutative, $x + y = y + x$, and associative, $(x + y) + z = x + (y + z)$. There exist a zero element $0 \in V$, and an inverse element $-x \in V$, to any $x \in V$ so that $x + 0 = x$ and $x + (-x) = 0$ holds for all $x \in V$.

(ii) $\alpha(\beta x) = (\alpha\beta)x$ and $1x = x$.

(iii) The summation and multiplication are distributive, $\alpha(x + y) = \alpha x + \alpha y$ and $(\alpha + \beta)x = \alpha x + \beta x$.

The elements of V are called *vectors*. The set of numbers (or *scalars*) in the definition can be replaced by any algebraic field F. Then we speak about a vector space over F, and in particular, about a *complex* and *real* vector space for $F = \mathbb{C}, \mathbb{R}$, respectively. A vector space without further specification in this book always means a complex vector space.

1.1.1 Examples: (a) The space $\mathbb{C}^n$ consists of n–tuples of complex numbers with the summation and scalar multiplication defined componentwise. In the same way, we define the real space $\mathbb{R}^n$.

(b) The space ℓ^p, $1 \le p \le \infty$, of all complex sequences $X := \{\xi_j\}_{j=1}^{\infty}$ such that $\sum_{j=1}^{\infty} |\xi_j|^p < \infty$ for $p < \infty$ and $\sup_j |\xi_j| < \infty$ if $p = \infty$, with the summation and scalar multiplication defined as above; the *Minkowski inequality* implies $X + Y \in \ell^p$ for $X, Y \in \ell^p$ (Problem 2).

(c) The space $C(J)$ of continuous complex functions on a closed interval $J \subset \mathbb{R}$ with $(\alpha f+g)(x) := \alpha f(x)+g(x)$. In a similar way, we define the space $C(X)$ of continuous functions on a compact X and spaces of *bounded* continuous functions on more general topological spaces (see the next two sections).

A *subspace* $L \subset V$ is a subset, which is itself a vector space with the same operations. A minimal subspace containing a given subset $M \subset V$ is called the *linear hull (envelope)* of M and denoted as M_{lin} or $\mathrm{lin}(M)$. Vectors $x_1, \dots, x_n \in V$ are **linearly independent** if $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$ implies that all the numbers $\alpha_1, \dots, \alpha_n$ are zero; otherwise they are *linearly dependent*, which means some of them can be expressed as a linear combination of the others. A set $M \subset V$ is *linearly independent* if each of its finite subsets consists of linearly independent vectors.

This allows us to introduce the *dimension* of a vector space V as a maximum number of linearly independent vectors in V. Among the spaces mentioned in Example 1.1.1, $\mathbb{C}^n$ and $\mathbb{R}^n$ are n–dimensional ($\mathbb{C}^n$ is $2n$–dimensional as a real vector space) while the others are infinite–dimensional. A *basis* of a finite–dimensional V is any linearly independent set $B \subset V$ such that $B_{lin} = V$; it is clear that $\dim V = n$ *iff* V has a basis of n elements. Vector spaces V, V' are said to be (algebraically) *isomorphic* if there is a bijection $f : V \to V'$, which is linear, $f(\alpha x + y) = \alpha f(x) + f(y)$. Isomorphic spaces have the same dimension; for finite–dimensional spaces the converse is also true (Problem 3).

There are various ways to construct new vector spaces from given ones. Let us mention two of them:

(i) If $V_1, \dots, V_N$ are vector spaces over the same field; then we can equip the Cartesian product $V := V_1 \times \cdots \times V_N$ with a summation and scalar multiplication defined by $\alpha[x_1, \dots, x_N] + [y_1, \dots, y_N] := [\alpha x_1 + y_1, \dots, \alpha x_N + y_N]$. The axioms are obviously satisfied; the resulting vector space is called the *direct sum* of $V_1, \dots, V_N$ and denoted as $V_1 \oplus \cdots \oplus V_N$ or $\sum_j^{\oplus} V_j$. The same term and symbols are used if $V_1, \dots, V_N$ are subspaces of a given space V such that each $x \in V$ has a unique decomposition $x = x_1 + \cdots + x_N$, $x_j \in V_j$.

(ii) If W is a subspace of a vector space V, we can introduce an equivalence relation on V by $x \sim y$ if $x - y \in W$. Defining the vector–space operations on the set $\tilde{V}$ of equivalence classes by $\alpha\tilde{x} + \tilde{y} := (\alpha x + y)\tilde{}$ for some $x \in \tilde{x}$, $y \in \tilde{y}$, we get a vector space, which is called the *factor space* of V with respect to W and denoted as V/W.

1.1.2 Example: The space $\mathcal{L}^p(M, d\mu)$, $p \ge 1$, where μ is a non–negative measure, consists of all measurable functions $f : M \to \mathbb{C}$ satisfying $\int_M |f|^p d\mu < \infty$ with pointwise summation and scalar multiplication — *cf.* Appendix A.3. The subset $\mathcal{L}_0 \subset \mathcal{L}^p$ of the functions such that $f(x) = 0$ for μ–almost all $x \in M$ is easily seen to be a subspace; the corresponding factor space $L^p(M, d\mu) := \mathcal{L}^p(M, d\mu)/\mathcal{L}_0$ is then formed by the classes of μ–equivalent functions.

A map $f : V \to \mathbb{C}$ on a vector space V is called a *functional*; if it maps into the reals we speak about a *real* functional. A functional f is *additive* if $f(x+y) = f(x)+f(y)$ holds for all $x, y \in V$, and *homogeneous* if $f(\alpha x) = \alpha f(x)$ or *antihomogeneous* if $f(\alpha x) = \bar{\alpha} f(x)$ for $x \in V$, $\alpha \in \mathbb{C}$. An additive (anti)homogeneous functional is called *(anti)linear*. A real functional p is called a *seminorm* if $p(x+y) \le p(x)+p(y)$ and $p(\alpha x) = |\alpha| p(x)$ holds for any $x, y \in V$, $\alpha \in \mathbb{C}$; this definition implies that p maps V into $\mathbb{R}^+$ and $|p(x)-p(y)| \le p(x-y)$. The following important result is valid (see the notes to this chapter).

1.1.3 Theorem (Hahn–Banach): Let p be a seminorm on a vector space V. Any linear functional f_0 defined on a subspace $V_0 \subset V$ and fulfilling $|f_0(y)| \le p(y)$ for all $y \in V_0$ can be extended to a linear functional f on V such that $|f(x)| \le p(x)$ holds for any $x \in V$.

A map $F := V \times \cdots \times V \to \mathbb{C}$ is called a *form*, in particular, a *real* form if its range is contained in $\mathbb{R}$. A form $F : V \times V \to \mathbb{C}$ is *bilinear* if it is linear in both arguments, and *sesquilinear* if it is linear in one of them and antilinear in the other. Most frequently we shall drop the adjective when speaking about sesquilinear forms; we shall use the "physical" convention assuming that such a form is antilinear in the *left* argument. For a given F we define the *quadratic form* (generated by F) by $q_F : q_F(x) = F(x,x)$; the correspondence is one–to–one as the *polarization formula*

$$F(x,y) = \frac{1}{4}\Big(q_F(x+y) - q_F(x-y)\Big) - \frac{i}{4}\Big(q_F(x+iy) - q_F(x-iy)\Big)$$

shows. A form F is *symmetric* if $F(x,y) = \overline{F(y,x)}$ for all $x, y \in V$; it is *positive* if $q_F(x) \ge 0$ for any $x \in V$ and *strictly positive* if, in addition, $F(x) = 0$ holds for $x = 0$ only. A positive form is symmetric (Problem 6) and fulfils the **Schwarz inequality**,

$$|F(x,y)|^2 \le q_F(x) q_F(y) .$$

A *norm* on a vector space V is a seminorm $\|\cdot\|$ such that $\|x\| = 0$ holds *iff* $x = 0$. A pair $(V, \|\cdot\|)$ is called a **normed space**; if there is no danger of misunderstanding we shall speak simply about a normed space V.

1.1.4 Examples: (a) In the spaces $\mathbb{C}^n$ and $\mathbb{R}^n$, we introduce

$$\|x\|_\infty := \max_{1\le j\le n} |\xi_j| \quad \text{and} \quad \|x\|_p := \left(\sum_{j=1}^n |\xi_j|^p\right)^{1/p}, \; p \ge 1 ,$$

for $x = \{\xi_1, \dots, \xi_n\}$; the norm $\|\cdot\|_2$ on $\mathbb{R}^n$ is often also denoted as $|\cdot|$. Analogous norms are used in ℓ^p (see also Problem 8).

(b) In $L^p(M, d\mu)$, we introduce

$$\|f\|_p := \left(\int_M |f|^p d\mu\right)^{1/p} .$$

The relation $\|f\|_p=0$ implies $f(x)=0$ μ–a.e. in M, so f is the zero element of $L^p(M,d\mu)$. If we speak about $L^p(M,d\mu)$ as a normed space, we always have in mind this natural norm though it is not, of course, the only possibility. If the measure μ is discrete with countable support, $L^p(M,d\mu)$ is isomorphic to ℓ^p and we recover the norm $\|\cdot\|_p$ of the previous example.

(c) By $L^\infty(M,d\mu)$ we denote the set of classes of μ–equivalent functions $f: M\to\mathbb{C}$, which are bounded a.e., *i.e.*, there is $c>0$ such that $|f(x)|\le c$ for μ–almost all $x\in M$. The infimum of all such numbers is denoted as $\sup\operatorname{ess}|f(x)|$. We can easily check that $L^\infty(M,d\mu)$ is a vector space and $f\mapsto \|f\|_\infty := \sup\operatorname{ess}_{x\in M}|f(x)|$ is a norm on it.

(d) The space $C(X)$ can be equipped with the norm $\|f\|_\infty := \sup_{x\in X}|f(x)|$.

A strictly positive sesquilinear form on a vector space V is called an **inner** (or **scalar**) **product**. In other words, it is a map $(\cdot,\cdot)$ from $V\times V$ to $\mathbb{C}$ such that the following conditions hold for any $x,y,z\in V$ and $\alpha\in\mathbb{C}$:

(i) $(x,\alpha y+z)=\alpha(x,y)+(x,z)$

(ii) $(x,y)=\overline{(y,x)}$

(iii) $(x,x)\ge 0$ and $(x,x)=0$ *iff* $x=0$

A vector space with an inner product is called a **pre–Hilbert space**. Any such space is at the same time a normed space with the norm $\|x\| := \sqrt{(x,x)}$; the *Schwarz inequality* then assumes the form

$$|(x,y)| \le \|x\|\,\|y\|\,.$$

The above norm is said to be *induced by the inner product*. Due to conditions (i) and (ii) it fulfils the *parallelogram identity,*

$$\|x+y\|^2+\|x-y\|^2 = 2\|x\|^2+2\|y\|^2\,;$$

on the other hand, it allows us to express the inner product by polarization,

$$(x,y) = \frac{1}{4}\Big(\|x+y\|^2-\|x-y\|^2\Big) - \frac{i}{4}\Big(\|x+iy\|^2-\|x-iy\|^2\Big)\,.$$

These properties are typical for a norm induced by an inner product (Problem 11).

Vectors x,y of a pre–Hilbert space V are called *orthogonal* if $(x,y)=0$. A vector x is *orthogonal to a set* M if $(x,y)=0$ holds for all $y\in M$; the set of all such vectors is denoted as $M^\perp$ and called the *orthogonal complement* to M. Inner–product linearity implies that it is a subspace, $(M^\perp)_{lin}=M^\perp$, with the following simple properties

$$(M_{lin})^\perp = M^\perp\,,\quad M_{lin}\subset (M^\perp)^\perp\,,\quad M\subset N \Rightarrow M^\perp\supset N^\perp\,.$$

A set M of nonzero vectors whose every two elements are orthogonal is called an *orthogonal set*; in particular, M is *orthonormal* if $\|x\| = 1$ for each $x \in M$. Any orthonormal set is obviously linearly independent, and in the opposite direction we have the following assertion, the proof of which is left to the reader.

1.1.5 Theorem (Gram-Schmidt): Let N be an at most countable linearly independent set in a pre–Hilbert space V, then there is an orthonormal set $M \subset V$ of the same cardinality such that $M_{lin} = N_{lin}$.

1.2 Metric and topological spaces

A *metric* on a set X is a map $\varrho : X \times X \to [0, \infty)$, which is symmetric, $\varrho(x,y) = \varrho(y,x)$, $\varrho(x,y) = 0$ *iff* $x = y$, and fulfils the *triangle inequality,*

$$\varrho(x,z) \le \varrho(x,y) + \varrho(y,z) \,,$$

for any $x, y, z \in X$; the pair (X, ϱ) is called a **metric space** (we shall again for simplicity often use the symbol X only). If X is a normed space, one can define a metric on X by $\varrho(x,y) := \|x - y\|$; we say it is *induced by the norm* (see also Problems 15 and 16).

Let us first recall some basic notions and properties of metric spaces. An ε–*neighborhood* of a point $x \in X$ is the open ball $U_\varepsilon(x) := \{\, y \in X \,:\, \varrho(y,x) < \varepsilon \}$. A point x is an *interior point* of a set M if there is a $U_\varepsilon(x) \subset M$. A set is *open* if all its points are interior points, in particular, any neighborhood of a given point is open. A union of an arbitrary family of open sets is again an open set; the same is true for *finite* intersections of open sets.

The *closure* $\overline{M}$ of a set M is the family of all points $x \in X$ such that the intersection $U_\varepsilon(x) \cap M \ne \emptyset$ for any $\varepsilon > 0$. A point $x \in \overline{M}$ is called *isolated* if there is $U_\varepsilon(x)$ such that $U_\varepsilon(x) \cap M = \{x\}$, otherwise x is a *limit* (or *accumulation*) point of M. The closure points of M which are not interior form the *boundary* $\mathrm{bd}\, M$ of M. A set is *closed* if it coincides with its closure, and $\overline{M}$ is *the smallest closed set* containing M (*cf.* Problem 17). In particular, the whole X and the empty set $\emptyset$ are closed and open at the same time.

A set M is said to be *dense in a set* $N \subset X$ if $\overline{M} \supset N$; it is *everywhere dense* if $\overline{M} = X$ and *nowhere dense* if $X \setminus \overline{M}$ is everywhere dense. A metric space which contains a countable everywhere dense set is called *separable*. An example is the space $\mathbb{C}^n$ with any of the norms of Example 1.1.4a where a dense set is formed, *e.g.*, by n–tuples of complex numbers with rational real and imaginary parts; other examples will be given in the next chapter (see also Problem 18).

A sequence $\{x_n\} \subset X$ *converges to a point* $x \in X$ if to any $U_\varepsilon(x)$ there is n_0 such that $x_n \in U_\varepsilon(x)$ holds for all $n > n_0$. Since any two mutually different points $x, y \in X$ have disjoint neighborhoods, each sequence has at most one limit. Sequences can also be used to characterize closure of a set (Problem 17).

Next we recall a few notions related to maps $f : X \to X'$ of metric spaces. The map f is *continuous at a point* $x \in X$ if to any $U'_\varepsilon(f(x))$ there is a $U_\delta(x)$ such that $f(U_\delta(x)) \subset U'_\varepsilon(f(x))$; alternatively we can characterize the local continuity using sequences (Problem 19). On the other hand, f is (*globally*) *continuous* if the pull–back $f^{(-1)}(G')$ of any open set $G' \subset X'$ is open in X.

An important class of continuous maps is represented by *homeomorphisms*, *i.e.*, bijections $f : X \to X'$ such that both f and f^{-1} are continuous. It is clear that in this way any family of metric spaces can be divided into equivalence classes. A homeomorphism maps, in particular, the family τ of open sets in X bijectively onto the family τ' of open sets in X'; we say that homeomorphic metric spaces are *topologically equivalent*. Such spaces can still differ in metric properties. As an example, consider the spaces $\mathbb{R}$ and $(-\frac{\pi}{2}, \frac{\pi}{2})$ with the same metric $\varrho(x,y) := |x-y|$; they are homeomorphic by $x \mapsto \arctan x$ but only the first of them contains unbounded sets. A bijection $f := X \to X'$ which preserves the metric properties, $\varrho'(f(x), f(y)) = \varrho(x,y)$, is called *isometry*; this last named property implies continuity, so any isometry is a homeomorphism.

A homeomorphism $f : V \to V'$ of normed spaces is called *linear homeomorphism* if it is simultaneously an isomorphism. Linearly homeomorphic spaces therefore also have the same algebraic structure; this is particularly simplifying in the case of finite dimension (Problem 21). In addition, if the identity $\|f(x)\|_{V'} = \|x\|_V$ holds for any $x \in V$ we speak about a *linear isometry*.

A sequence $\{x_n\}$ in a metric space X is called *Cauchy* if to any $\varepsilon > 0$ there is n_ε such that $\varrho(x_n, x_m) < \varepsilon$ for all $n, m > n_\varepsilon$. In particular, any convergent sequence is Cauchy; a metric space in which the converse is also true is called *complete*. Completeness is one of the basic "nontopological" properties of metric spaces: recall the spaces $\mathbb{R}$ and $(-\frac{\pi}{2}, \frac{\pi}{2})$ mentioned above; they are homeomorphic but only the first of them is complete.

1.2.1 Example: Let us check the *completeness of* $L^p(M, d\mu)\,, p \ge 1$, *with a* σ*–finite measure* μ. Suppose first $\mu(M) < \infty$ and consider a Cauchy sequence $\{f_n\} \subset L^p$. By the Hölder inequality, it is Cauchy also with respect to $\|\cdot\|_1$, so for any $\varepsilon > 0$ there is $N(\varepsilon)$ such that $\|f_n - f_m\|_1 < \varepsilon$ for $n, m > N(\varepsilon)$. We pick a subsequence, $g_n := f_{k_n}$, by choosing $k_1 := N(2^{-1})$ and $k_{n+1} := \max\{k_n + 1,\, N(2^{-n-1)})\}$, so $\|g_{n+1} - g_n\|_1 < 2^{-n}$, and the functions $\varphi_n := |g_1| + \sum_{\ell=1}^{n-1} |g_{\ell+1} - g_\ell|$ obey

$$\int_M \varphi_n d\mu \;\le\; \|g_1\|_1 + \sum_{\ell=1}^{n-1} 2^{-\ell} \;<\; 1 + \|g_1\|_1\,.$$

Since they are measurable and form a nondecreasing sequence, the monotone–convergence theorem implies existence of a finite $\lim_{n\to\infty} \varphi_n(x)$ for μ–a.a. $x \in M$. Furthermore, $|g_{n+p} - g_n| \le \varphi_{n+p} - \varphi_n$, so there is a function f which is finite μ-a.e. in M and fulfils $f(x) = \lim_{n\to\infty} g_n(x)$. The sequence $\{g_n\}$ has been picked from a Cauchy sequence and it is therefore Cauchy also, $\|g_n - g_m\|_p < \varepsilon$ for all $n, m > \tilde N(\varepsilon)$ for a suitable $\tilde N(\varepsilon)$. On the other hand, $\lim_{m\to\infty} |g_n(x) - g_m(x)|^p = |g_n(x) - f(x)|^p$ for μ–a.a. $x \in M$, so Fatou's lemma implies $\|g_n - f\|_p \le \varepsilon$ for all $n > \tilde N(\varepsilon)$; hence $f \in L^p$ and $\lim_{n\to\infty} \|f_n - f\|_p = 0$ (Problem 24).

If μ is σ–finite and $\mu(M) = \infty$, there is a disjoint decomposition $\bigcup_{j=1}^{\infty} M_j = M$ with $\mu(M_j) < \infty$. The already proven completeness of $L^p(M_j, d\mu)$ implies the existence of functions $f^{(j)} \in L^p(M_j, d\mu)$ which fulfil $\|f_n^{(j)} - f^{(j)}\|_p \to 0$ as $n \to \infty$; then we can proceed as in the proof of completeness of ℓ^p (*cf.* Problem 23).

Other examples of complete metric spaces are given in Problem 23. Any metric space can be extended to become complete: a complete space (X', ϱ') is called the *completion* of (X, ϱ) if (i) $X \subset X'$ and $\varrho'(x,y) = \varrho(x,y)$ for all $x, y \in X$, and (ii) the set X is everywhere dense in X' (this requirement ensures minimality — *cf.* Problem 25).

1.2.2 Theorem: Any metric space (X, ϱ) has a completion. If $(\tilde{X}, \tilde{\varrho})$ is another completion of (X, ϱ), there is an isometry $f : X' \to \tilde{X}$ which preserves X, *i.e.*, $f(x) = x$ for all $x \in X$.

Sketch of the proof: Uniqueness follows directly from the definition. Existence is proved constructively by the so–called *standard completion procedure* which generalizes the Cantor construction of the reals. We start from the set of all Cauchy sequences in (X, ϱ). This can be factorized if we set $\{x_j\} \sim \{y_j\}$ for the sequences with $\lim_{j\to\infty} \varrho(x_j, y_j) = 0$. The set of equivalence classes we denote as X^* and define $\varrho^*([x],[y]) := \lim_{j\to\infty} \varrho(x_j, y_j)$ to any $[x],[y] \in X^*$. Finally, one has to check that this definition makes sense, *i.e.*, that ϱ^* does not depend on the choice of sequences representing the classes $[x],[y]$, ϱ^* is a metric on X^*, and (X^*, ϱ^*) satisfies the requirements of the definition. ∎

The notion of topology is obtained by axiomatization of some properties of metric spaces. Let X be a set and τ a family of its subsets which fulfils the following conditions *(topology axioms):*

(t1) $X \in \tau$ and $\emptyset \in \tau$.

(t2) If I is any index set and $G_\alpha \in \tau$ for all $\alpha \in I$; then $\bigcup_{\alpha \in I} G_\alpha \in \tau$.

(t3) $\bigcap_{j=1}^{n} G_j \in \tau$ for any finite subsystem $\{G_1, \dots, G_n\} \subset \tau$.

The family τ is called a *topology*, its elements *open sets* and the set X equipped with a topology is a **topological space**; when it is suitable we write (X, τ).

A family of open sets in a metric space (X, ϱ) is a topology by definition; we speak about the *metric–induced topology* τ_ϱ, in particular, the *norm–induced topology* if X is a vector space and ϱ is induced by a norm. On the other hand, finding the conditions under which a given topology is induced by a metric is a nontrivial problem (see the notes). Two extreme topologies can be defined on any set X: the *discrete topology* $\tau_d := 2^X$, *i.e.*, the family of all subsets in X, and the *trivial topology* $\tau_0 := \{\emptyset, X\}$. The first of them is induced by the discrete metric, $\varrho_d(x,y) := 1$ for $x \neq y$, while (X, τ_0) is not metrizable unless X is a one–point set.

An open set in a topological space X containing a point x or a set $M \subset X$ is called a *neighborhood* of the point X or the set M, respectively. Using this concept, we can adapt to topological spaces most of the "metric" definitions presented above, as well as some simple results such as those of Problems 17a, c, 19b, topological equivalence of homeomorphic spaces, *etc.* On the other hand, equally elementary metric–space properties may not be valid in a general topological space.

1.2.3 Example: Consider the topologies τ_{fin} and τ_{count} on $X = [0,1]$ in which the closed sets are all finite and almost countable subsets of X, respectively. If $\{x_n\} \subset X$ is a simple sequence, $x_n \neq x_m$ for $n \neq m$; then any neighborhood $U(x)$ contains all elements of the sequence with the exception of a finite number; hence *the limit is not unique* in (X, τ_{fin}). This is not the case in (X, τ_{count}) but there only very few sequences converge, namely those with $x_n = x_{n_0}$ for all $n \geq n_0$, which means, in particular, that we cannot use sequences to characterize local continuity or points of the closure.

Some of these difficulties can be solved by introducing a more general notion of convergence. A partially ordered set I is called *directed* if for any $\alpha, \beta \in I$ there is $\gamma \in I$ such that $\alpha \prec \gamma$ and $\beta \prec \gamma$. A map of a directed index set I into a topological space X, $\alpha \mapsto x_\alpha$, is called a *net* in X. A net $\{x_\alpha\}$ is said to *converge* to a point $x \in X$ if to any neighborhood $U(x)$ there is an $\alpha_0 \in I$ such that $x_\alpha \in U(x)$ for all $\alpha \succ \alpha_0$. To illustrate that nets in a sense play the role that sequences played in metric spaces, let us mention two simple results the proofs of which we leave to the reader (Problem 29).

1.2.4 Proposition: Let (X, τ), (X', τ') be topological spaces; then

(a) A point $x \in X$ belongs to the closure of a set $M \subset X$ *iff* there is a net $\{x_\alpha\} \subset M$ such that $x_\alpha \to x$.

(b) A map $f : X \to X'$ is continuous at a point $x \in X$ *iff* the net $\{f(x_\alpha)\}$ converges to $f(x)$ for any net $\{x_\alpha\}$ converging to x.

Two topologies can be compared if there is an inclusion between them, $\tau_1 \subset \tau_2$, in which case we say that τ_1 is *weaker* (*coarser*) than τ_2; while the latter is *stronger* (*finer*) than τ_1. Such a relation between topologies has some simple consequences — see, *e.g.*, Problem 32. In particular, continuity of a map $f : X \to Y$ is preserved when we make the topology in Y weaker or in X stronger. In other cases it may not be preserved; for instance, Problem 3.9 gives an example of three topologies, $\tau_w \subset \tau_s \subset \tau_u$, on a set $X := \mathcal{B}(\mathcal{H})$ and a map $f : X \to X$ which is continuous with respect to τ_w and τ_u but not τ_s.

1.2.5 Example: A frequently used way to construct a topology on a given X employs a family $\mathcal{F}$ of maps from X to a topological space $(\tilde{X}, \tilde{\tau})$. Among all topologies such that each $f \in \mathcal{F}$ is continuous there is one which is the weakest; its existence follows from Problem 30, where the system $\mathcal{S}$ consists of the sets $f^{(-1)}G)$ for each $G \subset \tilde{\tau}$, $f \in \mathcal{F}$. We call this the *$\mathcal{F}$–weak topology.*

For any set M in a topological space (X,τ) we define the *relative topology* τ_M as the family of intersections $M \cap G$ with $G \subset \tau$; the space (M,τ_M) is called a *subspace* of (X,τ). Other important notions are obtained by axiomatization of properties of open balls in metric spaces. A family $\mathcal{B} \subset \tau$ is called a *basis* of a topological space (X,τ) if any nonempty open set can be expressed as a union of elements of $\mathcal{B}$. A family $\mathcal{B}_x$ of neighborhoods of a given point $x \in X$ is called a *local basis* at x if any neighborhood $U(x)$ contains some $B \in \mathcal{B}_x$. A trivial example of both a basis and a local basis is the topology itself; however, we are naturally more interested in cases where bases are rather a "small part" of it. It is easy to see that local bases can be used to compare topologies.

1.2.6 Proposition: Let a set X be equipped with topologies τ, τ' with local bases $\mathcal{B}_x$, $\mathcal{B}'_x$ at each $x \in X$. The inclusion $\tau \subset \tau'$ holds *iff* for any $B \in \mathcal{B}_x$ there is $B' \in \mathcal{B}'_x$ such that $B' \subset B$.

To be a basis of a topology or a local basis, a family of sets must meet certain consistency requirements (*cf.* Problem 30c, d); this is often useful when we define a particular topology by specifying its basis.

1.2.7 Example: Let (X_j,τ_j), $j=1,2$, be topological spaces. On the Cartesian product $X_1 \times X_2$ we define the standard topology $\tau_{X_1\times X_2}$ determined by τ_j, $j = 1,2$, as the weakest topology which contains all sets $G_1 \times G_2$ with $G_j \in \tau_j$, *i.e.*, $\tau_{X_1\times X_2} := \tau(\tau_1 \times \tau_2)$ in the notation of Problem 30b. Since $(A_1\times A_2)\cap(B_1\times B_2) = (A_1\cap B_1)\times(A_2\cap B_2)$, the family $\tau_1\times\tau_2$ itself is a basis of $\tau_{X_1\times X_2}$; a local basis at $[x_1,x_2]$ consists of the sets $U(x_1)\times V(x_2)$, where $U(x_1)\in\tau_1$, $V(x_2)\in\tau_2$ are neighborhoods of the points x_1,x_2, respectively. The space $(X_1\times X_2,\, \tau_{X_1\times X_2})$ is called the *topological product* of the spaces (X_j,τ_j), $j=1,2$.

Bases can also be used to classify topological spaces by the so–called countability axioms. A space (X,τ) is called *first countable* if it has a countable local basis at any point; it is *second countable* if the whole topology τ has a countable basis. The second requirement is actually stronger; for instance, a nonseparable metric space is first but not second countable (*cf.* Problem 18; some related results are collected in Problem 31). The most important consequence of the existence of a countable local basis, $\{U_n(x) :\ n=1,2,\dots\} \subset \tau$, is that one can pass to another local basis $\{V_n(x) :\ n=1,2,\dots\}$, which is ordered by inclusion, $V_{n+1}\subset V_n$, setting $V_1 := U_1$ and $V_{n+1} := V_n \cap U_{n+1}$. This helps to partially rehabilitate *sequences* as a tool in checking topological properties (Problem 33a).

The other problem mentioned in Example 1.2.3, namely the possible nonuniqueness of a sequence limit, is not related to the cardinality of the basis but rather to the degree to which a given topology separates points. It provides another classification of topological spaces through *separability axioms:*

T_1 To any $x,y\in X$, $x\neq y$, there is a neighborhood $U(x)$ such that $y \notin U(x)$.

T_2 To any $x,y\in X$, $x\neq y$, there are disjoint neighborhoods $U(x)$ and $U(y)$.

T_3 To any closed set $F \subset X$ and a point $x \notin F$, there are disjoint neighborhoods $U(x)$ and $U(F)$.

T_4 To any pair of disjoint closed sets F, F', there are disjoint neighborhoods $U(F)$ and $U(F')$.

A space (X, τ) which fulfils the axioms T_1 and T_j is called T_j*–space*, T_2–spaces are also called *Hausdorff*, T_3–spaces are *regular,* and T_4–spaces are *normal*. For instance, the spaces of Example 3 are T_1 but not Hausdorff; one can find examples showing that the whole hierarchy is nontrivial (see the notes). In particular, any metric space is normal. The question of limit uniqueness that we started with is answered affirmatively in Hausdorff spaces (see Problem 29).

1.3 Compactness

One of the central points in an introductory course of analysis is the Heine–Borel theorem, which claims that given a family of open intervals covering a closed bounded set $F \subset \mathbb{R}$, we can select a finite subsystem which also covers F. The notion of compactness comes from axiomatization of this result. Let M be a set in a topological space (X, τ). A family $\mathcal{P} := \{M_\alpha : \alpha \in I\} \subset 2^X$ is a *covering* of M if $\bigcup_{\alpha \in I} M_\alpha \supset M$; in dependence on the cardinality of the index set I the covering is called finite, countable, *etc.*We speak about an *open* covering if $\mathcal{P} \subset \tau$. The set M is **compact** if an arbitrary open covering of M has a *finite* subsystem that still covers M; if this is true for the whole of X we say that the topological space (X, τ) is compact. It is easy to see that compactness of M is equivalent to compactness of the space (M, τ_M) with the induced topology, so it is often sufficient to formulate theorems for compact spaces only.

1.3.1 Proposition: Let (X, τ) be a compact space, then

(a) Any infinite set $M \subset X$ has at least one accumulation point.

(b) Any closed set $F \subset X$ is compact.

(c) If a map $f : (X, \tau) \to (X', \tau')$ is continuous, then $f(X)$ is compact in (X', τ').

Proof: To check (a) it is obviously sufficient to consider countable sets. Suppose $M = \{x_n : n = 1, 2, \dots\}$ has no accumulation points; then the same is true for the sets $M_N := \{x_n : n \geq N\}$. They are therefore closed and their complements form an open covering of X with no finite subcovering. Further, let $\{G_\alpha\}$ be an open covering of F; adding the set $G := X \setminus F$ we get an open covering of X. Any finite subcovering $\mathcal{G}$ of X is either contained in $\{G_\alpha\}$ or it contains the set G; in the latter case $\mathcal{G} \setminus G$ is a finite covering of the set F. Finally, the last assertion follows from the appropriate definitions. ∎

Part (a) of the proposition represents a particular case of a more general result (see the notes) which can be used to define compactness; another alternative definition is given in Problem 36. Compactness has an important implication for the way in which the topology separates points.

1.3.2 Theorem: A compact Hausdorff space is normal.

Proof: Let F, R be disjoint closed sets and $y \in R$. By assumption, to any $x \in F$ one can find disjoint neighborhoods $U_y(x)$ and $U_x(y)$. The family $\{U_y(x) : x \in F\}$ covers the set F, which is compact in view of the previous proposition; hence there is a finite subsystem $\{U_y(x_j) : j = 1, \dots, n\}$ such that $U_y(F) := \bigcup_{j=1}^{n} U_y(x_j)$ is a neighborhood of F. Moreover, $U(y) := \bigcap_{j=1}^{n} U_{x_j}(y)$ is a neighborhood of the point y and $U(y) \cap U_y(F) = \emptyset$. This can be done for any point $y \in R$ giving an open covering $\{U(y) : y \in R\}$ of the set R; from it we select again a finite subsystem $\{U(y_k) : k = 1, \dots, m\}$ such that $U(R) := \bigcup_{k=1}^{m} U(y_k)$ is a neighborhood of R which has an empty intersection with $U(F) := \bigcap_{k=1}^{m} U_{y_k}(F)$. ∎

1.3.3 Theorem: Let X be a Hausdorff space, then

(a) Any compact set $F \subset X$ is closed.

(b) If the space X is compact, then any continuous bijection $f : X \to X'$ for X' Hausdorff is a homeomorphism.

Proof: If $y \notin F$, the neighborhood $U(y)$ from the preceding proof has an empty intersection with F, so $y \notin \overline{F}$. To prove (b) we have to check that $f(F)$ is closed in X' for any closed $F \subset X$; this follows easily from (a) and Proposition 1.3.1c. ∎

A set M in a topological space is called *precompact* (or *relatively compact*) if $\overline{M}$ is compact. A space X is *locally compact* if any point $x \in X$ has a precompact neighborhood; it is *σ–compact* if any *countable* covering has a finite subcovering.

Let us now turn to compactness in metric spaces. There, any compact set is closed by Theorem 1.3.3 and bounded — from an unbounded set we can always select an infinite subset which has no accumulation point. However, these conditions are not sufficient. For instance, the closed ball $S_1(0)$ in ℓ^2 is bounded but not compact: its subset consisting of the points $X_j := \{\delta_{jk}\}_{k=1}^{\infty}$, $j = 1, 2, \dots$, has no accumulation point because $\|X_j - X_k\| = \sqrt{2}$ holds for all $j \neq k$.

To be able to characterize compactness by metric properties we need a stronger condition. Given a set M in a metric space (X, ϱ) and $\varepsilon > 0$, we call a set N_ε an *ε–lattice* for M if to any $x \in M$ there is a $y \in N_\varepsilon$ such that $\varrho(x, y) \le \varepsilon$ (N_ε may not be a subset of M but by using it one is able to construct a 2ε–lattice for M which is contained in M). A set M is *completely bounded* if it has a *finite* ε–lattice for any $\varepsilon > 0$; if the set X itself is completely bounded we speak about a completely bounded metric space. If M is completely bounded, the same is obviously true for $\overline{M}$. Any completely bounded set is bounded; on the other hand, any infinite orthonormal set in a pre–Hilbert space represents an example of a set which is bounded but not completely bounded.

1.3.4 Proposition: A σ–compact metric space is completely bounded. A completely bounded metric space is separable.

Proof: Suppose that for some $\varepsilon > 0$ there is no finite ε–lattice. Then $X \setminus S_\varepsilon(x_1) \neq \emptyset$ for an arbitrarily chosen $x_1 \in X$, otherwise $\{x_1\}$ would be an ε–lattice for X. Hence there is $x_2 \in X$ such that $\varrho(x_1, x_2) > \varepsilon$ and we have $X \setminus (S_\varepsilon(x_1) \cup S_\varepsilon(x_2)) \neq \emptyset$ *etc.*; in this way we construct an infinite set $\{x_j : j = 1, 2, \dots\}$ which fulfils $\varrho(x_j, x_k) > \varepsilon$ for all $j \neq k$, and therefore it has no accumulation points. As for the second part, if N_n is a $(1/n)$–lattice for X, then $\bigcup_{n=1}^{\infty} N_n$ is a countable everywhere dense set. ∎

1.3.5 Corollary: Let X be a metric space; then the following conditions are equivalent:

(i) X is compact.

(ii) X is σ–compact.

(iii) Any infinite set in X has an accumulation point.

1.3.6 Theorem: A metric space is compact *iff* it is complete and completely bounded.

Proof: Let X be compact; in view of Proposition 1.3.4 it is sufficient to show that it is complete. If $\{x_n\}$ is Cauchy, the compactness implies existence of a convergent subsequence so $\{x_n\}$ is also convergent (Problem 24). On the other hand, to prove the opposite implication we have to check that any $M := \{x_n : n = 1, 2, \dots\} \subset X$ has an accumulation point. By assumption, there is a finite 1–lattice N_1 for X, hence there is $y_1 \in N_1$ such that the closed ball $S_1(y_1)$ contains an infinite subset of M. The ball $S_1(y_1)$ is completely bounded, so we can find a finite $(1/2)$–lattice $N_2 \subset S_1(y_1)$ and a point $y_2 \in N_2$ such that the set $S_{1/2}(y_2) \cap M$ is infinite. In this way we get a sequence of closed balls $S_n := S_{2^{1-n}}(y_n)$ such that each of them contains infinitely many points of M and their centers fulfil $y_{n+1} \in S_n$. The closed balls of doubled radii then satisfy $S_{2^{1-n}}(y_{n+1}) \subset S_{2^{2-n}}(y_n)$ and M has an accumulation point in view of Problem 26. ∎

1.3.7 Corollary: (a) A set M in a complete metric space X is precompact *iff* it is completely bounded. In particular, if X is a finite–dimensional normed space, then M is precompact *iff* it is bounded.

(b) A continuous real–valued function f on a compact topological space X is bounded and assumes its maximum and minimum values in X.

Proof: The first assertion follows from Problem 25. If $\overline{M}$ is compact, it is bounded so M is also bounded. To prove the opposite implication in a finite–dimensional normed space, we can use the fact that such a space is topologically isomorphic to $\mathbb{C}^n$ (or $\mathbb{R}^n$ in the case of a real normed space — see Problem 21). As for part (b), the set $f(X) \subset \mathbb{R}$ is compact by Proposition 1.3.1c, and therefore bounded. Denote

$\alpha := \sup_{x\in X} f(x)$ and let $\{x_n\} \subset X$ be a sequence such that $f(x_n) \to \alpha$. Since X is compact there is a subsequence $\{x_{k_n}\}$ converging to some x_s and the continuity implies $f(x_s) = \alpha$. In the same way we can check that f assumes a minimum value. ∎

1.4 Topological vector spaces

We can easily check that the operations of summation and scalar multiplication in a normed space are continuous. Let us now see what would follow from such a requirement when we combine algebraic and topological properties. A vector space V equipped with a topology τ is called a **topological vector space** if

(tv1) The summation maps continuously $(V\times V,\, \tau_{V\times V})$ to (V,τ).

(tv2) The scalar multiplication maps continuously $(\mathbb{C}\times V,\, \tau_{\mathbb{C}\times V})$ to (V,τ).

(tv3) (V,τ) is Hausdorff.

In the same way, we define a topological vector space over any field. Instead of (tv3), we may demand T_1–separability only because the first two requirements imply that T_3 is valid (Problem 39).

A useful tool in topological vector spaces is the family of translations, $t_x : V \to V$, defined for any $x \in V$ by $t_x(y) := x+y$. Since $t_x^{-1} = t_{-x}$, the continuity of summation implies that any translation is a homeomorphism; hence if G is an open set, then $x + G := t_x(G)$ is open for all $x \in V$; in particular, U is a neighborhood of a point x *iff* $U = x + U(0)$, where $U(0)$ is a neighborhood of zero. This allows us to define a topology through its local basis at a single point (Problem 40).

Suppose a map between topological vector spaces (V,τ) and (V',τ') is simultaneously an algebraic isomorphism of V, V' and a homeomorphism of the corresponding topological spaces, then we call it a *linear homeomorphism* (or *topological isomorphism*). As in the case of normed spaces (*cf.* Problem 21), the structure of a finite–dimensional topological vector space is fully specified by its dimension.

1.4.1 Theorem: Two finite–dimensional topological vector spaces, (V,τ) and (V',τ'), are linearly homeomorphic *iff* $\dim V = \dim V'$. Any finite–dimensional topological vector space is locally compact.

Proof: It is sufficient to construct a linear homeomorphism of a given n–dimensional (V,τ) to $\mathbb{C}^n$. We take a basis $\{e_1,\dots,e_n\} \subset V$ and construct $f : V \to \mathbb{C}^n$ by $f\left(\sum_{j=1}^n \xi_j e_j\right) := [\xi_1,\dots,\xi_n]$; in view of the continuity of translations we have to show that f and f^{-1} are continuous at zero. According to (tv1), for any $U(0) \in \tau$ we can find neighborhoods $U_j(0)$ such that $\sum_{j=1}^n x_j \in U(0)$ for $x_j \in U_j(0)\,,\ j = 1,\dots,n$ and f^{-1} is continuous by Problem 42a. To prove that f is continuous we use the fact that V is Hausdorff: Proposition 1.3.1 and Theorem 1.3.3 together with the already proven continuity of f^{-1} ensure that $S_\varepsilon := \{x\in V : \|f(x)\| = \varepsilon\}$ $= f^{(-1)}(K_\varepsilon)$ is closed for any $\varepsilon > 0$; we have denoted here by K_ε the ε–sphere in $\mathbb{C}^n$. Since $0 \notin S_\varepsilon$, the set $G := V \setminus S_\varepsilon$ is a neighborhood of zero, and by Problem 42b there is a balanced neighborhood $U \subset G$ of zero; this is possible only if $\|f(x)\| < \varepsilon$ for all $x \in U$. ∎

Next we want to discuss a class of topological vector spaces whose properties are closer to those of normed spaces. In distinction to the latter the topology in them is not specified generally by a single (semi)norm but rather by a family of them. Let $\mathcal{P} := \{p_\alpha : \alpha \in I\}$ be a family of seminorms on a vector space V where I is an arbitrary index set. We say that $\mathcal{P}$ *separates points* if to any nonzero $x \in V$ there is a $p_\alpha \in \mathcal{P}$ such that $p_\alpha(x) \neq 0$. It is clear that if $\mathcal{P}$ consists of a single seminorm p it separates points *iff* p is a norm. Given a family $\mathcal{P}$ we set

$$B_\varepsilon(p_1, \dots, p_n) := \{\, x \in V \,:\, p_j(x) < \varepsilon\,,\ j = 1, \dots, n \,\}\,;$$

the collection of these sets for any $\varepsilon > 0$ and all finite subsystems of $\mathcal{P}$ will be denoted as $\mathcal{B}_0^{\mathcal{P}}$. In view of Problem 40, $\mathcal{B}_0^{\mathcal{P}}$ defines a topology on V which we denote as $\tau^{\mathcal{P}}$.

1.4.2 Theorem: If a family $\mathcal{P}$ of seminorms on a vector space V separates points, then $(V, \tau^{\mathcal{P}})$ is a topological vector space.

Proof: By assumption, to a pair x, y of different points there is a $p \in \mathcal{P}$ such that $\varepsilon := \frac{1}{2}\,p(x-y) > 0$. Then $U(x) := x + B_\varepsilon(p)$ and $U(y) := y + B_\varepsilon(p)$ are disjoint neighborhoods, so the axiom T_2 is valid. The continuity of summation at the point $[0,0]$ follows from the inequality $p(x{+}y) \le p(x) + p(y)$; for the scalar multiplication we use $p(\alpha x - \alpha_0 x_0) \le |\alpha - \alpha_0|\, p(x_0) + |\alpha|\, p(x - x_0)$. ∎

A topological vector space with a topology induced by a family $\mathcal{P}$ separating points is called *locally convex*. This name has an obvious motivation: if $x, y \in B_\varepsilon(p_1, \dots, p_n)$, then $p_j(tx + (1-t)x) \le tp_j(x) + (1-t)p_j(x)$ holds for any $t \in [0,1]$ so the sets $B_\varepsilon(p_1, \dots, p_n)$ are convex. The convexity is preserved at translations, so the local basis of $\tau^{\mathcal{P}}$ at each $x \in V$ consists of convex sets (see also the notes).

1.4.3 Example: The family $\mathcal{P} := \{p_x := |(x, \cdot)| \,:\, x \in V\}$ in a pre–Hilbert space V generates a locally convex topology which is called the *weak topology* and is denoted as τ_w; it is easy to see that it is weaker than the "natural" topology induced by the norm.

1.4.4 Theorem: A locally convex space (V, τ) is metrizable *iff* there is a countable family $\mathcal{P}$ of seminorms which generates the topology τ.

Proof: If V is metrizable it is first countable. Let $\{U_j : j = 1, 2, \dots\}$ be a local basis of τ at the point 0. By definition, to any U_j we can find $\varepsilon > 0$ and a finite subsystem $\mathcal{P}_j \subset \mathcal{P}$ such that $\bigcap_{p \in \mathcal{P}_j} B_\varepsilon(p) \subset U_j$. The family $\mathcal{P}' := \bigcup_{j=1}^{\infty} \mathcal{P}_j$ is countable and generates a topology $\tau^{\mathcal{P}'}$ which is not stronger than $\tau := \tau^{\mathcal{P}}$; the above inclusion shows that $\tau^{\mathcal{P}'} = \tau$. On the other hand, suppose that τ is generated by a family $\{p_n : n = 1, 2, \dots\}$ separating points; then we can define a metric ϱ as in Problem 16 and show that the corresponding topology satisfies $\tau_\varrho = \tau$ (Problem 43). ∎

A locally convex space which is complete with respect to the metric used in the proof is called a *Fréchet space* (see also the notes).

1.4.5 Example: The set $\mathcal{S}(\mathbb{R}^n)$ consists of all infinitely differentiable functions $f:\mathbb{R}^n\to\mathbb{C}$ such that

$$\|f\|_{J,K} := \sup_{x\in\mathbb{R}^n} |x^J(D^Kf)(x)| < \infty$$

holds for any multi–indices $J:=[j_1,\dots,j_n]$, $K:=[k_1,\dots,k_n]$ with $j_r,\,k_r$ non–negative integers, where $x^J := \xi_1^{j_1}\dots\xi_n^{j_n}$, $D^K := \partial^{|K|}/\partial\xi_1^{k_1}\dots\partial\xi_n^{k_n}$ and $|K| := k_1+\dots+k_n$. It is easy to see that any such f and any derivative D^Kf (as well as polynomial combinations of them) tend to zero faster than $|x^J|^{-1}$ for each J; we speak about *rapidly decreasing* functions. It is also clear that any $\|\cdot\|_{J,K}$ is a seminorm, with $\|f\|_{0,0}=\|f\|_\infty$, and the family $\mathcal{P}:=\{\|\cdot\|_{J,K}\}$ separates points. The corresponding locally convex space $\mathcal{S}(\mathbb{R}^n)$ is called the *Schwartz space*; one can show that it is complete, *i.e.*, a Fréchet space (see the notes).

An important subspace in $\mathcal{S}(\mathbb{R}^n)$ consists of infinitely differentiable functions with a compact support; we denote it as $C_0^\infty(\mathbb{R}^n)$. It is dense,

$$\overline{C_0^\infty(\mathbb{R}^n)} = \mathcal{S}(\mathbb{R}^n)\,, \tag{1.1}$$

with respect to the topology of $\mathcal{S}(\mathbb{R}^n)$ (Problem 44).

1.5 Banach spaces and operators on them

A normed space which is complete with respect to the norm–induced metrics is called a **Banach space**. We have already met some frequently used Banach spaces — see Example 1.2.1 and Problem 23. In view of Problem 21, any finite–dimensional normed space is complete; in the general case we have the following completeness criterion, the proof of which is left to the reader (see also Example 1.5.3b below).

1.5.1 Theorem: A normed space V is complete *iff* to any sequence $\{x_n\}\subset V$ such that $\sum_{n=1}^\infty \|x_n\|<\infty$ there is an $x\in V$ such that $x=\lim_{n\to\infty}\sum_{k=1}^n x_k$ (or in short, *iff* any absolutely summable sequence is summable).

Given a noncomplete norm space, we can always extend it to a Banach space by the standard completion procedure (Problem 46). A set M in a Banach space $\mathcal{X}$ is called *total* if $\overline{M_{lin}}=\mathcal{X}$. Such a set is a basis if M is linearly independent and $\dim\mathcal{X}<\infty$, while an infinite–dimensional space can contain linearly independent total sets, which are not Hamel bases of $\mathcal{X}$ (*cf.* the notes to Section 1.1).

1.5.2 Lemma: (a) If M is total in a Banach space $\mathcal{X}$, then any set $N\subset\mathcal{X}$ dense in M is total in $\mathcal{X}$.

(b) A Banach space which contains a countable total set is separable.

Proof: Part (a) follows from the appropriate definitions. Suppose that $M = \{x_1, x_2, \dots\}$ is total in $\mathcal{X}$ and $\mathbb{C}_{rat}$ is the countable set of complex numbers with rational real and imaginary parts; then the set $L := \{\sum_{j=1}^n \gamma_j x_j \,:\, \gamma_j \in \mathbb{C}_{rat},\, n < \infty\}$ is countable. Since $\mathbb{C}_{rat}$ is dense in $\mathbb{C}$, we get $\overline{L} = \mathcal{X}$. ∎

1.5.3 Examples: (a) The set $\mathcal{P}(a,b)$ of all complex polynomials on (a,b) is an infinite–dimensional subspace in $C[a,b] := C([a,b])$. By the Weierstrass theorem, any $f \in C[a,b]$ can be approximated by a uniformly convergent sequence of polynomials; hence $C[a,b]$ is a complete envelope of $(\mathcal{P}(a,b), \|\cdot\|_\infty)$. The set $\{x^k : k = 0,1,\dots\}$ is total in $C[a,b]$, which is therefore separable.

(b) Consider the sequences $E_k := \{\delta_{jk}\}_{j=1}^\infty$ in ℓ^p, $p \ge 1$. For a given $X := \{\xi_j\} \in \ell^p$, the sums $X_n := \sum_{j=1}^n \xi_j E_j$ are nothing else than truncated sequences, so $\lim_{n\to\infty} \|X - X_n\|_p = 0$. Hence $\{E_k : k = 1,2\dots\}$ is a countable total set and ℓ^p is separable. Notice also that the sequence $\{\xi_j E_j\}_{j=1}^\infty$ is summable but it may not be absolutely summable for $p > 1$.

(c) Consider next the space $L^p(\mathbb{R}^n, d\mu)$ with an arbitrary Borel measure μ on $\mathbb{R}^n$. We use the notation of Appendix A. In particular, $\mathcal{J}^n$ is the family of all bounded intervals in $\mathbb{R}^n$; then we define $S^{(n)} := \{\chi_J : J \in \mathcal{J}^n\}$. It is a subset in L^p and the elements of its linear envelope are called *step functions*; we can check that $S^{(n)}$ is total in $L^p(\mathbb{R}^n, d\mu)$ (Problem 47). Combining this result with Lemma 1.5.2 we see that the subspace $C_0^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n, d\mu)$; in particular, for the Lebesgue measure on $\mathbb{R}^n$ the inclusions $C_0^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ yield

$$\overline{(C_0^\infty(\mathbb{R}^n))}_p = \overline{(\mathcal{S}(\mathbb{R}^n))}_p = L^p(\mathbb{R}^n)\,. \tag{1.2}$$

(d) Given a topological space (X,τ) we call $C_\infty(X)$ the set of all continuous functions on X with the following property: for any $\varepsilon > 0$ there is a compact set $K \subset X$ such that $|f(x)| < \varepsilon$ outside K. It is not difficult to check that $C_\infty(X)$ is a closed subspace in $C(X)$ and $\overline{C_0(X)} = C_\infty(X)$, where $C_0(X)$ is the set of continuous functions with compact support (Problem 48). In the particular case $X = \mathbb{R}^n$, $C_0^\infty(\mathbb{R}^n)$ is dense in $C_\infty(\mathbb{R}^n)$ (see the notes), so

$$\overline{(C_0^\infty(\mathbb{R}^n))}_\infty = \overline{(\mathcal{S}(\mathbb{R}^n))}_\infty = C_\infty(\mathbb{R}^n)\,. \tag{1.3}$$

There are various ways in, which it is possible to construct new Banach spaces from given ones. We mention two of them (see also Problem 49):

(i) Let $\{\mathcal{X}_j : j = 1,2,\dots\}$ be a countable family of Banach spaces. We denote by $\mathcal{X}$ the set of all sequences $x := \{x_j\}$, $x_j \in \mathcal{X}_j$, such that $\sum_j \|x_j\|_j < \infty$, and equip it with the "componentwise" defined summation and scalar multiplication. The norm $\|X\|_\oplus := \sum_j \|x_j\|_j$ turns it into a Banach space; the completeness can be checked as for ℓ^p (Problem 23). The space $(\mathcal{X}, \|\cdot\|_\oplus)$ is called the *direct sum* of the spaces $\mathcal{X}_j$, $j = 1,2,\dots$, and denoted as $\sum_j^\oplus \mathcal{X}_j$.

(ii) Starting from the same family $\{\mathcal{X}_j : j = 1, 2, \dots\}$, one can define another Banach space (which is sometimes also referred to as a direct sum) if we change the above norm to $\|X\|_\infty := \sup_j \|x_j\|_j$ replacing, of course, $\mathcal{X}$ by the set of sequences for which $\|X\|_\infty < \infty$. The two Banach spaces are different unless the family $\{\mathcal{X}_j\}$ is finite; the present construction can easily be adapted to families of any cardinality.

A map $B : V_1 \to V_2$ between two normed spaces is called an *operator*; in particular, it is called a **linear operator** if it is linear. In this case we conventionally do not use parentheses and write the image of a vector $x \in V_1$ as Bx. In this book we shall deal almost exclusively with linear operators, and therefore the adjective will usually be dropped. A linear operator $B : V_1 \to V_2$ is said to be **bounded** if there is a positive c such that $\|Bx\|_2 \le c\|x\|_1$ for all $x \in V_1$; the set of all such operators is denoted as $\mathcal{B}(V_1, V_2)$ or simply $\mathcal{B}(V)$ if $V_1 = V_2 := V$. One of the elementary properties of linear operators is the equivalence between continuity and boundedness (Problem 50).

The set $\mathcal{B}(V_1, V_2)$ becomes a vector space if we define on it summation and scalar multiplication by $(\alpha B + C)x := \alpha Bx + Cx$. Furthermore, we can associate with every $B \in \mathcal{B}(V_1, V_2)$ the non–negative number

$$\|B\| := \sup_{S_1} \|Bx\|_2 \,,$$

where $S_1 := \{x \in V_1 : \|x\|_1 = 1\}$ is the unit sphere in V_1 (see also Problem 51).

1.5.4 Proposition: The map $B \mapsto \|B\|$ is a norm on $\mathcal{B}(V_1, V_2)$. If V_2 is complete, the same is true for $\mathcal{B}(V_1, V_2)$, *i.e.*, it is a Banach space.

Proof: The first assertion is elementary. Let $\{B_n\}$ be a Cauchy sequence in $\mathcal{B}(V_1, V_2)$; then for all n, m large enough we have $\|B_n – B_m\| < \varepsilon$, and therefore $\|B_n x – B_m x\|_2 \le \varepsilon\|x\|_1$ for any $x \in V_1$. As a Cauchy sequence in V_2, $\{B_n x\}$ converges to some $B(x) \in V_2$. The linearity of the operators B_n implies that $x \mapsto Bx$ is linear, $B(x) = Bx$. The limit $m \to \infty$ in the last inequality gives $\|Bx – B_n x\|_2 \le \varepsilon\|x\|_1$, so $B \in \mathcal{B}(V_1, V_2)$ by the triangle inequality, and $\|B – B_n\| \le \varepsilon$ for all n large enough. ∎

The norm on $\mathcal{B}(V_1, V_2)$ introduced above is called the *operator norm*. It has an additional property: if $C : V_1 \to V_2$ and $B : V_2 \to V_3$ are bounded operators, and BC is the *operator product* understood as the composite mapping $V_1 \to V_3$, we have $\|B(Cx)\|_3 \le \|B\|\,\|Cx\|_2 \le \|B\|\,\|C\|\,\|x\|_1$ for all $x \in V_1$, so BC is also bounded and

$$\|BC\| \le \|B\|\,\|C\| \,. \tag{1.4}$$

Let V_1 be a subspace of a normed space $\tilde V_1$. An operator $B : V_1 \to V_2$ is called a *restriction* of $\tilde B : \tilde V_1 \to V_2$ to the subspace V_1 if $Bx = \tilde B x$ holds for all $x \in V_1$, and on the other hand, $\tilde B$ is said to be an *extension* of B; we write $B = \tilde B \upharpoonright V_1$ or $B \subset \tilde B$. Another simple property of bounded operators is that they can be extended uniquely by continuity.