Chapter 4

Operators on Hilbert space

In this chapter we first recall in section 4.1 some basic facts about matrix represention of linear mappings defined on finite dimensional real Hilbert spaces. In section 4.2 their immediate generalisation to finite dimensional complex Hilbert spaces is described. Linear mappings defined on infinite dimensional Hilbert spaces are introduced in section 4.3, including some illustrative examples. As is usual, we generally use the name *linear operator* or just *operator* instead of linear mapping in the following. For the sake of technical simplicity the main focus is on continuous (also called *bounded*) operators, although many operators relevant in physics, such as differential operators, are actually not bounded. The *adjoint* of an operator is defined and the basic properties of the adjoint opeation are established. This allows the introduction of *self-adjoint* operators (corresonding to symmetric (or Hermitean matrices) which together with *diagonalisable operators* (corresonding to diagonalisable matrices) are the subject of section 4.4. In section 4.5 we define *unitary operators* (corresponding to orthogonal matrices) and discuss the Fourier transformation as an important example. Finally, section 4.6 contains some remarks on Dirac notation.

4.1 Operators on finite dimensional real Hilbert spaces

In this section H denotes a real Hilbert space of dimension $\dim H = N < \infty$, and $\alpha = (e_1, \ldots, e_N)$ denotes an orthonormal basis for H. For any given vector $x \in H$ its coordinates w.r.t. α will be denoted by $x_1, \ldots, x_N$, that is $x = x_1e_1 + \ldots + x_Ne_N$, and by $\underline{x}$ we denote the corresponding coordinate clumn vector,

$$\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} .$$

We first note that every linear mapping $A : H \to H'$, where H' is a second Hilbert space whose norm we denote by $\|\cdot\|'$, is continuous. Indeed, for $x = x_1e_1+\ldots+x_Ne_N \in H$, we have

$$Ax = x_1Ae_1 + \ldots + x_NAe_N \tag{4.1}$$

and therefore

$$\begin{aligned}\|Ax\|' &= \|x_1 Ae_1 + \ldots + x_N Ae_N\|' \\ &\leq |x_1|\ \|Ae_1\|' + \ldots + |x_N|\ \|Ae_N\|' \\ &\leq (\|Ae_1\|' + \ldots + \|Ae_N\|') \max\{|x_1|, \ldots, |x_N|\} \\ &\leq (\|Ae_1\|' + \ldots + \|Ae_N\|')\|x\|\ ,\end{aligned}$$

where we have used the triangle inequality and

$$\|x\|^2 = |x_1|^2 + \ldots + |x_N|^2 \geq (\max\{|x_1|, \ldots, |x_N|\})^2.$$

This shows that A is *bounded* in the sense that there exists a constant $c \geq 0$, such that

$$\|Ax\|' \leq c\,\|x\| \quad \text{for all } x \in H\,. \tag{4.2}$$

The *norm* $\|A\|$ of a bounded operator $A : H \to H'$ is by definition the smallest number c for which eq. (4.2) holds. Equivalently,

$$\|A\| = \sup\{\|Ax\|' \mid \|x\| = 1\}\ , \tag{4.3}$$

see Exercise 5.

It is a general fact, shown in Exercise 6, that an operator $A : H \to H'$, where H and H' are arbitrary Hilbert spaces, is continuous if and only if it is bounded. Thus we have shown that, if H is of finite dimension, then every operator $A : H \to H'$ is bounded and hence continuous.

For the sake of simplicity we now assume that $H = H'$. As is well known from linear algebra (see section 6.3 in [M]) a linear operator $A : H \to H$ is represented w.r.t. the basis α by an $N \times N$-matrix $\underline{\underline{A}}$ in the sense that the relation between the coordinate set for a vector $x \in H$ and its image $y = Ax$ is given by

$$\underline{y} = \underline{\underline{A}}\ \underline{x}\ . \tag{4.4}$$

Moreover, matrix multiplication is defined in such a way that, if the operators A and B on H are represented by the matrices $\underline{\underline{A}}$ og $\underline{\underline{B}}$ w.r.t. α, then the composition AB is represented by the product $\underline{\underline{A}}\ \underline{\underline{B}}$ w.r.t. α (Theorem 6.13 in [M]).

The basis α being orthonormal, the matrix elements a_{ij}, $1 \leq i, j \leq N$, of $\underline{\underline{A}}$ are given by the formula

$$a_{ij} = (Ae_j, e_i) \tag{4.5}$$

as a consequence of the following calculation:

$$\begin{aligned}y_i &= (Ax, e_i) \\ &= (x_1 Ae_1 + \ldots + x_N Ae_N, e_i) \\ &= \sum_{j=1}^{N} (Ae_j, e_i) x_j\ .\end{aligned}$$

Now let A^* denote the operator on H represented by the transpose $\underline{\underline{A}}^t$ of $\underline{\underline{A}}$. According to eq. (4.5) we have

$$(Ae_i, e_j) = a_{ji} = (A^* e_j, e_i) = (e_i, A^* e_j)\ ,$$

where in the last step we have used the symmetry of the inner produkt. Since the inner product is linear in both variables, we get for $x = \sum_{i=1}^{N} x_i e_i$ and $z = \sum_{j=1}^{N} z_j e_j$ in H that

$$(Ax, z) = \sum_{i,j=1}^{N} x_i z_j (Ae_i, e_j) = \sum_{i,j=1}^{N} x_i z_j (e_i, A^* e_j) = (x, A^* z) \; . \tag{4.6}$$

We claim that the validity of this identity for all $x, z \in H$ implies that A^* is uniquely determined by A and does not depend on the choice of orthonormal basis α. Indeed, if the operator B on H satisfies $(Ax, z) = (x, Bz)$ for all $x, z \in H$, then $(x, A^*z - Bz) = 0$ for all $x, z \in H$, i.e. $A^*z - Bz \in H^\perp = \{\underline{0}\}$ for all $z \in H$. This shows that $A^*z = Bz$ for all $z \in H$, that is $A^* = B$. The operator A^* is called the *adjoint operator* of A. If $A = A^*$, we say that A is *self-adjoint*. By the definition of A^* we have that *the self-adjoint operators on a real finite dimensional Hilbert space are precisely those operators that are represented by symmetric matrices w.r.t. an arbitrary orthonormal basis for H.*

It is known from linear algebra (see section 8.4 in [M]), that every symmetric $N \times N$-matrix $\underline{\underline{A}}$ can be diagonalised by an orthognal matrix, that is there exists an orthogonal $N \times N$-matrix $\underline{\underline{O}}$ such that

$$\underline{\underline{O}}^{-1} \underline{\underline{A}}\, \underline{\underline{O}} = \underline{\underline{D}} \; , \tag{4.7}$$

where $\underline{\underline{D}} = \Delta(\lambda_1, \ldots, \lambda_N)$ is an $N \times N$ diagonal matrix with the eigenvalues $\lambda_1, \ldots, \lambda_N$ of $\underline{\underline{A}}$ in the diagonal. That $\underline{\underline{O}}$ is orthogonal means that the columns (and therefore also the rows) of $\underline{\underline{O}}$ form an orthonormal basis for $\mathbb{R}^N$, which is equivalent to

$$\underline{\underline{O}}^t \underline{\underline{O}} = \underline{\underline{I}} \; , \tag{4.8}$$

where I denotes the $N \times N$ identity matrix. Thus we have that $\underline{\underline{O}}$ is orthogonal if and only if $\underline{\underline{O}}$ is invertible and

$$\underline{\underline{O}}^t = \underline{\underline{O}}^{-1} \; . \tag{4.9}$$

Eq. (4.7) expresses that the columns of $\underline{\underline{O}}$ are eigenvectors for $\underline{\underline{A}}$. Therefore, we have shown that for every symmetric matrix $\underline{\underline{A}}$ there exists an orthonormal basis for $\mathbb{R}^N$ consisting of egenvectors for $\underline{\underline{A}}$.

Let now $\underline{\underline{A}}$ represent a self-adjoint operator A w.r.t. the basis α as above, and let the orthogonal matrix $\underline{\underline{O}} = (o_{ij})$ be chosen according to (4.7). We denote by $O : H \to H$ the operator represented by the matrix $\underline{\underline{O}}$ w.r.t. α. Similarly, let $D : H \to H$ denote the operator represented by $\underline{\underline{D}}$ w.r.t. α, that is

$$De_i = \lambda_i e_i \; , \qquad i = 1, \ldots, N \; . \tag{4.10}$$

Then eq. (4.7) is equivalent to

$$O^{-1} A O = D \; , \tag{4.11}$$

since the matrix representing $O^{-1}AO$ w.r.t. α is $\underline{\underline{O}}^{-1} \underline{\underline{A}}\, \underline{\underline{O}}$.

Similarly, eqs. (4.8) and (4.9) are equivalent to

$$O^*O = 1 \tag{4.12}$$

and

$$O^* = O^{-1} \; , \tag{4.13}$$

respectively, where 1 denotes the identity operator on H. An operator O fulfilling (4.13) is called an *orthogonal operator*. Thus *orthogonal operators are precisely those operators that are represented by orthogonal matrices w.r.t. an arbitrary orthonormal basis.*

Setting

$$f_i = Oe_i \, , \qquad i = 1, \ldots, N \; ,$$

we get

$$(f_i, f_j) = (Oe_i, Oe_j) = (e_i, O^*Oe_j) = (e_i, e_j) = \delta_{ij} \; , \tag{4.14}$$

which means that $(f_1, \ldots, f_N)$ is an orthonormal basis for H. In other words, *orthogonal operators map orthonormal bases to orthonormal bases.*

It now follows from (4.10) and (4.11) that

$$Af_i = AOe_i = ODe_i = O(\lambda_i e_i) = \lambda_i Oe_i = \lambda_i f_i \; . \tag{4.15}$$

A vector $x \in H \setminus \{0\}$ such that the image Ax er proportional to x, i.e. such that there exists a $\lambda \in \mathbb{R}$, such that

$$Ax = \lambda x \; , \tag{4.16}$$

is called an *eigenvector* for A, and λ is called the corresponding *eigenvalue*. From (4.14) and (4.15) we thus conclude that *for every self-adjoint operator A on a finite dimensianal real Hilbert space there exists an orthonormal basis consisting of eigenvctors for A.* We say that such a basis diagonalises A, since the matrix representing A w.r.t. this basis is the diagonal matrix $\underline{\underline{D}}$, whose diagonal elements are the eigenvalues of A.

4.2 Operators on finite dimensional complex Hilbert spaces

In this section H denotes a finite dimensional complex Hilbert space and $\alpha = (e_1, \ldots, e_N)$ again denotes an orthonormal basis for H.

By the same argument as in the previous section (see (4.1)) every operator $A : H \to H$ is bounded. And by the same calculations as those leading to (4.6) A is represented w.r.t. α by the complex matrix $\underline{\underline{A}} = (a_{ij})$, where

$$a_{ij} = (Ae_j, e_i) \, , \qquad 1 \leq i, j \leq N \; .$$

Let A^* denote the operator, which is represented w.r.t. α by the matrix $\underline{\underline{A}}^*$, obtained from $\underline{\underline{A}}$ by transposition and complex conjugation. We say that $\underline{\underline{A}}^*$ is the *Hermitean conjugate* of $\underline{\underline{A}}$. That is, we have

$$(Ae_i, e_j) = a_{ji} = \overline{(A^*e_j, e_i)} = (e_i, A^*e_j) \; .$$

For $x = \sum_{i=1}^{N} x_i e_i$ and $z = \sum_{j=1}^{N} z_j e_j$ in H we then get

$$(Ax, z) = \sum_{i,j=1}^{N} x_i \overline{z}_j (Ae_i, e_j) = \sum_{i,j=1}^{N} x_i \overline{z}_j (e_i, A^* e_j) = (x, A^* z) \ .$$

By a similar argument as in the previous section it follows that the operator A^* is uniquely determined by A. It is called the adjoint operator of A. If $A = A^*$, we say that A is self-adjoint. It follows, in particular, that if A is self-adjoint, then $\underline{\underline{A}}^* = \underline{\underline{A}}$, and then the matrix $\underline{\underline{A}}$ is called *Hermitean.*

We note that addition and multiplication of comples matrices are defined in the same way and satisfy the same baisc laws of calculation as for real matrices. Moreover, well known concepts such as rank and determinant are defined in an analogous manner. Results and proofs can be transported directly from the real to the complex case. It suffices here to mention that *a quadratic matrix is regular, or invertible, if and only if its determinant is different from zero.* We note, moreover, that for a quadratic matrix $\underline{\underline{A}}$ we have

$$\det \underline{\underline{A}}^* = \overline{\det \underline{\underline{A}}} \ ,$$

which follows immediately from $\det \underline{\underline{A}}^t = \det \underline{\underline{A}}$.

The complex counterpart of an ortogonal operator is called a *unitary operator.* Thus, we say that $U : H \to H$ is unitary, if it fulfills

$$U^* U = 1 \ , \tag{4.17}$$

that is

$$U^* = U^{-1} \ . \tag{4.18}$$

Letting $\underline{\underline{U}}$ denote the matrix representing U w.r.t. α, eq. (4.17) is equivalent to

$$\underline{\underline{U}}^* \underline{\underline{U}} = \underline{\underline{I}} \ . \tag{4.19}$$

A matrix $\underline{\underline{U}}$ fulfilling (4.19), is called *unitary.* Evidently, a real unitary matrix is the same thing as an orthogonal matrix. We also remark that (4.19) is equivalent to the statement that the columns of $\underline{\underline{U}}$ form an orthonormal basis for $\mathbb{C}^N$.

Eq. (4.17) implies

$$(Ux, Uy) = (x, U^* U y) = (x, y) \ , \quad x, y \in H \ , \tag{4.20}$$

i.e. any unitary operator U preserves the inner product. In particular, U is an isometry,

$$\|Ux\| = \|x\| \ , \qquad x \in H \ . \tag{4.21}$$

The following result shows, among other things, that (4.21) is in fact equivalent to U being unitary.

Theorem 4.1 *Let $U : H \to H$ be a linear operator. The following four statements are equivalent.*

a) U is unitary.

b) U preserves the inner product.

c) U is an isometry.

d) U maps an orthonormal basis into an orthonormal basis for H.

Proof. Above we have shown a) $\Rightarrow$ b) $\Rightarrow$ c).

c) $\Rightarrow$ b): This follows by using the polarisation identity for the two sesquilinear forms (x, y) and (Ux, Uy). Indeed, if U is an isometry, we have

$$(Ux, Uy) = \tfrac{1}{4}\sum_{\nu=0}^{3} i^\nu \|U(x + i^\nu y)\|^2 = \tfrac{1}{4}\sum_{\nu=0}^{3} i^\nu \|x + i^\nu y\|^2 = (x, y)$$

for arbitrary $x, y \in H$.

b) $\Rightarrow$ a): We have that $(Ux, Uy) = (x, y)$ for all $x, y \in H$ if and only if $(x, U^*Uy) = (x, y)$ holds for all $x, y \in H$, which is the case if and only if $U^*Uy = y$ for all $y \in H$. This means that $U^*U = 1$ as claimed.

We have now demonstrated that a), b) and c) are equivalent.

b) $\Rightarrow$ d): Setting $f_i = Ue_i$, we get

$$(f_i, f_j) = (Ue_i, Ue_j) = (e_i, e_j) = \delta_{ij}\,, \quad 1 \le i, j \le N\,,$$

which shows that $(f_1, \ldots, f_N)$ is an orthonormal basis for H.

d) $\Rightarrow$ c): For $x = x_1e_1 + \ldots + x_Ne_N \in H$ we have

$$\|Ux\|^2 = \|x_1Ue_1 + \ldots + x_NUe_N\|^2 = |x_1|^2 + \ldots + |x_N|^2 = \|x\|^2\,,$$

where we have used Pythagoras' theorem twice. □

In relation to Theorem 4.1 we note that for any pair of orthonormal bases $(e_1, \ldots, e_N)$ and $(f_1, \ldots, f_N)$ for H there is exactly one operator U, which maps the first basis onto the second, i.e. such that

$$Ue_i = f_i\,, \qquad i = 1, \ldots, N\,,$$

and it is given by

$$U(x_1e_1 + \ldots + x_Ne_N) = x_1f_1 + \ldots + x_Nf_N\,.$$

This operator is unitary by Theorem 4.1.

We define eigenvectors and eigenvalues for an operator $A : H \to H$ as in the previous section (cf. (4.16)), except that the eigenvalue now may assume complex values instead of real values only. We say that A is diagonalisable if there exists an orthonormal basis for H consisting of eigenvectors for A. Equivalently, this means that there exists an orthonormal basis $(f_1,\dots,f_N)$ w.r.t. which A is represented by a diagonal matrix. Indeed, if $(f_1,\dots,f_N)$ is an orthonormal basis consisting of eigenvectors, such that $Af_i = \lambda_i f_i,\ i = 1,\dots,N$, then

$$a_{ij} = (Af_j, f_i) = (\lambda_j f_j, f_i) = \lambda_j \delta_{ij}\,, \qquad 1 \le i,j \le N\ ,$$

and, conversely, if $(f_1,\dots,f_N)$ is an orthonormal basis, such that

$$a_{ij} = (Af_j, f_i) = \lambda_j \delta_{ij}\,, \qquad 1 \le i,j \le N\ ,$$

then

$$Af_j = \sum_{i=1}^{N} (Af_j, f_i) f_i = \sum_{i=1}^{N} \lambda_j \delta_{ij} f_i = \lambda_j f_j$$

for $1 \le j \le N$.

In view of the discussion in the previous section it is natural to ask if every self-adjoint operator A on H is diagonalisable. The answer is yes, and the same holds for unitary operators, as we shall see in section 4.4.

We end this section by looking at a couple of simple examples.

Example 4.2 a) Let H be a 2-dimensional complex Hilbert space and let $\alpha = (e_1, e_2)$ be an orthonormal basis for H. We consider the operator A on H, which is represented w.r.t. α by the matrix

$$\underline{\underline{A}} = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}\ .$$

Then A is self-adjoint since $\underline{\underline{A}}^* = \underline{\underline{A}}$. We determine the eigenvalues of A by solving the characteristic equation $\det(\underline{\underline{A}} - \lambda \underline{\underline{I}}) = 0$, since this condition ensures (as in the case of real matrices) that the system of linear equations $(\underline{\underline{A}} - \lambda \underline{\underline{I}})\underline{x} = 0$ has a non-trivial solution $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, which is equivalent to stating that the vector $x = x_1 e_1 + x_2 e_2$ satisfies $Ax = \lambda x$.

The characteristic equation is

$$\det \begin{pmatrix} 1-\lambda & i \\ -i & 1-\lambda \end{pmatrix} = (1-\lambda)^2 - 1 = 0\ ,$$

which gives $\lambda = 0$ or $\lambda = 2$. For $\lambda = 0$ the equation $(\underline{\underline{A}} - \lambda \underline{\underline{I}})\binom{x_1}{x_2} = 0$ is equivalent to $x_1 + i\,x_2 = 0$, whose solutions are of the form $t\binom{1}{i}$, $t \in \mathbb{C}$.

For $\lambda = 2$ we get similarly the solutions $t\binom{1}{-i}$, $t \in \mathbb{C}$.

Two normalized eigencolumns corresponding to $\lambda = 0$ og $\lambda = 2$, respectively, are then

$$\underline{x}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \quad \text{og} \quad \underline{x}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}\ .$$

The matrix

$$\underline{\underline{U}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$$

is seen to be unitary, and we conclude that A is diagonalisable, since

$$\underline{\underline{U}}^{-1}\underline{\underline{A}}\,\underline{\underline{U}} = \underline{\underline{U}}^{*}\underline{\underline{A}}\,\underline{\underline{U}} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} .$$

The unitary operator, represented by $\underline{\underline{U}}$ w.r.t. α, maps the basis α onto the orthonormal basis $\big(\frac{1}{\sqrt{2}}(e_1 + ie_2), \frac{1}{\sqrt{2}}(e_1 - ie_2)\big)$ consisting of eigenvectors for A.

b) The real matrix

$$\underline{\underline{O}} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

is seen at to be orthogonal by inspection. It has no real eigenvalues and is therefore not diagonalisable over $\mathbb{R}$.

On the other hand, considering it as a complex matrix representing an operator O on a 2-dimensional complex Hilbert space H w.r.t. an orthonormal basis $\alpha = (e_1, e_2)$, then O is unitary. We find its eigenvalues by solving

$$\det \begin{pmatrix} \frac{1}{\sqrt{2}} - \lambda & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} - \lambda \end{pmatrix} = 0 .$$

This gives $\lambda = \frac{1}{\sqrt{2}}(1 \pm i)$. Corresponding eigencolumns for $\underline{\underline{O}}$ are as in the previous example $\underline{x}_1 = \frac{1}{\sqrt{2}}\binom{1}{i}$ and $\underline{x}_2 = \frac{1}{\sqrt{2}}\binom{1}{-i}$.

Since

$$\underline{\underline{U}}^{*}\underline{\underline{O}}\,\underline{\underline{U}} = \begin{pmatrix} \frac{1+i}{\sqrt{2}} & 0 \\ 0 & \frac{1-i}{\sqrt{2}} \end{pmatrix} ,$$

we conclude that O is diagonalisable and that $\big(\frac{1}{\sqrt{2}}(e_1 + ie_2), \frac{1}{\sqrt{2}}(e_1 - ie_2)\big)$ is an orthonormal basis for H, that diagonalises O.

4.3 Operators on Hilbert space. Adjoint operators.

In the remainder of this chapter H, H_1 and H_2 denote arbitrary separable Hilbert spaces over $\mathbb{L}$, where $\mathbb{L}$ denotes either $\mathbb{R}$ or $\mathbb{C}$. Results are mostly formulated for the case where the Hilbert spaces are infinite dimensional, but all results hold with obvious modifications for finite dimensional Hilbert spaces as well.

We shall use the notation $\mathcal{B}(H_1, H_2)$ for the set of bounded operators from H_1 to H_2 and set $\mathcal{B}(H) = \mathcal{B}(H, H)$. See (4.2) for the definition af a bounded operator. In Exercise 1 it is shown that $\mathcal{B}(H_1, H_2)$ is a vector space and that the norm defined by (4.3) fulfills the standard requirements (see eqs. (3.3-5)).

Given an operator $A \in \mathcal{B}(H)$ and an orthonormal basis $(e_i)_{i\in\mathbb{N}}$ for H we define the matrix

$$(a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1n} & \cdots \\ \vdots & & \vdots & \\ a_{n1} & \cdots & a_{nn} & \cdots \\ \vdots & & \vdots & \end{pmatrix} \tag{4.22}$$

that represents A w.r.t. $(e_i)_{i\in\mathbb{N}}$ by the same formula as in the finite dimensional case

$$a_{ij} = (Ae_j, e_i) \ . \tag{4.23}$$

For $x = \sum\limits_{j=1}^{\infty} x_j e_j \in H$ we then have

$$Ax = \sum_{i=1}^{\infty} y_i e_i \ ,$$

where

$$y_i = (Ax, e_i) = \big(\sum_{j=1}^{\infty} x_j Ae_j, e_i\big) = \sum_{j=1}^{\infty} x_j (Ae_j, e_i) \ . \tag{4.24}$$

In the second step we have used linearity and continuity of A to obtain

$$\begin{aligned} Ax &= A\big(\lim_{N\to\infty} \sum_{j=1}^{N} x_j e_j\big) = \lim_{N\to\infty} A\big(\sum_{j=1}^{N} x_j e_j\big) \\ &= \lim_{N\to\infty} \sum_{j=1}^{N} x_j Ae_j = \sum_{j=1}^{\infty} x_j Ae_j \ , \end{aligned} \tag{4.25}$$

and in the last step we have similarly used linearity and continuity of the map $x \to (x, e_i)$. From (4.23) and (4.24) it is seen that

$$y_i = \sum_{j=1}^{\infty} a_{ij} x_j \ ,$$

which is the infinite dimensional version of (4.4) and, in particular, shows that the matrix (4.22) determines the operator A uniquely.

In the infinite dimensional case matrix representations are generally of rather limited use since calculations involve infinite series and can only rarely can be performed explicitely. Moreover, the notion of determinant is not immediately generalisable to the infinite dimensional case, which implies that the standard method for determining eigenvalues and eigenvectors is not available any more. The coordinate independent operator point of view to be developed in the following will turn out more advantageous. First we shall consider a few imortant examples of operators on infinite dimensional Hilbert spaces.

Example 4.3 a) Let $H = \ell_2(\mathbb{N})$. For $x = (x_n)_{n\in\mathbb{N}} \in \ell_2(\mathbb{N})$ we define

$$Ax = (\frac{1}{n} x_n)_{n\in\mathbb{N}} \ . \tag{4.26}$$

Since

$$\sum_{n=1}^{\infty} \frac{1}{n^2}|x_n|^2 \leq \sum_{n=1}^{\infty} |x_n|^2 = \|x\|^2 \ , \tag{4.27}$$

it follows that $Ax \in \ell_2(\mathbb{N})$. We conclude that eq. (4.26) defines a mapping A from H to H. It is easily seen that A is linear (verify this!), and from (4.27) it then follows that $\|A\| \leq 1$. In fact, $\|A\| = 1$, because equality holds in (4.27) for the sequence $(1, 0, 0, \ldots)$.

More generelly, let $a = (a_n)_{n\in\mathbb{N}}$ be a *bounded* (complekx) sequence of numbers and set $\|a\|_u = \sup\{|a_n| \mid n \in \mathbb{N}\}$. For $x = (x_n)_{n\in\mathbb{N}} \in \ell_2(\mathbb{N})$ we define

$$M_a x = (a_n x_n)_{n\in\mathbb{N}} \ . \tag{4.28}$$

Noting that

$$\sum_{n=1}^{\infty} |a_n x_n|^2 \leq \|a\|_u^2 \|x\|^2 \ ,$$

it follows that $M_a x \in \ell_2(\mathbb{N})$. and we may conclude that eq. (4.28) deines a mapping M_a from H to H. It is easily seen that M_a is linear and the previous inequality then shows that $\|M_a\| \leq \|a\|_u$. In fact, $\|M_a\| = \|a\|_u$, because $M_a e_n = a_n e_n$, where e_n denotes the n'th vector in the canonical orthonormal basis for $\ell_2(\mathbb{N})$, and as a consequence we get $\|M_a\| \geq \|M_a e_n\| = |a_n|$ for all $n \in \mathbb{N}$, which yields $\|M_a\| \geq \|a\|_u$.

Viewing $(a_n)_{n\in\mathbb{N}}$ and $x = (x_n)_{n\in\mathbb{N}}$ in (4.28) as functions on $\mathbb{N}$, the operator M_a is defined by multiplication by the funkcion $(a_n)_{n\in\mathbb{N}}$. For this reason it is called the *multiplication operator defined by the sequence* $(a_n)_{n\in\mathbb{N}}$. Note, that w.r.t. the canonical basis the operator M_a is represented by the diagonal matrix $\Delta(a_1, a_2, \dots)$ (verify this!).

b) Let $H = L_2([a, b])$, where $[a, b]$ is a closed bounded interval, and let $f : [a, b] \to \mathbb{C}$ be a continuous function and set

$$M_f g = f \cdot g \ , \quad g \in H \ . \tag{4.29}$$

Since f is continuous, it is bounded. We define

$$\|f\|_u = \max\{|f(x)| \mid x \in [a, b]\}$$

and call $\|f\|_u$ the *uniform norm of* f. For $g \in H$ we then have

$$\int_a^b |f(x)g(x)|^2 dx \leq \|f\|_u^2 \int_a^b |g(x)|^2 dx; , \tag{4.30}$$

from which we conclude that $f \cdot g \in H$, such that (4.29) defines a mapping M_f from H into H. Clearly this is a linear mapping and (4.30) then shows that $M_f \in \mathcal{B}(H)$ and that $\|M_f\| \leq \|f\|_u$. In fact, we have $\|M_f\| = \|f\|_u$ (see Exercise 7). The operator M_f is called the *multiplication operator defined by the function* f.

c) Let $H = \ell_2(\mathbb{N})$ and set

$$T(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, x_3, \ldots)$$

for $(x_1, x_2, x_3, \ldots) \in \ell_2(\mathbb{N})$. It is evident that $T(x_1, x_2, x_3, \ldots) \in \ell_2(\mathbb{N})$ and that

$$\|T(x_1, x_2, x_3, \ldots)\| = \|(0, x_1, x_2, x_3, \ldots)\| \ .$$

Furthermore, it is easily verified that the mapping $T : H \to H$ so defined is linear. Hence, T is a bounded operator on H with norm 1, in fact an isometry. T is sometimes called the *right shift operator* on $\ell_2(\mathbb{N})$.

d) Let $H = L_2([-\pi, \pi])$ with the standard inner product normalised by the factor $\frac{1}{2\pi}$ as in section 3.6, and let $(e_n)_{n\in\mathbb{Z}}$ denote the orthonormal basis, where $e_n(\theta) = e^{in\theta}$. Setting $D = -i\frac{d}{d\theta}$, we have

$$De_n = ne_n \ , \tag{4.31}$$

and D acts, of course, linearly on the subspace $\operatorname{span}\{e_n \mid n \in \mathbb{Z}\}$, whose closure is $L_2([-\pi, \pi])$. However, D is not a bounded operator on this space, since $\|D\| \geq \|De_n\| = |n|$ for all $n \in \mathbb{Z}$. As a consequence, D cannot be extended to a bounded operator on the whole space H. This is a general feature of so-called differential operators. As will be discussed later, D has an extension to a self-adjoint operator (see section 4.4), which is of great importance in quantum mechanics.

e) Let again $H = L_2(I)$ where $I = [a, b]$ is a closed bounded interval and let $\varphi : I \times I \to \mathbb{C}$ be continuous. For $x \in I$ and $f \in H$ we define

$$(\phi f)(x) = \int_I \varphi(x, y)f(y)dy \ . \tag{4.32}$$

In order to see that $(\phi f)(x)$ is well defined by this formula we note that the continuous function $\varphi_x : y \to \varphi(x, y)$ belongs to $L_2(I)$ for every $x \in I$, and that the righhand side of (4.32) is the inner product of φ_x and $\overline{f}$ in H. Applying Cauchy-Schwarz we get

$$|(\phi f)(x)|^2 \leq \int_I |\varphi(x, y)|^2 dy \cdot \|f\|^2 \ , \tag{4.33}$$

which implies that (4.32) defines a bounded funktion ϕf on I.

We now show that ϕf is continuous. We may assume $f \neq \underline{o}$ since otherwise $(\phi f)(x) = 0 \ , x \in I$, which is continuous. First we note that f integrable, since

$$\int_a^b |f(x)|dx \leq (b - a)^{\frac{1}{2}} \left(\int_a^b |f(x)|^2 dx \right)^{\frac{1}{2}}$$

by Cauchy-Schwarz. We set $K = \int_I |f(x)|dx$. Let now $\epsilon > 0$. Since φ is uniformly continuous there exists a $\delta > 0$, such that $|\varphi(x, y) - \varphi(x', y')| < \frac{\epsilon}{K}$ for all $(x, y), (x', y') \in I \times I$ fulfilling $|x - x'|, |y - y'| < \delta$. In particular, $|\varphi(x, y) - \varphi(x', y)| < \frac{\epsilon}{K}$ for all $y \in I$ and all $x, x' \in I$, such that $|x - x'| < \delta$. From this we infer

$$|(\phi f)(x) - (\phi f)(x')| \leq \int_I |(\varphi(x, y) - \varphi(x', y))f(y)|dy \leq \frac{\epsilon}{K} \int_I |f(x)|dx = \epsilon \ ,$$

when $|x - x'| < \delta$. This shows that, ϕf is continuous.

It now follows from (4.33), that $\phi f \in H$, and that

$$\|\phi f\|^2 \leq \int_I \int_I |\varphi(x, y)|^2 dy dx \cdot \|f\|^2 \ . \tag{4.34}$$

Since ϕf clearly depends linearly on f, we have shown that (4.32) defines a bounded operator ϕ on $L_2(I)$, whose norm fulfills

$$\|\phi\| \leq \Big(\int_I \int_I |\varphi(x,y)|^2 dy dx \Big)^{\frac{1}{2}} .$$

In general, equality does not hold here.

The operator ϕ is called an *integral operator* and the function φ is called its *kernel*. Operators of this kind play an important role in the theory of differential equations.

We now aim at defining, for an arbitrary operator $A \in \mathcal{B}(H_1, H_2)$, the adjoint operator $A^* \in \mathcal{B}(H_2, H_1)$. It is not convenient to make use of a matrix representation of A, since it is not clear in the infinite dimensional case that the Hermitean conjugate matrix represents a bounded every where defined operator. Instead, we take (4.6) as the defining equation for A^* (se Theorem 4.5 below). We need, however, some preparation first.

The *dual space* to the Hilbert space H, which is denoted by H^*, is defined as

$$H^* = \mathcal{B}(H, \mathbb{L}) .$$

The elements of H^* are thus continuous linear functions from H to $\mathbb{L}$. These are also called *continuous linear forms* on H. Let us define, for $y \in H$, the function $\ell_y : H \to \mathbb{L}$ by

$$\ell_y(x) = (x, y) , \tag{4.35}$$

where $(\cdot\ ,\ \cdot)$ is the inner product on H. Then ℓ_y is linear. By Cauchy-Schwarz

$$|\ell_y(x)| \leq \|x\|\ \|y\| , \quad y \in H ,$$

from which it follows that $\ell_y \in H^*$ and that

$$\|\ell_y\| \leq \|y\| . \tag{4.36}$$

The *Riesz representation theorem* that follows tells us that all continuous linear forms on H are of the form (4.35).

Theorem 4.4 *Let H be a Hilbert space over $\mathbb{L}$. The mapping $y \to \ell_y$, where $\ell_y \in H^*$ is given by (4.35), is a bijective isometry from H onto H^*, that is for every $\ell \in H^*$ there exists exactly one vector $y \in H$, such that $\ell = \ell_y$, and we have*

$$\|\ell_y\| = \|y\| . \tag{4.37}$$

Proof. We have already seen that $y \to \ell_y$ is a well defined mapping from H into H^*. By (4.35)

$$\ell_{y+z} = \ell_y + \ell_z \ \text{ og } \ \ell_{\lambda y} = \overline{\lambda} \ell_y \tag{4.38}$$

for $y, z \in H$ and $\lambda \in \mathbb{L}$, i.e. the mapping $y \to \ell_y$ is conjugate linear.

That it is isometric can be seen as follows. If $y = 0$ then $\ell_y = 0$, so $\|\ell_y\| = \|y\| = 0$. For $y \neq 0$ we set $x = \|y\|^{-1} y$, such that $\|x\| = 1$. Since

$$|\ell_y(x)| = |(\|y\|^{-1} y, y)| = \|y\|$$

it follows that $\|\ell_y\| \geq \|y\|$. Together with (4.36) this proves (4.37). That (4.37) is equivalent to the isometry of $y \to \ell_y$ now follows from (4.38), since

$$\|\ell_y - \ell_z\| = \|\ell_{y-z}\| = \|y - z\|$$

for $y, z \in H$.

Knowing that the mapping $y \to \ell_y$ is isometric and conjugate linear, it is obviously also injective. It remains to show that it is surjective. For this purpose let $\ell \in H^*$. We wish to find $y \in H$, such that $\ell = \ell_y$. If $\ell = 0$ we can evidently choose $y = 0$. Suppose therefore $\ell \neq 0$. Then $X = \ell^{-1}(\{\underline{0}\}) \neq H$, and since $\{\underline{0}\}$ is a closed subspace of $\mathbb{L}$ and ℓ is continuous, it follows that X is a closed subspace of H. Hence $H = X \oplus X^\perp$ by Theorem 3.15 and $X^\perp \neq \{\underline{0}\}$.

In fact $X^\perp$ is a one-dimensional subspace of H: Choose $e \in X^\perp \setminus \{\underline{0}\}$ with $\|e\| = 1$, and let $z \in X^\perp$ be arbitrary. Then $\ell(e) \neq 0$ and

$$\ell(z) = \frac{\ell(z)}{\ell(e)}\ell(e) = \ell\left(\frac{\ell(z)}{\ell(e)}e\right)$$

and hence $\ell\left(z - \frac{\ell(z)}{\ell(e)}e\right) = 0$, i.e. $z - \frac{\ell(z)}{\ell(e)}e \in X$. But since $z, e \in X^\perp$ we also have $z - \frac{\ell(z)}{\ell(e)}e \in X^\perp$. Using $X \cap X^\perp = \{\underline{0}\}$, we conclude that $z = \frac{\ell(z)}{\ell(e)}e$, which shows that e is a basis for $X^\perp$.

Every vector in H can therefore be written in the form $x + \lambda e$, where $x \in X$ and $\lambda \in \mathbb{L}$. We then have

$$\ell(x + \lambda e) = \ell(x) + \lambda\ell(e) = \lambda\ell(e) = (x + \lambda e, \overline{\ell(e)}e) \ ,$$

such that $\ell = \ell_y$ hvor $y = \overline{\ell(e)}e$. Here we have used linearity of ℓ and that $x \perp e$. □

Note that (4.37) is equivalent to

$$\|x\| = \sup\{|(x, y)| \mid \|y\| \leq 1\} \tag{4.39}$$

for $x \in H$. For later use we remark that, as a consequence, we have for $A \in \mathcal{B}(H_1, H_2)$ that

$$\|A\| = \sup\{|(Ax, y)| \mid x \in H_1, y \in H_2 \, , \, \|x\| \, , \, \|y\| \leq 1\} \ . \tag{4.40}$$

We are now ready to introduce the adjoint operator.

Theorem 4.5 *Let $A \in \mathcal{B}(H_1, H_2)$. There exists a unique operator $A^* \in \mathcal{B}(H_2, H_1)$ which fulfills*

$$(Ax, y) = (x, A^*y) \ , \quad x \in H_1 \, , \, y \in H_2 \ . \tag{4.41}$$

Moreover,

$$\|A\| = \|A^*\| \ . \tag{4.42}$$

Proof. For any given $y \in H_2$ the mapping $x \to (Ax, y)$ belongs to H_1^*, being a composition of A and ℓ_y. By Theorem 4.4 there exists a unique vector $z \in H_1$, such that

$$(Ax, y) = (x, z) \ , \quad x \in H_1 \ .$$

Since z depends only on y for the given operator A, a mappping $A^* : H_2 \to H_1$ is defined by setting $A^*y = z$. Obviously, this mapping satisfies (4.41).

That A^* is a linear mapping can be seen as follows. For given $y, z \in H_2$ we have on the one hand

$$(Ax, y+z) = (x, A^*(y+z)) \ , \quad x \in H_1 \ , \tag{4.43}$$

and on the other hand

$$\begin{aligned} (Ax, y+z) &= (Ax, y) + (Ax, z) \\ &= (x, A^*y) + (x, A^*z) \\ &= (x, A^*y + A^*z) \ , \quad x \in H_1 \ . \end{aligned} \tag{4.44}$$

Since A^* is uniquely determined by (4.41), it follows by comparing (4.43) and (4.44), that

$$A^*(y+z) = A^*y + A^*z \ .$$

Similarly,

$$A^*(\lambda y) = \lambda A^* y$$

for $\lambda \in \mathbb{L}$ and $y \in H_1$ follows from

$$(Ax, \lambda y) = (x, A^*(\lambda y))$$

and

$$(Ax, \lambda y) = \overline{(Ax, y)} = \overline{(x, A^*y)} = (x, \lambda A^* y)$$

for $x \in H_1$.

That A^* is bounded and that $\|A\| = \|A^*\|$ follows immediately from (4.40) and (4.41). □

In case $H_1 = H_2 = H$ we say that an operator $A \in \mathcal{B}(H)$ is *self-adjoint* if $A = A^*$, which by (4.41) means that

$$(Ax, y) = (x, Ay) \ , \quad x, y \in H \ .$$

Self-adjoint operators play a particularly important role in operator theory, similar to that of symmetric matrices in linear algebra, as will be further discussed in the next section.

Example 4.6 a) Let M_a be the multiplication operator on $\ell_2(\mathbb{N})$ defined in Example 4.3 a). For $x = (x_n)_{n\in\mathbb{N}}$ and $y = (y_n)_{n\in\mathbb{N}}$ we have

$$(M_a x, y) = \sum_{n=1}^{\infty} a_n x_n \overline{y_n} = \sum_{n=1}^{\infty} x_n \overline{\overline{a_n} y_n} = (x, M_{\overline{a}} y) \ ,$$

where $\overline{a} = (\overline{a_n})_{n\in\mathbb{N}}$. Hence

$$M_a^* = M_{\overline{a}} \ .$$

In particular, M_a is self-adjoint if and only if a is a real sequence.

b) Let M_f be the multiplication operator on $L_2([a,b])$ defined in Example 4.3 b). For $g,h \in L_2([a,b])$ we have

$$(M_f g, h) = \int_a^b f(x)g(x)\overline{h(x)}dx = \int_a^b g(x)\overline{\overline{f(x)}h(x)}dx = (g, M_{\overline{f}}h) ,$$

from which it follows that

$$M_f^* = M_{\overline{f}} ,$$

and, in particular, that M_f is self-adjoint if and only if f is a real function.

c) Let ϕ be the integral operator on $L_2(I)$ defined in Example 4.3 e). For $f,g \in L_2(I)$ we have

$$\begin{aligned}(\phi f, g) &= \int_I \Big(\int_I \phi(x,y)f(y)dy\Big)\overline{g(x)}dx = \int_I\int_I f(y)\phi(x,y)\overline{g(x)}dydx \\ &= \int_I\int_I f(y)\overline{\overline{\phi(x,y)}g(x)}dxdy = \int_I \Big(f(x)\overline{\int_I \overline{\phi(y,x)}g(y)dy}\Big)dx ,\end{aligned}$$

where we have interchanged the order of integrations (Fubini's theorem) at one stage. From this calculation we read off the action of ϕ^*:

$$(\phi^* g)(x) = \int_I \overline{\phi(y,x)}g(y)dy .$$

Hence ϕ^* is the integral operator on $L_2(I)$, whose kernel is $\overline{\varphi(y,x)}$.

d) Let T denote the right shift operator on $\ell_2(\mathbb{N})$ defined in Example 4.4 c). For $x = (x_n)_{n\in\mathbb{N}}$ and $y = (y_n)_{n\in\mathbb{N}}$ we have

$$(Tx, y) = x_1\overline{y_2} + x_2\overline{y_3} + x_3\overline{y_4}\ldots = (x, (y_2, y_3, y_4, \ldots)) ,$$

from which we get

$$T^* y = (y_2, y_3, y_4, \ldots) .$$

We call T^* the *left shift operator* on $\ell_2(\mathbb{N})$.

e) Consider the differential operator D defined in Example 4.4 d). By partial integration one gets for functions $f,g \in \operatorname{span}\{e_n \mid n \in \mathbb{Z}\}$ (with notation as in Example 4.4)

$$\begin{aligned}(Df, g) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} -i\frac{df}{d\theta}\overline{g(\theta)}d\theta \\ &= \Big[-if(\theta)\overline{g(\theta)}\Big]_{-\pi}^{\pi} + i\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\overline{\frac{dg}{d\theta}}d\theta = (f, Dg) ,\end{aligned}$$

where we have used that the first term after the second equality sign vanishes since f and g are periodic with period 2π. It is possible to define the adjoint for unbounded operators, such as D, in general. We shall, however, refrain from doing that at this stage. Suffice to mention here that despite the identity above, D is not self-adjoint, but it can be extended to an operator $\bar{D}$, such that $\bar{D} = \bar{D}^*$, as will be further discussed is the next section.

We note the following useful properties of adjoint operators:

(i) $(A+B)^* = A^* + B^*, \quad \text{for} \quad A, B \in \mathcal{B}(H_1, H_2)$

(ii) $(\lambda A)^* = \overline{\lambda} A^*, \quad \text{for} \quad A \in \mathcal{B}(H_1, H_2)$

(iii) $(BA)^* = A^* B^*, \quad \text{for} \quad A \in \mathcal{B}(H_1, H_2)\,, B \in \mathcal{B}(H_2, H_3)$

(iv) $A^{**} = A, \quad \text{for} \quad A \in \mathcal{B}(H_1, H_2)\,,$

where H_1, H_2 and H_3 are (separable) Hilbert spaces. Properties i) and ii) can be shown in a way similar to the linearity of A^* in the proof of Theorem 4.5 and are left for the reader. Property (iii) follows similarly by comparing

$$(BAx, y) = (x, (BA)^* y)$$

and

$$(BAx, y) = (B(Ax), y) = (Ax, B^* y) = (x, A^*(B^* y)) = (x, A^* B^* y)$$

for $x \in H_1$, $y \in H_3$.

By complex conjugation of both sides of (4.41) we obtain

$$(A^* y, x) = (y, Ax)\,, \quad y \in H_2\,, \, x \in H_1\,.$$

Comparison with (4.41) finally shows that $A^{**} = A$.

We end this section with a brief discussion of adjoints of unbounded operators. With the exception of the operator D in Example 4.3 d) we have up to now discussed operators defined everywhere on a Hilbert space. An operator A from H_1 into H_2 is called *densely defined* if its domain of definition $\mathcal{D}(A)$ is a subspace whose closure equals H_1 or, equivalently, whose orthogonal complement is $\{\underline{o}\}$. If A is densely defined and bounded, i.e. fulfills (4.2) for all $x \in \mathcal{D}(A)$, then it can be extended uniquely to a bounded operator defined everywhere on H_1, see Exercise 16. In this case we may define the adjoint of A simply as the adjoint of its extension to H_1. For unbounded densely defined operators, such as D in Example 4.3 d), this method is not applicable. Nevertheless, we can define the adjoint A^* as follows.

As the domain of A^* we take the subspace

$$\mathcal{D}(A^*) = \{y \in H_2 \mid x \to (Ax, y) \text{ is a bounded linear form on } \mathcal{D}(A)\}\,.$$

If $y \in \mathcal{D}(A^*)$ the extension result quoted above implies that the linear form $x \to (Ax, y)$ has a unique extension to a bounded linear form on H_1. Hence, by Theorem 4.4, there exists a unique vector $A^* y$ in H_1 such that

$$(AX, y) = (x, A^* y)\,, \quad x \in \mathcal{D}(A)\,. \tag{4.45}$$

This defines A^* on $\mathcal{D}(A^*)$. That $\mathcal{D}(A^*)$ is a subspace of H_2 and that A^* is linear on this subspace is shown in the same way as for bounded operators. Note that it may happen that A^* is not densely defined.

With this definition we say that a densely defined operator A from H into H is self-adjoint if $A = A^*$, that is if

$$\mathcal{D}(A^*) = \mathcal{D}(A) \quad \text{and} \quad (Ax, y) = (x, Ay) \text{ for } x, y \in \mathcal{D}(A)\,.$$

As seen in Example 4.6 e) the operator D with domain $\text{span}\{e_n \mid n \in \mathbb{Z}\}$ satisfies the latter of thee requirements, which we express by saying that D is *symmetric*. However, it does not satisfy the first requirement concerning domains. Although, D is densely defined, since $\{e_n \mid n \in \mathbb{Z}\}$ is an orthonormal basis for $L_2([-\pi, \pi])$, the domain of D^* is bigger, as will be further discussed in Example 4.19. Thus D is symmetric but not self-adjoint. It can, however, be extended to a self-adjoint operator, see Example 4.19.

4.4 Diagonalisable operators and self-adjoint operators.

In this section we discuss some basic notions and results relating to diagonalisable operators on sepabable Hilbert spaces. We use a rather straight-foreward generalisation of the finite dimensional version of the notion of a diagonalisable operator already discussed in sections 4.1 and 4.2. It is a basic result from linear algebra (see Theorem 8.24 in [M]) that every self-adjoint operator on a finite dimensional real Hilbert space is diagonalisable. The corresponding result for the infinite dimensional case (the spectral theorem) requires a further extension of the notion of a diagonalisable operator then we offer in these notes. In technical terms we focus attention on operators with discrete spectrum (consisting of the eigenvalues of the operator), although examples of operators with continuous spectrum will also occur.

Definition 4.7 Let $A \in \mathcal{B}(H)$. A scalar $\lambda \in \mathbb{L}$ is called an *eigenvalue* of A, if there exists a vector $x \neq \underline{0}$ in H, such that

$$Ax = \lambda x\,. \tag{4.46}$$

A vector $x \neq \underline{0}$ fulfilling (4.46) is called an *eigenvector* for A corresponding to the eigenvalue λ, and the subspace of all vectors fulfilling (4.46), (that is the set of eigenvectors and the null-vector) is called the *eigenspace* corresponding to λ and is denoted by $E_\lambda(A)$.

The *kernel* or *null-space* $N(A)$ for an operator $A \in \mathcal{B}(H_1, H_2)$ is defined as

$$N(A) = \{x \in H_1 \mid Ax = \underline{0}\}\,.$$

By continuity of A it follows that $N(A) = A^{-1}(\{\underline{0}\})$ is a closed subspace of H_1, since $\{\underline{0}\}$ is a closed subspace of H_2. This also follows from

$$N(A) = (A^*(H_2))^\perp\,,$$

which is a direct consequence of the definitionen of A^*.

In particular, it follows that $\lambda \in \mathbb{L}$ is an eigenvalue of $A \in \mathcal{B}(H)$ if and only if

$$N(A - \lambda) \neq \{\underline{0}\}\,,$$

and the eigenspace $E_\lambda(A) = N(A - \lambda)$ is a closed subspace of H.

Definition 4.8 An operator $A \in \mathcal{B}(H)$ is called (unitarily) *diagonalisable*, if there exists an orthonormal basis $(f_i)_{i\in\mathbb{N}}$ for H consisting of eigenvectors for A. Equivalently, this means that A is represented by a diagonal matrix w.r.t. $(f_i)_{i\in\mathbb{N}}$ (which can be seen as in the previous section). The basis $(f_i)_{i\in\mathbb{N}}$ is then said to *diagonalise* A.

The following characterisation of bounded diagonalisable operators is useful.

Theorem 4.9 *An operator $A \in \mathcal{B}(H)$ is diagonalisable if and only if there exists an orthonormal basis $(f_i)_{i\in\mathbb{N}}$ for H and a bounded sequence $(\lambda_i)_{i\in\mathbb{N}}$ in $\mathbb{L}$, such that*

$$Ax = \sum_{i=1}^{\infty} \lambda_i (x, f_i) f_i \,, \quad x \in H \,. \tag{4.47}$$

In that case $\lambda_1, \lambda_2, \lambda_3, \ldots$ are exactly the eigenvalues of A (possibly with repetitions) and the eigenspace corresponding to a given eigenvalue $\lambda \in \mathbb{L}$ is given by

$$E_\lambda(A) = \overline{span\{f_i \mid i \in \mathbb{N}\,, \lambda_i = \lambda\}} \,. \tag{4.48}$$

Moreover,

$$\|A\| = \sup\{|\lambda_i| \mid i \in \mathbb{N}\} \,. \tag{4.49}$$

Proof. Suppose $A \in \mathcal{B}(H)$ is diagonalisable and let $(f_i)_{i\in\mathbb{N}}$ be an orthonormal basis consisting of eigenvectors for A, such that $Af_i = \lambda_i f_i, \quad i \in \mathbb{N}$. For $x = \sum\limits_{i=1}^{\infty}(x, f_i)f_i \in H$ we then have (see (4.25))

$$Ax = \sum_{i=1}^{\infty} (x, f_i) Af_i = \sum_{i=1}^{\infty} \lambda_i (x, f_i) f_i$$

as desired. Since

$$|\lambda_i| = \|\lambda_i f_i\| = \|Af_i\| \leq \|A\| \, \|f_i\| = \|A\|$$

for $i \in \mathbb{N}$, we also have

$$\sup\{|\lambda_i| \mid i \in \mathbb{N}\} \leq \|A\| \,. \tag{4.50}$$

Assume, conversely, that A is given by (4.47), where $(f_i)_{i\in\mathbb{N}}$ is an orthonormal basis, and the sequence $(\lambda_i)_{i\in\mathbb{N}}$ is bounded, such that

$$\sup\{|\lambda_i| \mid i \in \mathbb{N}\} \equiv M < +\infty \,.$$

Then

$$\|Ax\|^2 = \sum_{i=1}^{\infty} |\lambda_i (x, f_i)|^2 \leq M^2 \|x\|^2 \,,$$

which shows that

$$\|A\| \leq M = \sup\{|\lambda_i| \mid i \in \mathbb{N}\} \,. \tag{4.51}$$

Inserting $x = f_i$ into (4.47) it is seen that

$$Af_i = \lambda_i f_i \,.$$

Hence $\lambda_1, \lambda_2, \dots$ are eigenvalues of A.

That there are no other eigenvalues is seen as follows. Suppose

$$Ax = \lambda x$$

where $\lambda \in \mathbb{L}$ and $x \in H \setminus \{\underline{0}\}$. Then

$$\sum_{i=1}^{\infty} \lambda_i (x, f_i) f_i = \lambda x = \sum_{i=1}^{\infty} \lambda (x, f_i) f_i \ ,$$

which impies

$$\lambda_i (x, f_i) = \lambda (x, f_i) \ , \qquad i \in \mathbb{N} \ .$$

Since $x \neq \underline{0}$ there exists $i_0 \in \mathbb{N}$, such that $(x, f_{i_0}) \neq 0$, and this implies $\lambda = \lambda_{i_0}$ and that $(x, f_i) = 0$ for $\lambda_i \neq \lambda_{i_0}$. This proves (4.48).

Finally, (4.49) follows from (4.50) and (4.51). □

Example 4.10 Let H_0 be a closed subspace of H and let P denote the orthogonal projection onto H_0, that is $Px \in H_0$ is determined by $x - Px = (1 - P)x \in H_0^{\perp}$ for $x \in H$.

Let $(f_i)_{i \in I}$ be an orthonormal basis for H_0 and let $(f_j)_{j \in J}$ be an orthonormal basis for $H_0^{\perp}$. Then the two orthonormal bases together form an orthonormal basis $(f_i)_{i \in I \cup J}$ for H (and since H is separable we can assume that $I \cup J = \mathbb{N}$). We have

$$Px = \sum_{i \in I} (x, f_i) f_i \ , \qquad x \in H \ , \tag{4.52}$$

cf. the remark after Theorem 3.15.

This shhows that P has the form (4.47) with $\lambda_i = 1$ for $i \in I$ and $\lambda_i = 0$ for $i \in J$. in particular, the only eigenvalues of P are 1 and 0, provided $I \neq \emptyset$ and $J \neq \emptyset$, and the corresonding eigenspaces are H_0 and $H_0^{\perp}$, respectively. (For $I = \emptyset$ we have $P = 0$ whereas $P = 1$ for $J = \emptyset$).

Considering again the diagonalisable operator A given by (4.47) we have, in particular, according to Example 4.10 that the orthogonal projektion P_i onto the finite dimensional subspace $H_i = \text{span}\{f_i\}$ is given by

$$P_i x = (x, f_i) f_i \ , \qquad x \in H \ , \tag{4.53}$$

such that (4.47) can be written in the form

$$Ax = \sum_{i=1}^{\infty} \lambda_i P_i x \ , \qquad x \in H \ , \tag{4.54}$$

which is expressed by writing

$$A = \sum_{i=1}^{\infty} \lambda_i P_i \ . \tag{4.55}$$

It is worth noting that the series (4.55) does generally not converge w.r.t. the norm on $\mathcal{B}(H)$, i.e. the operator norm. In fact, this is the case only if $\lambda_i \to 0$ for $i \to \infty$. Otherwise, (4.55) only makes sense when interpreted as (4.54).

The operator A given by (4.55) is represented by the diagonal matrix $\Delta(\lambda_1, \lambda_2, \dots)$ w.r.t. $(f_i)_{i\in\mathbb{N}}$. Hence the adjoint operator A^* is represented by the diagonal matrix $\Delta(\overline{\lambda}_1, \overline{\lambda}_2, \dots)$ w.r.t. $(f_i)_{i\in\mathbb{N}}$. That is

$$A^* = \sum_{i=1}^{\infty} \overline{\lambda}_i P_i \; . \tag{4.56}$$

In particular, A is self-adjoint if and only if its eigenvalues are real. This is e.g. the case for ortogonal projektions, cf. Example 4.10.

That the eigenvalues of a self-adjoint operator are real, holds generally (and not only for diagonalisable operators). Likewise, the eigenspaces corresponding to different eigenvalues are orthogonal, which for diagonalisable operators is evident from (4.48) in Theorem 4.9. These facts are contained in the following lemma.

Lemma 4.11 *Let $A \in \mathcal{B}(H)$ be self-adjoint. Then*

a) Every eigenvalue of A is real.

b) If λ_1 and λ_2 are two different eigenvalues of A, then the two corresponding eigenspaces are orthogonal.

c) If H_0 is a subspace of H such that $A\, H_0 \subseteq H_0$, then $A(H_0^\perp) \subseteq H_0^\perp$.

Proof. a) Assume $Ax = \lambda x$, $x \in H \setminus \{\underline{0}\}$. Then

$$\lambda \|x\|^2 = (\lambda x, x) = (Ax, x) = (x, Ax) = (x, \lambda x) = \overline{\lambda} \|x\|^2 \; ,$$

which implies that $\lambda = \overline{\lambda}$, since $\|x\| \neq 0$.

b) Assume $Ax_1 = \lambda_1 x_1$ and $Ax_2 = \lambda_2 x_2$. Since $\lambda_1, \lambda_2 \in \mathbb{R}$ by a) we get

$$\begin{aligned}(\lambda_1 - \lambda_2)(x_1, x_2) &= (\lambda_1 x_1, x_2) - (x_1, \lambda_2 x_2) \\ &= (Ax_1, x_2) - (x_1, Ax_2) \\ &= (x_1, Ax_2) - (x_1, Ax_2) = 0 \; ,\end{aligned}$$

which gives $(x_1, x_2) = 0$, since $\lambda_1 - \lambda_2 \neq 0$. Hence $x_1 \perp x_2$.

c) Let $x \in H_0^\perp$ and $y \in H_0$. Then

$$(Ax, y) = (x, Ay) = 0$$

since $Ay \in H_0$. This being true for arbitrary $y \in H_0$ we conclude that $Ax \in H_0^\perp$, and the claim is proven. □

A subspace H_0 as in Lemma 4.11 c) is called an *invariant subspace* for A. Thus we have seen that the orthogonal complement to an invariant subspace for a self-adjoint operator $A \in \mathcal{B}(H)$ is also an invariant subspace for A. If H_0 is closed such that $H = H_0 \oplus H_0^\perp$, this means that A can e split into two operators $A_1 \in \mathcal{B}(H_0)$ and $A_2 \in \mathcal{B}(H_0^\perp)$ in the sense that

$$Ax = A_1 x_1 + A_2 x_2$$

for $x = x_1 + x_2$, $x_1 \in H_0$, $x_2 \in H_0^\perp$. We then write

$$A = A_1 \oplus A_2 \ .$$

These remarks suffice to prove the finite dimensional versions of the spectral theorem.

Theorem 4.12 *Every self-adjoint operator $A : H \to H$ on a finite dimensional complex Hilbert space H is diagonalisable.*

Proof. Choose an arbitrary orthonormal basis α for H and let $\underline{\underline{A}}$ be the matrix representing A w.r.t. α. By the fundamental theorem of algebra the characteristic polynomial

$$p(\lambda) = \det(\underline{\underline{A}} - \lambda I)$$

has a root, and hence A has an eigenvalue. Call it λ_1 and let e_1 be a corresponding normalised eigenvector. Setting $H_0 = \mathrm{span}\{e_1\}$ we clearly have that H_0 is an invariant subspace for A. Writing accordingly $A = A_1 \oplus A_2$ as above, it is clear that A_2 is a self-adjoint operator on the subspace $H_0^\perp$ of one dimension lower than H. Hence the proof can be completed by induction. □

The assumption that H is a complex vector space is essential in the argument above, since the fundamental theorem of calculus only garantees the existence of complex roots. However, by Lemma 4.11 a) the roots are real for any symmetric real matrix and the argument carries through also in the real case. For completeness, we state the result in the following theorem which is a restatement of Theorem 8.24 in [M].

Theorem 4.13 *Every self-adjoint operator $A; H \to H$ on a finite dimensional real Hilbert space H is diagonalisable.*

Theorem 4.12 is easily generalisable to so-called *normal operators*, i.e. operators commuting with their adjoint.

Theorem 4.14 *Suppose $A : H \to H$ is an operator on a finite dimensional complex Hilbert space such that*

$$AA^* = A^*A$$

Then A is diagonalisable.

Proof. Write

$$A = U + iV \quad \text{where} \quad U = \frac{1}{2}(A + A^*) \quad \text{and} \quad V = \frac{1}{2i}(A - A^*) \, .$$

Then U and V are self-adjoint and they commute,

$$UV = VU \ ,$$

since A and A^* commute. By Theorem 4.12, U is diagonalizable, and we can write

$$U = \lambda_1 P_1 + \cdots + \lambda_k P_k \ ,$$

where $\lambda_1, \ldots, \lambda_k$ are the different eigenvalues of U and P_i is the orthogonal projection onto $E_{\lambda_i}(U)$. Since V commutes with U it follows that each eigenspace $E_{\lambda_i}(U)$ is invariant under V (see Exercise 8), and evidently V is self-adjoint when restricted to this subspace. By Theorem 4.12 there exists an orthonormal basis of eigenvectors for V for each of the eigenspaces $E_{\lambda_i}(U)$. Together these bases form an orthonormal basis consisting of eigenvectors for both U and V and hence they are also eigenvectors for A. □

Since unitary operators clearly are normal we get the following

Corollary 4.15 *Any unitary operator on a finite dimensional complex Hilbert space is diagonalisable.*

It is worth noting that the corresponding result does not hold for orthogonal operators on a real Hilbert space. F.ex. a rotation in the plane through an angle different from 0 and π has no eigenvalues at all (see also Example 4.2 b)).

We next quote a result without proof about diagonalisability of so-called Hilbert-Schmidt operators which turns out to by quite usefull in various contexts.

Definition 4.16 An operator $A : H \to H$ on a Hilbert space H is said to be a Hilbert-Schmidt operator if

$$\sum_{i \in \mathbb{N}} \|Ae_i\|^2 < \infty$$

for any orthonormal basis $(e_i)_{i\in\mathbb{N}}$ for H.

It can be shown that if the condition holds for one orthonormal basis then it holds for all orthonormal bases. Important examples of Hilbert-Schmidt operators are provided by the integral opertors as defined in Example 4.3 e) for which one can show that

$$\sum_{i \in \mathbb{N}} \|Ae_i\|^2 = \int_a^b \int_a^b |\varphi(x,y)|^2 dx dy \, .$$

We then have the following result.

Theorem 4.17 *Every self-adjoint Hilbert-Schmidt operator $A : H \to H$ on a separable Hilbert space H is diagonalisable.*

The difficult step in the proof of this result is to establish the existence of an eigenvalue. Having done so, the proof can be completed in a similar way as in the finite dimensional case. Details of the proof can be found in e.g. B.Durhuus: Hilbert rum med anvendelser, Lecture notes 1997.

We end this section by two examples of which the first contains some further discussion of multiplication operators providing prototypes of self-adjoint operators that cannot be diagonalised in the sence of Definition 4.8, and the second one discusses self-adjointness of the unbounded differential operator of Example 4.3 d).

Example 4.18 Let $H = L_2([0,3])$, and let $f : [0,3] \to \mathbb{R}$ denote the function

$$f(x) = x\,, \qquad x \in [0,3]\,.$$

Since f is real, M_f is a self-adjoint operator by Example 4.6 b).

M_f has no eigenvalues, which is seen as follows. Assume $M_f g = \lambda g$ for some $\lambda \in \mathbb{R}$ and some function $g \in H$. This means that $(f - \lambda)g = 0$ almost everywhere. But since $f - \lambda$ is $\neq 0$ almost everywhere, it follows that $g = 0$ almost everywhere, that is $g = \underline{0} \in H$, which shows that λ is not an eigenvalue.

In particular, it follows that M_f is not diagonalisable in the sense of Definition 4.8.

Let us next consider the multiplication operator M_{f_1} defined by the function $f_1 : [0,3] \to \mathbb{R}$, where

$$f_1(x) = \begin{cases} x & \text{for } 0 \le x \le 1 \\ 1 & \text{for } 1 \le x \le 2 \\ x - 1 & \text{for } 2 \le x \le 3\,. \end{cases}$$

Note that f_1 is constant equal to 1 on the interval $[1,2]$ and strictly increasing outside this interval. Clearly, M_{f_1} is self-adjoint. We show that 1 is the only eigenvalue of M_{f_1}. Indeed, if $\lambda \neq 1$ we have $f_1 - \lambda \neq 0$ almost everywhere in $[0,3]$, and it follows as above that λ is not an eigenvaluee. On the other hand, $f_1 - 1 = 0$ on $[1,2]$ and $\neq 0$ outside $[1,2]$. It follows that $M_{f_1} g = g$, if and only if $g = 0$ almost everywhere in $[0,3] \setminus [1,2]$. The set of functions fulfilling this requirement is an infinite dimensional closed subspace of H, that can be identified with $L_2([1,2])$, and which hence equals the eigenspace $E_1(M_{f_1})$.

Clearly, M_{f_1} is not diagonalisable in the sense of Definition 4.8.

The primary lesson to be drawn from this example is that self-adjoint operators in general are not diagonalisable in the sense of Definition 4.8. However, as mentioned previously, it is a principal result of analysis, called the spectral theorem for self-adjoint operators, that they are diagonalisable in a generalised sense. It is outside the scope of this course to formulate and even less to prove this result. Interested readers are referred to e.g. M.Reed and B.Simon: Methods of modern mathematical physics, Vol. I and II, Academic Press 1972.

Example 4.19 Consider again the operator D defined in Example 4.3 d). Keeping the same notation, it follows from eq. (4.31) that D is given on the domain span$\{e_n \mid n \in \mathbb{Z}\}$ by

$$Df = \sum_{n \in \mathbb{Z}} n(f, e_n)\, e_n\,. \tag{4.57}$$

This formula can be used to define an extension $\bar{D}$ of D by extending the domain of definition to the largest subspace of $L_2([-\pi,\pi])$ on which the righthand side is convergent, i.e. to the set

$$\mathcal{D}(\bar{D}) = \{f \in L_2([0,2\pi]) \mid \sum_{n \in \mathbb{Z}} n^2 |(f, e_n)|^2 < \infty\}\,,$$

on which $\bar{D}f$ is given by the righthand side of (4.57). Defining the adjoint $\bar{D}^*$ as at the end of section 4.3 it is then not difficult to show that $\bar{D}$ is a self-adjoint operator, see Exercise 17.

4.5 The Fourier transformation as a unitary operator.

Some of the fundamental properties of the Fourier transformation, which we define below, are conveniently formulated in operator language as will briefly be explained in this section.

We first generalise the definition of a unitary operator, given in section 4.2 in the finite dimensional case.

Definition 4.20 An operator $U \in \mathcal{B}(H_1, H_2)$ is unitary if it is bijective and

$$U^{-1} = U^*$$

or, equivalently, if

$$U^*U = \mathrm{id}_{H_1} \quad \text{and} \quad UU^* = \mathrm{id}_{H_2}\,.$$

Note that, contrary to the finite dimensional case, the relation $U^*U = \mathrm{id}_{H_1}$ does not imply the relation $UU^* = \mathrm{id}_{H_2}$, or vice versa, if H_1 and H_2 are of infinite dimension, see Exercise 14. As a consequence, Theorem 4.1 needs to be modified as follows in order to cover the infinite dimensional case.

Theorem 4.21 *Let $U : H_1 \to H_2$ be a bounded linear operator. The following four statements are equivalent.*

a) U is unitary.

b) U is surjective and preserves the inner product.

c) U is surjective and isometric.

d) U maps an orthonormal basis for H_1 into an orthonormal basis for H_2.

The proof is identical to that of Theorem 4.1 with minor modifications and is left to the reader.

Example 4.22 Let H be a Hilbert space with orthonormal basis $(e_i)_{i\in\mathbb{N}}$. The mapping $C : H \to \ell_2(\mathbb{N})$ which maps the vector x to its coordinate sequence w.r.t. $(e_i)_{i\in\mathbb{N}}$, that is

$$Cx = ((x, e_i))_{i\in\mathbb{N}}$$

is unitary. Indeed, C is an isometry by Parseval's equality (3.27), and C is surjective because, if $a = (a_i)_{i\in\mathbb{N}} \in \ell_2(\mathbb{N})$ then $Cx = a$, where $x = \sum_{i=1}^{\infty} a_i e_i$.

Thus every separable Hilbert space can be mapped onto $\ell_2(\mathbb{N})$ by a unitary operator. We say that every separable Hilbert space is unitarily isomorphic to $\ell_2(\mathbb{N})$

As a further, highly non-trivial, example of a unitary operator we discuss next the Fourier transformation. For detailed proofs of the statements below the reader is referred to e.g. M.Reed and B. Simon: Methods of modern mathematical physics, Vol. II, Academic press, 1972.

If f is an integrable function on $\mathbb{R}$, its *Fourier transform* $\hat{f}$ is the function on $\mathbb{R}$ defined by

$$\hat{f}(p) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ipx}\, dx\,, \quad p \in \mathbb{R}\,. \tag{4.58}$$

The first fundamental result about the Fourier transform is that, if f is a C^∞-function vanishing outside some bounded interval, then $\hat{f}$ is a C^∞-function that is both integrable and square integrable, and

$$\int_{\mathbb{R}} |f(x)|^2\, dx = \int_{\mathbb{R}} |\hat{f}(p)|^2\, dp\,. \tag{4.59}$$

It is clear that the set $C_0^\infty(\mathbb{R})$ of C^∞-functions that vanish outside some bounded interval is a subspace of $L_2(\mathbb{R})$ and that the mapping $f \to \hat{f}$ is linear on this subspace. As mentioned in Example 3.9 3), the closure of $C_0^\infty(\mathbb{R})$ equals $L_2(\mathbb{R})$. Hence, by the extension result of Exercise 16, the mapping $f \to \hat{f}$ has a unique extension to an operator $\mathcal{F} : L_2(\mathbb{R}) \to L_2(\mathbb{R})$, which moreover is isometric by (4.59).

The second fundamental result about the Fourier transform is the *inversion theorem*, which states that for $f \in C_0^\infty(\mathbb{R})$ we have

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(p) e^{ipx}\, dp\,, \quad x \in \mathbb{R}\,. \tag{4.60}$$

In operator language this equation can also be written as

$$\mathcal{F} \circ \bar{\mathcal{F}}(f) = f\,, \tag{4.61}$$

where $\bar{\mathcal{F}}$ is defined by the same procedure as $\mathcal{F}$ except that p is replaced by $-p$ in (4.58), i.e.

$$\bar{\mathcal{F}}(f)(p) = \mathcal{F}(f)(-p)\,.$$

By continuity of $\mathcal{F}$ and $\bar{\mathcal{F}}$ eq. (4.61) holds for all $f \in L_2(\mathbb{R})$. In particular, $\mathcal{F}$ is surjective such that, by Theorem 4.20, we have shown that $\mathcal{F}$ is a unitary operator on $L_2(\mathbb{R})$, whose adjoint is $\bar{\mathcal{F}}$.

We collect our findings in the following

Theorem 4.23 *The Fourier transformation $\mathcal{F}$ defined for $f \in C_0^\infty(\mathbb{R})$ by*

$$\mathcal{F}(f)(p) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ipx}\, dx\,, \quad p \in \mathbb{R}\,,$$

extends uniquely to a unitary operator on $L_2(\mathbb{R})$, whose adjoint is given by

$$\mathcal{F}^*(f)(p) = \mathcal{F}(f)(-p)\,, \quad p \in \mathbb{R}\,.$$

4.6 A note on Dirac notation.

In physics litterature on quantum mechanics the inner product is traditionally denoted by $\langle\cdot|\cdot\rangle$ instead of $(\cdot,\cdot)$. Moreover, it is assumed to be linear in the second variable rather than in the first variable, cf. Definition 3.1 iii) and iv). This convention, however, has no influence beyond notational on the general results derived above and is closely linked to the notation introduced by Dirac in his classic work P.A.M. Dirac: The Principles of Quantum Mechanics, Oxford, 1930. Its essentials are the following:

A vector x in the Hilbert space H is called a *ket*-vector and denoted by $|x\rangle$. The linear form ℓ_y corresponding to the vector $y \in H$ is called a *bra*-vector and is denoted by $\langle y|$. By Theorem 4.4, the mapping

$$|x\rangle \to \langle x|$$

from H to H^* is bijective. With this notation, $\ell_y(x)$ takes the form

$$\langle y|(|x\rangle) = \langle y|x\rangle\,,$$

which explains the names of bra and ket making up a bracket $\langle\cdot|\cdot\rangle$.

Given a ket-vector $|x\rangle$ and a bra-vector $\langle y|$, the linear operator on H defined by

$$|z\rangle \to \langle y|z\rangle|x\rangle$$

is naturally denoted by $|x\rangle\langle y|$. In particular, if $|e\rangle$ is a unit vector in H, the orthogonal projection onto the subspace spanned by $|e\rangle$ is $|e\rangle\langle e|$. More generally, if A is a diagonalisable operator it has the form (4.47), which in Dirac notation is written as

$$A = \sum_{i=1}^{\infty} \lambda_i |f_i\rangle\langle f_i|\,,$$

interpreted in a way similar to (4.55). In particular, for $A = 1$, the identity operator on H, we get

$$1 = \sum_{i=1}^{\infty} f_i\rangle\langle f_i|\,,$$

where $(f_i)_{i\in\mathbb{N}}$ is an arbitrary orthonormal basis for H. When applied to $|x\rangle$, this yields the the orthonormal expansion of $|x\rangle$ in Dirac notation

$$|x\rangle = \sum_{i=1}^{\infty} |f_i\rangle\langle f_i|x\rangle\,.$$

Exercises

Exercise 1 Let H be a 3-dimensional complexn Hilbert space and let $A : H \to H$ be the operator represented by the matrix

$$\begin{pmatrix} 2 & 3+i & 2i \\ 0 & 2 & 0 \\ i & 1-i & 2 \end{pmatrix}$$

w.r.t. an orthonormal basis $\alpha = (e_1, e_2, e_3)$. Find the matrix, that represents A^* w.r.t. α, and determine $A^*(e_1 + ie_2 - e_3)$.

Exercise 2 Let A be the linear operator on $\mathbb{C}^3$ represented by the matrix

$$\begin{pmatrix} 2 & 0 & -i \\ 0 & 2 & 0 \\ i & 0 & 2 \end{pmatrix}$$

w.r.t. the canonical basis for $\mathbb{C}^3$. Show that A is self-adjoint and find its eigenvalues as well as an orthonormal basis for $\mathbb{C}^3$, which diagonalises A.

Exercise 3 Let H and α be as in Exercise 1 and let U be the linear operator on H represented by the matrix

$$\begin{pmatrix} \frac{1}{6}+\frac{i}{2} & \frac{1}{6}-\frac{i}{2} & -\frac{2}{3} \\ \frac{1}{6}-\frac{i}{2} & \frac{1}{6}+\frac{i}{2} & -\frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{pmatrix}.$$

w.r.t. α.

Show that U is unitary.

Show that 1, -1, i are eigenvalues of U and find an orthonormal basis for H, which diagonalises U.

Exercise 4 Let H be a finite dimensional Hilbert space over $\mathbb{L}$. We define

$$GL(H) = \{A \in \mathcal{B}(H) \mid A \text{ er invertible, i.e. } A \text{ is bijectiv}\}$$

and

$$O(H) = \{O \in \mathcal{B}(H) \mid O \text{ is orthogonal}\}\text{, if } \mathbb{L} = \mathbb{R},$$

and

$$U(H) = \{U \in \mathcal{B}(H) \mid U \text{ is unitary}\}\text{, if } \mathbb{L} = \mathbb{C}.$$

Show that

1) If $A, B \in GL(H)$ then $AB \in GL(H)$ and $A^{-1} \in GL(H)$,

2) $O(H) \subseteq GL(H)$ and $U(H) \subseteq GL(H)$,

3) If $A, B \in O(H)$ then $AB \in O(H)$ and $A^{-1} \in O(H)$, and similarly if $O(H)$ is replaced by $U(H)$.

Show likewise the corresponding statements for $GL(n)$, $O(n)$ and $U(n)$, denoting the sets of invertible, orthogonal and unitary $n \times n$-matrices, respectively.

$GL(H)$, $O(H)$ and $U(H)$ are called the *general linear*, the *orthogonal* and the *unitary group* over H, respectively.

Exercise 5 Let $A : E \to E'$ be a bounded operator between inner product spaces as defined by eq. (4.2).

Show that the norm $\|A\|$ is given by (4.3). Show also the following properties of the norm:

$$\begin{aligned} \|A\| > \ & 0 \text{ if } A \neq 0\,, \\ \|\lambda A\| &= |\lambda| \|A\|\,, \\ \|A + B\| &\leq \|A\| + \|B\|\,, \\ \|CA\| &\leq \|C\| \|A\|\,, \end{aligned}$$

where $\lambda \in \mathbb{L}$ and the operators $B : E \to E'$, $C : E' \to E''$ are bounded (E'' being an inner product space).

Exercise 6 Show that a linear operator $A : E \to E'$ between inner product spaces is bounded if and only if it is continuous.

Hint. Show first that by linearity A is continuous if and only if it is continuous at $\underline{0}$. To show that continuity implies boundedness it may be useful to argue by contradiction.

Exercise 7 Show that the norm of the multiplication operator in Example 4.3 b) is given by

$$\|M_f\| = \|f\|_u\,.$$

Exercise 8 Assume that U and V are two commuting bounded operators on a Hilbert space H. Show that each eigenspace $E_\lambda(U)$ is invariant under V, i.e. if x is an eigenvector for U with eigenvalue λ or equals $\underline{0}$, then the same holds for Vx.

Exercise 9 Let H be a Hilbert space and let f be a function given by a power seires

$$f(z) = \sum_{n=0}^{\infty} c_n z^n\,, \qquad |z| < \rho\,,$$

where $\rho > 0$ is the series' radius of convergence.

Show that for an operator $A \in \mathcal{B}(H)$ with $\|A\| < \rho$ the series

$$\sum_{n=0}^{\infty} |c_n|\, \|A^n\|$$

is convergent (in $\mathbb{R}$) and that the series

$$\sum_{n=0}^{\infty} c_n A^n x$$

is convergent in H for all $x \in H$. The sum is then denoted by $f(A)x$, that is

$$f(A)x = \sum_{n=0}^{\infty} c_n A^n x\,.$$

Verify that $f(A)$ is a bounded linear operator on H.

Assume now that H is finite dimensional and that $\alpha = (e_1, \ldots, e_n)$ is an orthonormal basis for H, and let $\underline{\underline{A}}$ be the matrix representing A w.r.t. α. Show that the series

$$\sum_{n=0}^{\infty} c_n (\underline{\underline{A}}^n)_{ij}$$

is convergent in $\mathbb{C}$ for all $1 \leq i, j \leq n$, and that the so defined matrix $f(\underline{\underline{A}})$, where

$$(f(\underline{\underline{A}}))_{ij} = \sum_{n=0}^{\infty} c_n (\underline{\underline{A}}^n)_{ij} \ ,$$

represents the operator $f(A)$ w.r.t. α.

Exercise 10 Let A and f be given as in Exercise 9.

Show that if A is diagonalisable, such that $A = \sum\limits_{i \in I} \lambda_i P_i$ in the sense of eq.(4.55), then $f(A)$ is also diagonalisable and

$$f(A) = \sum_{i \in I} f(\lambda_i) P_i \ .$$

Exercise 11 Let H be a complex Hilbert space and let $A \in \mathcal{B}(H)$. Show that

$$e^{(t+s)A} = e^{tA} e^{sA} \ , \quad s, t \in \mathbb{C} \ ,$$

and conclude that e^A is invertible, and that

$$(e^A)^{-1} = e^{-A} \ .$$

Does the identity $e^{A+B} = e^A e^B$ hold for arbitrary $A, B \in \mathcal{B}(H)$?

Exercise 12 Let H be a Hilbert space. A function $x : \mathbb{R} \to H$ with values in H is called differentiable at $t_0 \in \mathbb{R}$ if there exists a vector $a \in H$, such that

$$\|(t - t_0)^{-1}(x(t) - x(t_0)) - a\| \to 0 \ \text{ for } \ t \to t_0 \ .$$

As usual, we then denote a by $x'(t_0)$ or by $\frac{dx}{dt}(t_0)$:

$$\lim_{t \to t_0} (t - t_0)^{-1}(x(t) - x(t_0)) = x'(t_0) \ \text{ in } \ H \ .$$

Show, using notation as in Exercise 11 that the function $t \to e^{tA}x$ is differentiable on $\mathbb{R}$ for all $x \in H$ and that

$$\frac{de^{tA}x}{dt} = Ae^{tA}x \, ; \quad t \in \mathbb{R} \ . \quad (*) \ .$$

(In case A is self-adjoint the equation $(*)$ is a special case of the Schrödinger equation describing the time-evolution of a quantum mechanical system, for which the energy operator (Hamiltonian) A in suitable units is bounded and time-independent. If A is diagonalisable, its eigenvectors are called the energy eigenstates of the system.)

Exercise 13 Let σ_j, $j = 1, 2, 3$ denote the so-called Pauli matrices given by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

Verify that

$$\sigma_j^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} , \quad j = 1, 2, 3 ,$$

and use this to calculate $e^{i\theta\sigma_j}$ for $\theta \in \mathbb{C}$ and $j = 1, 2, 3$.

Exercise 14 Let H be an infinite dimensional Hilbert space and let $(e_i)_{i\in\mathbb{N}}$ be an orthonormal basis for H.

Show that there exists exactly one operator $T \in \mathcal{B}(H)$ such that

$$Te_i = e_{i+1} , \quad i \in \mathbb{N} ,$$

and that T is an isometry, that is $\|Tx\| = \|x\|$, $x \in H$.

Find T^* and show that

$$T^*T = 1 \quad \text{og} \quad TT^* \neq 1 .$$

Exercise 15 Show that if $P \in \mathcal{B}(H)$ fulfills $P^* = P$ and $P^2 = P$, then P is the orthogonal projection onto $X = P(H)$.

Exercise 16 Let A be a densely defined operator from H_1 to H_2 and assume A is bounded.

Show that A has a unique extension to a bounded operator $\bar{A}$ defined everywhere on H_1 in the following two steps.

1) Since A is densely defined there exists for every $x \in H_1$ a sequence (x_i) in $\mathcal{D}(A)$ that converges to x. Show that the sequence (Ax_i) is convergent in H_2 and that its limit only depends on x and not on the choice of sequence (x_i) converging to x.

2) With notation as in 1) set
$$\bar{A}x = \lim_{i\to\infty} Ax_i$$
and show that $\bar{A}$ is a bounded linear operator that extends A.

Exercise 17 Let H be a Hilbert space with orthonormal basis $(e_i)_{i\in\mathbb{N}}$ and let $(\lambda_i)_{i\in\mathbb{N}}$ be a sequence in $\mathbb{L}$, not necessarily bounded.

a) Show that the equation
$$Ax = \sum_{i=1}^{\infty} \lambda_i (x, e_i) e_i . \tag{4.62}$$
defines a densely defined operator with domain
$$\mathcal{D}(A) = \{x \in H \mid \sum_{i=1}^{\infty} |\lambda_i (x, e_i)|^2 < \infty\}$$

b) Show that the adjoint of the operator A in a) has the same domain of definition as A and is given by the formula

$$A^* x = \sum_{i=1}^{\infty} \overline{\lambda}_i (x, e_i) e_i \, . \tag{4.63}$$

c) Deduce from a) and b) that A is self-adjoint if and only if λ_i is real for all $i \in \mathbb{N}$.