Graduate Texts in Mathematics

Joachim Weidmann

Linear Operations in Hilbert Spaces

Springer

Graduate Texts in Mathematics

68

Joachim Weidmann

Linear Operators in Hilbert Spaces

Translated by
Joseph Szücs

Springer-Verlag
New York Heidelberg Berlin

Joachim Weidmann

Mathematisches Seminar der
Johann-Wolfgang Goethe-Universität
Institut für Angewandte Mathematik
Robert-Mayer-Strasse 10
6 Frankfurt a. M.
Federal Republic of Germany

Joseph Szücs

American Mathematical Society
P.O. Box 6248
Providence, RI 02940
USA

AMS Subject Classifications: 47Axx, 47B05, 47B10, 47B15, 47B20, 47B25, 47B30, 47E05, 81A09, 81A10, 81A45

With 1 Figure.

Library of Congress Cataloging in Publication Data

Weidmann, Joachim.
Linear operators in Hilbert spaces.

(Graduate texts in mathematics; 68)
Translation of Lineare Operatoren in Hilberträumen.
Bibliography: p.
Includes index.
1. Linear operators. 2. Hilbert space. I. Title. II. Series.
QA329.2.W4413 515'.72 79-12649

Exclusively authorized English translation of the original German edition *Lineare Operatoren in Hilberträumen* published in the series Mathematische Leitfäden edited by G. Köthe by B. G. Teubner, Stuttgart, 1976.

9 8 7 6 5 4 3 2

ISBN-13: 978-1-4612-6029-5 e-ISBN-13: 978-1-4612-6027-1
DOI: 10.1007/978-1-4612-6027-1

To the memory of
Konrad Jörgens

Preface to the English edition

This English edition is almost identical to the German original *Lineare Operatoren in Hilberträumen*, published by B. G. Teubner, Stuttgart in 1976. A few proofs have been simplified, some additional exercises have been included, and a small number of new results has been added (e.g., Theorem 11.11 and Theorem 11.23). In addition, a great number of minor errors has been corrected.

Frankfurt, January 1980 — J. Weidmann

Preface to the German edition

The purpose of this book is to give an introduction to the theory of linear operators on Hilbert spaces and then to proceed to the interesting applications of differential operators to mathematical physics. Besides the usual introductory courses common to both mathematicians and physicists, only a fundamental knowledge of complex analysis and of ordinary differential equations is assumed. The most important results of Lebesgue integration theory, to the extent that they are used in this book, are compiled with complete proofs in Appendix A. I hope therefore that students from the fourth semester on will be able to read this book without major difficulty. However, it might also be of some interest and use to the teaching and research mathematician or physicist, since among other things it makes easily accessible several new results of the spectral theory of differential operators.

In order to limit the length of the text, I present the results of abstract functional analysis only insofar as they are significant for this book. I prove those theorems (for example, the closed graph theorem) that also hold in more general Banach spaces by Hilbert space methods whenever this leads to simplification. The typical concepts of Hilbert space theory, "orthogonal" and "self-adjoint," stand clearly at the center. The spectral theorem for self-adjoint operators and its applications are the central topics of this book. A detailed exposition of the theory of expansions in terms of generalized eigenfunctions and of the spectral theory of ordinary differential operators (Weyl–Titchmarsh–Kodaira) was not possible within the framework of this book.

In the first three chapters pre-Hilbert spaces and Hilbert spaces are introduced, and their basic geometric and topologic properties are proven. Chapters 4 and 5 contain the fundamentals of the theory of (not necessarily bounded) linear operators on Hilbert spaces, including general spectral theory. Besides the numerous examples scattered throughout the text, in Chapter 6 certain important classes of linear operators are studied in detail. Chapter 7 contains the spectral theory of self-adjoint operators (first for compact operators, and then for the general case), as well as some important consequences and a detailed characterization of the spectral points. In Chapter 8 von Neumann's extension theory for symmetric operators is developed and is applied to, among other things, the Sturm-Liouville operators. Chapter 9 provides some important results of perturbation theory for self-adjoint operators. Chapter 10 begins with proofs of the most significant facts about Fourier transforms in $L_2(\mathbb{R}^m)$; applications to partial differential operators, in particular to Schrödinger and Dirac operators, follow. Finally, Chapter 11 gives a short introduction to (time dependent) scattering theory with some typical results; to my regret, I could only touch upon the far reaching results of recent years.

Exercises are not used later in the text, with a few exceptions. They mainly serve to deepen understanding of the material and give opportunity for practice; however, I often use them to formulate further results which I cannot treat in the text. The level of difficulty of the exercises varies widely; because I give many exercises with detailed hints, they can be solved in general without much difficulty.

Now I want to very heartily thank all those who helped me with the production of this book. Mrs. Hose turned my notes into an excellent typed manuscript with infinite diligence. Messrs. R. Holstein, D. Keim and H. Kiloh spent much time reading the whole manuscript and discussing with me their suggestions for improvement. Messrs. R. Colgen and W. Stork helped me with the proofreading. I thank the publisher and the editors for their pleasant cooperation.

My teacher Konrad Jörgens inspired me to study this subject; he influenced the present exposition in several ways. I dedicate this volume to his memory.

Hattersheim am Main, the summer of 1976 Joachim Weidmann

Contents

1 Vector spaces with a scalar product, pre-Hilbert spaces

In what follows we consider vector spaces over a field $\mathbb{K}$, where $\mathbb{K}$ is either the field $\mathbb{C}$ of complex numbers or the field $\mathbb{R}$ of real numbers; accordingly, we speak of a complex or a real vector space. For every $c \in \mathbb{K}$ let c^* be the complex conjugate of c; so for $c \in \mathbb{R}$ the star has no significance.

As a rule we assume the most important notions and results of linear algebra to be known.

1.1 Sesquilinear forms

Let H be a vector space over $\mathbb{K}$. A mapping $s: H \times H \to \mathbb{K}$ is called a *sesquilinear form on H* if for all $f, g, h \in H$ and $a, b \in \mathbb{K}$ we have

$$s(f, ag + bh) = as(f, g) + bs(f, h), \tag{1.1}$$

$$s(af + bg, h) = a^* s(f, h) + b^* s(g, h). \tag{1.2}$$

If (1.2) holds without stars, then s is called a bilinear form on H; in particular every sesquilinear form on a real vector space is a bilinear form.

Property (1.1) is obviously equivalent to the two properties

$$s(f, g + h) = s(f, g) + s(f, h), \tag{1.1'}$$

$$s(f, ag) = as(f, g). \tag{1.1''}$$

Similarly (1.2) is equivalent to

$$s(f + g, h) = s(f, h) + s(g, h), \tag{1.2'}$$

$$s(af, g) = a^* s(f, g). \tag{1.2''}$$

If s is a sesquilinear form on H, then the mapping $q: H \to \mathbb{K}$ that is defined by $q(f) = s(f, f)$ for each $f \in H$ is called the *quadratic form* on H generated or induced by s. For each quadratic form q we obviously have

$$q(af) = |a|^2 q(f) \quad \text{for all} \quad f \in H, a \in \mathbb{K}; \tag{1.3}$$

so we have, in particular, $q(af) = q(f)$ for every $a \in \mathbb{K}$ with $|a| = 1$.

The following theorem shows that in a complex vector space the generating sesquilinear form is uniquely determined by the quadratic form, for real vector spaces this is not necessarily true in general; see Exercise 1.2.

Theorem 1.1 (Polarization identity). *Let H be a complex vector space, s a sesquilinear form on H, and q the quadratic form generated by s. Then for all $f, g \in H$ we have*

$$s(f, g) = \tfrac{1}{4}\{q(f+g) - q(f-g) + iq(f-ig) - iq(f+ig)\}. \tag{1.4}$$

The proof of this identity may be given by calculating the right side of (1.4) according to the rules (1.1) and (1.2).

Theorem 1.2 (Parallelogram law). *Let s be a sesquilinear form on a vector space H, and let q be the corresponding quadratic form on H. Then for all $f, g \in H$ we have*

$$q(f+g) + q(f-g) = 2[q(f) + q(g)]. \tag{1.5}$$

PROOF. For every $f, g \in H$ we have

$$\begin{aligned} q(f+g) + q(f-g) &= s(f,f) + s(f,g) + s(g,f) + s(g,g) \\ &\quad + s(f,f) - s(f,g) - s(g,f) + s(g,g) \\ &= 2q(f) + 2q(g). \end{aligned}$$

□

A sesquilinear form s on H is said to be *Hermitian* provided that for every $f, g \in H$ we have

$$s(f, g) = s(g, f)^*. \tag{1.6}$$

A Hermitian bilinear form on a real vector space is said to be *symmetric*.

If s is a Hermitian sesquilinear form, and q the quadratic form generated by s, then we obviously have $q(f) \in \mathbb{R}$ for all $f \in H$; we say briefly that q is *real*. The following theorem shows, among other things, that Hermitian sesquilinear forms can be characterized by this property of their associated quadratic forms. We also obtain that symmetric bilinear forms are uniquely determined by the corresponding quadratic forms.

Theorem 1.3. *Let H be a vector space over $\mathbb{K}$, s a sesquilinear form on H, and q the quadratic form generated by s.*

(a) *If* $\mathbb{K} = \mathbb{C}$, *then the following statements are equivalent*:

(i) *s is symmetric,*

(ii) *q is real,*

(iii) *for all* $f, g \in H$ *we have*

$$\operatorname{Re} s(f, g) = \tfrac{1}{4}\{q(f+g) - q(f-g)\}, \tag{1.7}$$

(iv) *for all* $f, g \in H$ *we have*

$$\operatorname{Im} s(f, g) = \tfrac{1}{4}\{q(f-ig) - q(f+ig)\}. \tag{1.7'}$$

(b) *If* $\mathbb{K} = \mathbb{R}$, *then the following statements are equivalent*:

(i) *s is symmetric,*

(ii) *for all* $f, g \in H$ *we have*

$$s(f, g) = \tfrac{1}{4}\{q(f+g) - q(f-g)\}. \tag{1.8}$$

PROOF.

(a) (ii) *follows from* (i): $q(f)^* = s(f, f)^* = s(f, f) = q(f)$, i.e., $q(f)$ is real.

(iii) *follows from* (ii): Because $q(h) \in \mathbb{R}$ for all $h \in H$, it follows from (1.4) that

$$\begin{aligned}\operatorname{Re} s(f, g) &= \tfrac{1}{4}\operatorname{Re}\{q(f+g) - q(f-g) + iq(f-ig) - iq(f+ig)\}\\ &= \tfrac{1}{4}\{q(f+g) - q(f-g)\}.\end{aligned}$$

(iv) *follows from* (iii): Because of (iii) we have

$$\begin{aligned}\operatorname{Im} s(f, g) &= -\operatorname{Re}\{is(f, g)\}\\ &= \operatorname{Re} s(f, -ig) = \tfrac{1}{4}\{q(f-ig) - q(f+ig)\}.\end{aligned}$$

(i) *follows from* (iv):

$$\begin{aligned}s(g, f)^* &= \operatorname{Re} s(g, f) - i \operatorname{Im} s(g, f) = \operatorname{Im} s(g, if) - i \operatorname{Im} s(g, f)\\ &= \tfrac{1}{4}\{q(g+f) - q(g-f) - iq(g-if) + iq(g+if)\}\\ &= \tfrac{1}{4}\{q(f+g) - q(f-g) + iq(f-ig) - iq(f+ig)\} = s(f, g);\end{aligned}$$

here we have used (1.3) with $\alpha = -1$, $\alpha = i$, and $\alpha = -i$.

(b) (ii) *follows from* (i) by calculating the right side of (ii) while using the symmetry of s.

(i) *follows from* (ii):

$$\begin{aligned}s(g, f) &= \tfrac{1}{4}\{q(g+f) - q(g-f)\}\\ &= \tfrac{1}{4}\{q(f+g) - q(f-g)\} = s(f, g).\end{aligned} \qquad \square$$

A Hermitian sesquilinear form is said to be *non-negative* when

$$s(f, f) \geqslant 0 \quad \text{for all} \quad f \in H; \tag{1.9}$$

it is said to be *positive* when

$$s(f,f)>0 \quad \text{for all} \quad f\in H \quad \text{with} \quad f\neq 0. \tag{1.10}$$

Since we have $s(0,0)=0$, every positive sesquilinear form is non-negative. We also say that the corresponding quadratic forms are *non-negative*, respectively *positive*. (Because of Theorem 1.3, the word "Hermitian" may be omitted from this definition in the complex case; this does not hold in the real case, cf. Exercise 1.5.)

Theorem 1.4. *If s is a non-negative sesquilinear form on H, and q denotes the quadratic form generated by s, then for every $f, g\in H$ we have the Schwarz inequality*

$$|s(f,g)|\leq [q(f)q(g)]^{1/2}. \tag{1.11}$$

If s is positive, then the equality sign in (1.11) holds if and only if f and g are linearly dependent; the equality $s(f,g)=[q(f)q(g)]^{1/2}$ holds if and only if there exists a $c\geq 0$ such that $f=cg$ or $g=cf$.

PROOF. Let $f,g\in H$. For all $t\in\mathbb{R}$ we have

$$0\leq q(f+tg)=q(f)+2t\,\mathrm{Re}\,s(f,g)+t^2q(g).$$

This second degree polynomial in t has either no root or a double root. Since this holds for a polynomial $at^2+2bt+c$ if and only if $b^2-ac\leq 0$, it follows that

$$[\mathrm{Re}\,s(f,g)]^2\leq q(f)q(g). \tag{1.12}$$

If one chooses $a\in\mathbb{K}$ such that $|a|=1$ and $as(f,g)=|s(f,g)|$ holds, then it follows from (1.12) with $h=ag$ that

$$\begin{aligned}|s(f,g)|^2 &= [\mathrm{Re}\,as(f,g)]^2=[\mathrm{Re}\,s(f,h)]^2\\ &\leq q(f)q(h)=q(f)q(ag)=q(f)q(g);\end{aligned}$$

this is the *Schwarz inequality*.

Let s now be positive and let $s(f,g)=[q(f)q(g)]^{1/2}$ be true. If $g=0$, then the equality $g=0f$ proves the assertion. Consequently, let $g\neq 0$. Because of the equality $[\mathrm{Re}\,s(f,g)]^2-q(f)q(g)=0$, the polynomial considered above has a double root t_0; hence we have $q(f+t_0g)=0$, i.e., $f=-t_0g$. From $-t_0s(g,g)=s(f,g)\geq 0$ it follows that $-t_0\geq 0$. If we have $|s(f,g)|=[q(f)q(g)]^{1/2}$ and choose a and h as above, then $s(f,h)=[q(f)q(h)]^{1/2}$ follows. According to the part just proved we then have either $g=h=0f$, or there exists a $c\geq 0$ such that $f=ch=acg$. In both cases f and g are linearly dependent. One can verify the converses of the last two assertions by simple calculation. □

EXAMPLE 1. For each $m\in\mathbb{N}$ ($\mathbb{N}$ denotes the set $\{1,2,3,\ldots\}$ of positive integers) let $\mathbb{C}^m$ be the complex vector space of the m-tuples $f=$

$(f_1, f_2, \ldots, f_m)$, $g = (g_1, g_2, \ldots, g_m), \ldots$ of complex numbers with the addition

$$f + g = (f_1 + g_1, f_2 + g_2, \ldots, f_m + g_m)$$

and multiplication by $a \in \mathbb{C}$

$$af = (af_1, af_2, \ldots, af_m).$$

If $(a_{jk})_{j,k=1,\ldots,m}$ is a complex $m \times m$ matrix, then

$$s(f, g) = \sum_{j,k=1}^{m} a_{jk} f_j^* g_k \quad \text{for} \quad f, g \in \mathbb{C}^m$$

defines a sesquilinear form on $\mathbb{C}^m$. s is Hermitian if and only if the matrix (a_{jk}) is Hermitian, i.e., if for every $j, k = 1, 2, \ldots, m$ we have $a_{kj} = a_{jk}^*$. s is non-negative (positive) if, for example, (a_{jk}) is a diagonal matrix with non-negative (positive) entries in the diagonal. An important special case of a positive sesquilinear form on $\mathbb{C}^m$ occurs when (a_{jk}) is the unit matrix. Then

$$s(f, g) = \sum_{j=1}^{m} f_j^* g_j.$$

EXAMPLE 2. On the real vector space $\mathbb{R}^m$ (symmetric, non-negative, positive) bilinear forms can be given accordingly.

EXAMPLE 3. Let $C[0, 1]$ be the complex vector space of complex-valued continuous functions defined on $[0, 1]$ with the addition

$$(f + g)(x) = f(x) + g(x)$$

and multiplication by $a \in \mathbb{C}$

$$(af)(x) = af(x).$$

If $r : [0, 1] \to \mathbb{C}$ is continuous, then by

$$s(f, g) = \int_0^1 f(x)^* g(x) r(x) \, \mathrm{d}x \qquad f, g \in C[0, 1]$$

a sesquilinear form is defined on $C[0, 1]$. It is Hermitian if and only if r is real-valued; it is non-negative if and only if $r(x) \geq 0$ for all $x \in [0, 1]$; it is positive if and only if $r(x) \geq 0$ for all $x \in [0, 1]$ and r does not vanish identically on any non-trivial interval.

EXAMPLE 4. Let $C_{\mathbb{R}}[0, 1]$ be the real vector space of real-valued continuous functions defined on $[0, 1]$. For each continuous function $r : [0, 1] \to \mathbb{R}$ the bilinear form

$$s(f, g) = \int_0^1 f(x) g(x) r(x) \, \mathrm{d}x \qquad f, g \in C[0, 1]$$

is symmetric. Concerning non-negativity and positivity the same assertions hold as in Example 3.

EXAMPLE 5. If $k:[0, 1]\times[0, 1]\to\mathbb{C}$ is continuous, then by

$$s(f, g) = \int_0^1\int_0^1 k(x, y) f(x)^* g(y)\, dy\, dx$$

a sesquilinear form is defined on $C[0, 1]$. This is Hermitian if and only if the kernel k is Hermitian, i.e., if for every $x, y\in[0, 1]$ we have $k(x, y) = k(y, x)^*$.

Exercises

1.1. Prove the assertions given in Examples 1–5.

1.2. The matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

generates a non-zero sesquilinear form on $\mathbb{R}^2$ (cf. Example 2), the quadratic form of which vanishes. Consequently, in a real vector space sesquilinear forms are not determined uniquely by the corresponding quadratic forms.

1.3. Let s be the sesquilinear form on $\mathbb{R}^2$ generated by the matrix

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}.$$

If $|a|<2$ ($|a|\leq 2$), then we have $s(f, f)>0$ for all $f\in\mathbb{R}^2$ such that $f\neq 0$ ($s(f, f)\geq 0$ for all $f\in\mathbb{R}^2$). If $a\neq 0$, then s is not symmetric.

1.4. Let s be a non-negative sesquilinear form on H, q the quadratic form generated by s, and $N=\{f\in H: q(f)=0\}$. Show that
(a) N is a subspace (sub-vectorspace) of H.
(b) If $f\in N$ and $g\in H$, then we have $s(f, g)=0$ and $q(f+g)=q(g)$.
(c) In the Schwarz inequality the equality sign holds if and only if f and g are linearly dependent modulo N, i.e., if there are numbers $a, b\in\mathbb{K}$ not vanishing simultaneously and such that $af+bg\in N$.
(d) We have $s(f, g)=[q(f)q(g)]^{1/2}$ if and only if there is a $c\geq 0$ such that $f-cg\in N$ or $g-cf\in N$.

1.5. Prove the *Cauchy inequality*

$$\left|\sum_{j=1}^m f_j^* g_j\right|^2 \leq \sum_{j=1}^m |f_j|^2 \sum_{j=1}^m |g_j|^2$$

with the aid of Example 1 and the Schwarz inequality.

1.2 Scalar products and norms

A positive sesquilinear form on H is called a *scalar product* (or *inner product*) on H. In what follows scalar products will be denoted mostly by $\langle\,.\,,\,.\,\rangle$ and occasionally they will be given an index in order to distinguish between them. A non-negative sesquilinear form is called a *semi-scalar product*. Examples for (semi-) scalar products may be obtained from the exercises in Section 1.1.

The mapping $s: H \times H \to \mathbb{K}$ is a scalar product if and only if for all $f, g, h \in H$ and $a \in \mathbb{K}$ we have

(i) $s(f, g+h) = s(f, g) + s(f, h)$,
(ii) $s(f, ag) = as(f, g)$,
(iii) $s(f, g) = s(g, f)^*$, (1.13)
(iv) $s(f, f) \geq 0$,
(v) $s(f, f) > 0$ if $f \neq 0$.

For the *proof* we only have to observe that the properties (1.1) and (1.2) follow from (i), (ii) and (iii). Similarly, *a mapping $s: H \times H \to \mathbb{K}$ is a semi-scalar product if and only if s satisfies properties* [(1.13) (i–iv)].

A mapping $p: H \to \mathbb{R}$ is called a *norm* on H if for all $f, g \in H$ and $a \in \mathbb{K}$ we have

(i) $p(f) \geq 0$,
(ii) $p(af) = |a| p(f)$, (1.14)
(iii) $p(f+g) \leq p(f) + p(g)$ (*triangle inequality*),
(iv) $p(f) > 0$ provided $f \neq 0$.

A mapping $p: H \to \mathbb{R}$ is called a *seminorm* on H if it satisfies the properties [(1.14) (i–iii)]. In what follows norms will mostly be denoted by $\|\cdot\|$ and for more precise distinctions they will occasionally be given different indices.

Remark. If p is a seminorm on H, then for all $f, g \in H$ we have

$$p(f \pm g) \geq |p(f) - p(g)|.$$

Proof. The triangle inequality implies

$$p(f) = p(f - g + g) \leq p(f-g) + p(g),$$

thus

$$p(f) - p(g) \leq p(f-g).$$

Similarly, $p(g) = p(g - f + f) \leq p(g-f) + p(f)$; thus

$$-(p(f) - p(g)) \leq p(f-g).$$

From these two inequalities $p(f-g) \geq |p(f) - p(g)|$ follows. One can show the inequality $p(f+g) \geq |p(f) - p(g)|$ in a similar way. □

Example 1. In $\mathbb{C}^m$ (or $\mathbb{R}^m$) let us define two norms by

$$\|f\|_1 = \sum_{j=1}^{m} |f_j| \quad \text{and} \quad \|f\|_\infty = \max\{|f_j| : j = 1, \ldots, m\}.$$

If $c_j \geq 0$ for $j = 1, 2, \ldots, m$, then by

$$p_1(f) = \sum_{j=1}^{m} c_j |f_j| \quad \text{and} \quad p_\infty(f) = \max\{c_j |f_j| : j = 1, \ldots, m\}$$

two seminorms are defined. These seminorms are norms if all the c_j are positive.

EXAMPLE 2. If r is a non-negative continuous function on $[0, 1]$, then by

$$p_1(f) = \int_0^1 r(x)|f(x)|\,dx$$

and

$$p_\infty(f) = \max\{r(x)|f(x)| : 0 \leq x \leq 1\}$$

two seminorms are defined on $C[0, 1]$. These are norms if r does not vanish identically on any non-trivial interval. For $r(x) = 1$ these norms will be denoted by $\|\cdot\|_1$ and $\|\cdot\|_\infty$, respectively:

$$\|f\|_1 = \int_0^1 |f(x)|\,dx,$$

$$\|f\|_\infty = \max\{|f(x)| : 0 \leq x \leq 1\}.$$

A large number of norms can be generated with the aid of scalar products because of the following theorem.

Theorem 1.5. *If s is a semi-scalar product on H, then $p(f) = [s(f,f)]^{1/2}$ defines a seminorm on H.*

If $\langle .\,,.\rangle$ is a scalar product on H, then $\|f\| = \langle f, f\rangle^{1/2}$ defines a norm on H.

PROOF. Property [(1.14) (i)] follows immediately from [(1.13) (iv)]; [(1.14) (iv)] follows from [(1.13) (v)]. It is sufficient to prove the remaining properties for the first case. Because of [(1.13) (ii)] and [(1.13) (iii)] we have

$$p(af) = [s(af, af)]^{1/2} = [|a|^2 s(f,f)]^{1/2} = |a|[s(f,f)]^{1/2} = |a|p(f),$$

which is [(1.14) (ii)]. With the aid of the Schwarz inequality it follows that

$$\begin{aligned} p(f+g)^2 &= p(f)^2 + 2\,\mathrm{Re}\, s(f,g) + p(g)^2 \\ &\leq p(f)^2 + 2|s(f,g)| + p(g)^2 \leq p(f)^2 + 2p(f)p(g) + p(g)^2 \\ &= (p(f) + p(g))^2, \end{aligned}$$

which is the triangle inequality [(1.14) (iii)]. □

From the Schwarz inequality for non-negative sesquilinear forms we obtain for the norm $\|\cdot\|$ (seminorm p) induced by a scalar product $\langle .\,,.\rangle$ (semi-scalar product s) that

$$|\langle f, g\rangle| \leq \|f\|\,\|g\|, \tag{1.15}$$

$$|s(f,g)| \leq p(f)p(g). \tag{1.15'}$$

Proposition. *If $\langle .\,,.\rangle$ is a scalar product on H and $\|\,.\,\|$ denotes the norm generated by it (cf. Theorem 1.5), then $\|f+g\| = \|f\| + \|g\|$ if and only if there exists an $a \geqslant 0$ such that $f = ag$ or $g = af$.*

PROOF. If $f = ag$ with $a \geqslant 0$, then we have

$$\|f+g\| = \|(1+a)g\| = (1+a)\|g\| = \|ag\| + \|g\| = \|f\| + \|g\|$$

(this part of the assertion holds for any norm). Conversely, if $\|f+g\| = \|f\| + \|g\|$ then

$$\|f\|^2 + 2\|f\|\,\|g\| + \|g\|^2 = \|f+g\|^2 = \|f\|^2 + 2\,\mathrm{Re}\,\langle f, g\rangle + \|g\|^2.$$

Thus $\mathrm{Re}\,\langle f, g\rangle = \|f\|\,\|g\|$. Using (1.15) this implies $\langle f, g\rangle = \|f\|\,\|g\|$. Now Theorem 1.4 gives the assertion. □

For a norm $\|\,.\,\|$ (seminorm p) induced by a scalar product (semi-scalar product) the *parallelogram identity*

$$\|f+g\|^2 + \|f-g\|^2 = 2(\|f\|^2 + \|g\|^2), \tag{1.16}$$

respectively

$$p(f+g)^2 + p(f-g)^2 = 2(p(f)^2 + p(g)^2), \tag{1.16'}$$

follows from Theorem 1.2.

If one considers a (semi-)norm as the length of a vector, then these equalities have the following geometric meaning: In a parallelogram the sum of the squares of the diagonals equals the sum of the squares of the sides. According to (1.4) [respectively (1.8)] the scalar product $\langle .\,,.\rangle$ (respectively semi-scalar product s) which we started with is given by the *polarization identity*

$$\langle f, g\rangle = \begin{cases} \frac{1}{4}\{\|f+g\|^2 - \|f-g\|^2 + i\|f-ig\|^2 - i\|f+ig\|^2\}, & \mathbb{K} = \mathbb{C}, \\ \frac{1}{4}\{\|f+g\|^2 - \|f-g\|^2\}, & \mathbb{K} = \mathbb{R}, \end{cases} \tag{1.17}$$

respectively

$$s(f, g) = \begin{cases} \frac{1}{4}\{p(f+g)^2 - p(f-g)^2 + ip(f-ig)^2 - ip(f+ig)^2\}, & \mathbb{K} = \mathbb{C}, \\ \frac{1}{4}\{p(f+g)^2 - p(f-g)^2\}, & \mathbb{K} = \mathbb{R}. \end{cases} \tag{1.17'}$$

The following theorem enables us to decide if a given (semi-)norm is generated by a (semi-)scalar product.

Theorem 1.6 (Jordan and von Neumann). *A norm $\|\,.\,\|$ on a vector space H is generated by a scalar product $\langle .\,,.\rangle$ in the sense of Theorem 1.5 if and only if the parallelogram identity (1.16) is satisfied. If this is so then the scalar product $\langle .\,,.\rangle$ is given by (1.17). A corresponding statement holds true for seminorms and semi-scalar products.*

Proof. If the norm $\|\,.\,\|$ is induced by the scalar product $\langle\,.\,,\,.\,\rangle$, then (1.16) holds true and the scalar product can be recaptured from the norm by means of (1.17). It remains to be shown that if $\|\,.\,\|$ satisfies the parallelogram identity and $\langle\,.\,,\,.\,\rangle$ is defined by (1.17), then $\langle\,.\,,\,.\,\rangle$ is a scalar product and generates the norm $\|\,.\,\|$. We restrict ourselves to the proof in the complex case; the real case goes analogously and is even a little simpler.

Let $\langle\,.\,,\,.\,\rangle$ be defined by (1.17). We show that $\langle\,.\,,\,.\,\rangle$ is a scalar product.

[(1.13) (iv–v)]: For all $f \in H$ by virtue of the definition of $\langle\,.\,,\,.\,\rangle$ we have

$$\begin{aligned}\langle f, f\rangle &= \tfrac{1}{4}\{\|f+f\|^2 - \|0\|^2 + i\|f - if\|^2 - i\|f + if\|^2\}\\ &= \tfrac{1}{4}\{4\|f\|^2 - 0 + 2i\|f\|^2 - 2i\|f\|^2\} = \|f\|^2.\end{aligned}$$

The properties [(1.13) (iv–v)] of $\langle\,.\,,\,.\,\rangle$ now follow from the corresponding properties of the norm $\|\,.\,\|$. At the same time we obtain that $\|\,.\,\|$ is generated by $\langle\,.\,,\,.\,\rangle$.

[(1.13) (iii)]: For all $f, g \in H$ we have

$$\begin{aligned}\langle g, f\rangle^* &= \tfrac{1}{4}\{\|g+f\|^2 - \|g-f\|^2 + i\|g - if\|^2 - i\|g + if\|^2\}^*\\ &= \tfrac{1}{4}\{\|f+g\|^2 - \|f-g\|^2 + i\|f + ig\|^2 - i\|f - ig\|^2\}\\ &= \langle f, g\rangle.\end{aligned}$$

[(1.13) (i)]: For all $f, g, h \in H$ because of (1.16) we have

$$\begin{aligned}&\langle f, g\rangle + \langle f, h\rangle\\ &\quad= \tfrac{1}{4}\{\|f+g\|^2 - \|f-g\|^2 + i\|f - ig\|^2 - i\|f + ig\|^2\\ &\qquad + \|f+h\|^2 - \|f-h\|^2 + i\|f - ih\|^2 - i\|f + ih\|^2\}\\ &\quad= \tfrac{1}{4}\Big\{\Big\|\Big(f + \frac{g+h}{2}\Big) + \frac{g-h}{2}\Big\|^2 + \Big\|\Big(f + \frac{g+h}{2}\Big) - \frac{g-h}{2}\Big\|^2\\ &\qquad - \Big\|\Big(f - \frac{g+h}{2}\Big) + \frac{g-h}{2}\Big\|^2 - \Big\|\Big(f - \frac{g+h}{2}\Big) - \frac{g-h}{2}\Big\|^2\\ &\qquad + i\Big\|\Big(f - i\frac{g+h}{2}\Big) + i\frac{g-h}{2}\Big\|^2 + i\Big\|\Big(f - i\frac{g+h}{2}\Big) - i\frac{g-h}{2}\Big\|^2\\ &\qquad - i\Big\|\Big(f + i\frac{g+h}{2}\Big) + i\frac{g-h}{2}\Big\|^2 - i\Big\|\Big(f + i\frac{g+h}{2}\Big) - i\frac{g-h}{2}\Big\|^2\Big\}\\ &\quad= \frac{1}{2}\Big\{\Big\|f + \frac{g+h}{2}\Big\|^2 + \Big\|\frac{g-h}{2}\Big\|^2 - \Big\|f - \frac{g+h}{2}\Big\|^2 - \Big\|\frac{g-h}{2}\Big\|^2\\ &\qquad + i\Big\|f - i\frac{g+h}{2}\Big\|^2 + i\Big\|i\frac{g-h}{2}\Big\|^2 - i\Big\|f + i\frac{g+h}{2}\Big\|^2 - i\Big\|i\frac{g-h}{2}\Big\|^2\Big\}\\ &\quad= 2\Big\langle f, \frac{g+h}{2}\Big\rangle. \qquad (1.18)\end{aligned}$$

Since by (1.17) we obviously have $\langle f, 0\rangle = 0$, from (1.18) it follows by substituting $h = 0$ that

$$2\langle f, \tfrac{1}{2}g\rangle = \langle f, g\rangle. \qquad (1.19)$$

From (1.18) and (1.19) it follows that

$$\langle f, g\rangle + \langle f, h\rangle = 2\left\langle f, \frac{g+h}{2}\right\rangle = \langle f, g+h\rangle,$$

which is the required property.

[(1.13) (iii)]: We already know that $\langle f, g\rangle = 2\langle f, g/2\rangle$. From this and from property [(1.13) (ii)] we obtain by induction that

$$\langle 2^{-n} m f, g\rangle = \langle f, 2^{-n} m g\rangle \quad \text{for all} \quad n, m \in \mathbb{N}_0$$

($\mathbb{N}_0$ is the set of non-negative integers $\{0, 1, 2, \ldots\}$). If $a \geq 0$, then there exist numbers $a_k = 2^{-n(k)} m(k)$ such that $a_k \to a$ as $k \to \infty$. By the proposition preceding Example 1 we have

$$\big|\|f \pm a_k g\| - \|f \pm a g\|\big| \leq |a_k - a|\,\|g\|,$$
$$\big|\|f \pm i a_k g\| - \|f \pm i a g\|\big| \leq |a_k - a|\,\|g\|,$$

therefore because of (1.17)

$$\langle f, a_k g\rangle \to \langle f, a g\rangle \quad \text{as} \quad k \to \infty.$$

From this it follows that

$$a\langle f, g\rangle = \lim_{k\to\infty} a_k\langle f, g\rangle = \lim_{k\to\infty} \langle f, a_k g\rangle = \langle f, a g\rangle.$$

Furthermore, we have

$$\begin{aligned}\langle f, -g\rangle &= \tfrac{1}{4}\left(\|f-g\|^2 - \|f+g\|^2 + i\|f+ig\|^2 - i\|f-ig\|^2\right)\\ &= -\langle f, g\rangle;\end{aligned}$$

consequently $\langle f, ag\rangle = a\langle f, g\rangle$ for all $a \in \mathbb{R}$. As we also have

$$\begin{aligned}\langle f, ig\rangle &= \tfrac{1}{4}\left(\|f+ig\|^2 - \|f-ig\|^2 + i\|f+g\|^2 - i\|f-g\|^2\right)\\ &= i\langle f, g\rangle,\end{aligned}$$

The equality $\langle f, ag\rangle = a\langle f, g\rangle$ follows for all $a \in \mathbb{C}$. The proof for semi-norms is completely analogous. □

If H is a (complex or real) vector space and $\langle\cdot,\cdot\rangle$ is a scalar product on H, then we call the pair $(H, \langle\cdot,\cdot\rangle)$ a *vector space with scalar product* or a *pre-Hilbert space*. If it is clear which scalar product is meant on H, then we shall briefly write H for the pair mentioned. If $\|\cdot\|$ is a norm on H, then we call the pair $(H, \|\cdot\|)$ a normed space. Here we shall also only write H in most cases. By Theorem 1.5 the norm $\|f\| = \langle f, f\rangle^{1/2}$ is defined in a natural way on every pre-Hilbert space. Therefore in what follows we shall consider every pre-Hilbert space as a normed space.

EXAMPLE 3. On $\mathbb{C}^m$ respectively $\mathbb{R}^m$ by

$$\langle f, g\rangle = \sum_{j=1}^{m} f_j^* g_j$$

a scalar product is defined. The corresponding norm

$$\|f\| = \left\{ \sum_{j=1}^{m} |f_j|^2 \right\}^{1/2}$$

is the *Euclidean length* of the vector f, thus $\|f - g\|$ is the *Euclidean distance* of the points f and g.

EXAMPLE 4. On $C[0, 1]$ by

$$\langle f, g \rangle = \int_0^1 f(x)^* g(x)\, dx, \quad \|f\| = \left\{ \int_0^1 |f(x)|^2\, dx \right\}^{1/2}$$

a scalar product and the corresponding norm are defined.

EXAMPLE 5. Let l_2 be the *Hilbert sequence space*, i.e., the set of (real or complex) sequences $f = (f_n) = (f_1, f_2, \ldots)$ for which $\sum_{n=1}^{\infty} |f_n|^2 < \infty$. Then l_2 will be a (real or complex) vector space if one defines addition and multiplication as follows:

$$f + g = (f_n + g_n), \; af = (af_n) \quad \text{for} \quad f, g \in l_2 \quad \text{and} \quad a \in \mathbb{K}.$$

It is clear that this definition of multiplication is meaningful since along with $\sum_{n=1}^{\infty} |f_n|^2 < \infty$ we also have $\sum_{n=1}^{\infty} |af_n|^2 < \infty$. If f and g are in l_2, then for every $N \in \mathbb{N}$ we have

$$\sum_{n=1}^{N} |f_n + g_n|^2 \leq 2\left\{ \sum_{n=1}^{N} |f_n|^2 + \sum_{n=1}^{N} |g_n|^2 \right\} \leq 2\left\{ \sum_{n=1}^{\infty} |f_n|^2 + \sum_{n=1}^{\infty} |g_n|^2 \right\},$$

consequently we also have

$$\sum_{n=1}^{\infty} |f_n + g_n|^2 \leq 2\left\{ \sum_{n=1}^{\infty} |f_n|^2 + \sum_{n=1}^{\infty} |g_n|^2 \right\} < \infty$$

i.e., $f + g \in l_2$. It is easy to see that by

$$\langle f, g \rangle = \sum_{n=1}^{\infty} f_n^* g_n, \quad f, g \in l_2$$

a scalar product is defined on l_2; the series converges, because $|f_n^* g_n| \leq (|f_n|^2 + |g_n|^2)/2$. The induced norm is

$$\|f\| = \left\{ \sum_{n=1}^{\infty} |f_n|^2 \right\}^{1/2}$$

Unless otherwise stated, in what follows l_2 will always denote the complex sequence space.

Exercises

1.6. The norms in Examples 1 and 2 are not generated by scalar products.

1.7. The proposition after Theorem 1.5 does not hold true in general for norms that are not generated by scalar products.

1.8. Let p be a seminorm on H generated by a semi-scalar product and let $N = \{f \in H : p(f) = 0\}$. We have $p(f+g) = p(f) + p(g)$ if and only if there exists an $\alpha \geq 0$ such that $f - \alpha g \in N$ or $g - \alpha f \in N$.

1.9. (a) Let A^2 be the set of functions f holomorphic on $\mathbb{C}_1 = \{z \in \mathbb{C} : |z| < 1\}$ for which

$$\int_{\mathbb{C}_1} |f(x+iy)|^2 \, dx \, dy < \infty$$

(the integral can be understood as an improper Riemann integral or as a Lebesgue integral). A^2 is a vector space. By

$$\langle f, g \rangle_1 = \int_{\mathbb{C}_1} f(x+iy)^* g(x+iy) \, dx \, dy, \quad \|f\|_1^2 = \int_{\mathbb{C}_1} |f(x+iy)|^2 \, dx \, dy$$

a scalar product and the corresponding norm are defined on A^2.

(b) Let H^2 be the set of functions f holomorphic on $\mathbb{C}_1$ for which the limit

$$\lim_{r \to 1-0} \int_0^{2\pi} |f(re^{it})|^2 \, dt$$

is finite. H^2 is a vector space (*Hardy class*). By

$$\|f\|_2 = \left\{ \lim_{r \to 1-0} \int_0^{2\pi} |f(re^{it})|^2 \, dt \right\}^{1/2}$$

a norm is defined on H^2. This norm is generated by the scalar product

$$\langle f, g \rangle_2 = \lim_{r \to 1-0} \int_0^{2\pi} f(re^{it})^* g(re^{it}) \, dt.$$

(c) If $f(z) = \sum_{n=0}^{\infty} f_n z^n$, $g(z) = \sum_{n=0}^{\infty} g_n z^n$ are the Taylor series of f and g, then we have

$$\langle f, g \rangle_1 = \pi \sum_{n=0}^{\infty} \frac{1}{n+1} f_n^* g_n, \quad \langle f, g \rangle_2 = 2\pi \sum_{n=0}^{\infty} f_n^* g_n.$$

(d) H^2 is a subspace of A^2 and we have $\|f\|_1^2 \leq \frac{1}{2}\|f\|_2^2$ for $f \in H^2$.

(e) For all $f \in H^2$ we have

$$\|f\|_2^2 = \sup\left\{ \int_0^{2\pi} |f(re^{it})|^2 \, dt : 0 \leq r < 1 \right\}.$$

1.10. Let A be an arbitrary set, let $\mu : A \to (0, \infty)$, and let $l_2(A; \mu)$ be the set of functions $f : A \to \mathbb{C}$ that vanish outside a countable set (that may vary with f) and for which $\sum_{\alpha \in A} \mu(\alpha) |f(\alpha)|^2 < \infty$.

(a) $l_2(A; \mu)$ is a subspace of the space of all complex-valued functions on A.

(b) By

$$\langle f, g \rangle = \sum_{\alpha \in A} \mu(\alpha) f(\alpha)^* g(\alpha), \quad f, g \in l_2(A; \mu),$$

a scalar product is defined on $l_2(A; \mu)$.