Selected Topics in Mathematical Physics

Erwin Brüning
School of Mathematical Sciences

May 2014

Contents

Chapter 1

Sobolev spaces

Abstract: This chapter discusses classical aspects of the theory of Sobolev spaces. We begin with basic definitions. In order to prepare the important embedding results for Sobolev space we prove Morrey's inequality and the inequality of Gagliardo-Nirenberg-Sobolev. Now the proof of continuous embeddings of Sobolev space is straight forward. After recalling the Kolmogorov-Riesz compactness criterion for sets in $L^q(\mathbb{R}^n)$ spaces we proceed to prove for bounded sets $\Omega \subset \mathbb{R}^n$ compactness of the embeddings of Sobolev spaces $W^{k,p}(\Omega)$ into $L^q(\Omega)$ for a suitable range of the exponents p, q.

1.1 Motivation

As we will learn in the Introduction to Part C on variational methods, all major developments in the calculus of variations were driven by concrete problems, mainly in Physics. In these applications the underlying Banach space is a suitable function space, depending on the context as we are going to see explicitly later. Major parts of the existence theory of solutions of nonlinear partial differential use variational methods (some are treated in Chapter 32). Many other applications can be found for instance in the e-book 10. Here the function spaces which are used are often the so-called **Sobolev spaces** and the successful application of variational methods rests on various types of **embeddings** for these spaces. Accordingly we present here very briefly the classical aspects of the theory of Sobolev spaces as they are used in later applications. Some parts of our presentation will just be a brief sketch of important results; this applies in particular to the results on the approaximation of elements of a Sobolev space by smooth functions. A comprehensive treatment can for instance be found in the books 2, 1 and a short introduction in 10.

We assume that the reader is familiar with the basics aspects of the theory of Lebesgue spaces.

1.2 Basic definitions

Let $\Omega \subseteq \mathbb{R}^n$ be a nonempty openset, and for $k = 0,1,2,\ldots$ and $1 \leq p \leq \infty$ introduce the vector space

$$\mathcal{C}^{[k,p]}(\Omega) = \left\{ u \in \mathcal{C}^k(\Omega) : D^\alpha u \in L^p(\Omega), |\alpha| \leq k \right\}.$$

Here $\alpha = (\alpha_1, \ldots, \alpha_n)$ is an n-tuple of number $\alpha_i = 0,1,2,\ldots$ and $|\alpha| = \sum_{i=1} \alpha_i$, and $D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n}}$. On this vector space define a norm for $1 \leq p < \infty$ by

$$\|u\|_{k,p} = \left(\sum_{|\alpha| \leq k} \|D^\alpha u\|_p^p \right)^{1/p}. \tag{1.1}$$

and for $p = \infty$ by

$$\|f\|_{k,\infty} = \sum_{|\alpha| \leq k} \|D^\alpha f\|_{L^\infty(\Omega)}. \tag{1.2}$$

The *Sobolev space* $W^{k,p}(\Omega)$ is by definition the completion of $\mathcal{C}^{[k,p]}(\Omega)$ with respect to this norm. These Banach spaces are naturally embedded into each other according to

$$W^{k,p}(\Omega) \subset W^{k-1,p}(\Omega) \cdots \subset W^{0,p}(\Omega) = L^p(\Omega).$$

Since the Lebesgue spaces $L^p(\Omega)$ are separable for $1 \leq p < \infty$ one can show that these Sobolev spaces are separable too. For $1 < p < \infty$ the spaces $L^p(\Omega)$ are reflexive, and it follows that for $1 < p < \infty$ the Sobolev spaces $W^{k,p}(\Omega)$ are separable reflexive Banach spaces.

There is another equivalent definition of the Sobolev spaces in terms of weak (or distributional) derivatives due to Meyers and Serrin (1964) 17, 1:

$$W^{k,p}(\Omega) = \{f \in L^p(\Omega) : D^\alpha f \in L^p(\Omega) \text{ (weakly) for all } |\alpha| \leq k\}. \tag{1.3}$$

Here $D^\alpha f$ stands for the weak derivative of f, i.e. for all $\phi \in C_c^\infty(\Omega)$ one has in the sense of Schwartz distributions on Ω

$$\langle D^\alpha f, \phi \rangle = (-1)^{|\alpha|} \int f(x) D^\alpha \phi(x) \boldsymbol{D}x.$$

Theorem 1.2.1 *Equipped with the norms (1.1) respectively (1.2) the set $W^{k,p}(\Omega)$ is a Banach space. In the case $p = 2$ the space $W^{k,2}(\Omega) = H^k(\Omega)$ is actually a Hilbert space with the inner product*

$$\langle f, g \rangle_{H^k(\Omega)} = \sum_{|\alpha| \leq k} \int_\Omega \overline{D^\alpha f}(x) \cdot D^\alpha g(x) \boldsymbol{D}x. \tag{1.4}$$

The spaces $W^{k,p}(\Omega)$ are called **Sobolev spaces of order (k,p)**.

Proof. Since the space $L^p(\Omega)$ is a vector space, the set $W^{k,p}(\Omega)$ is a vector space too, as a subspace of $L^p(\Omega)$. The norm properties of $\|\cdot\|_{L^p(\Omega)}$ easily imply that $\|\cdot\|_{W^{k,p}(\Omega)}$ is also a norm. □

The **local** Sobolev spaces $W^{k,p}_{\mathrm{loc}}(\Omega)$ are obtained when in the above construction the Lebesgue space $L^p(\Omega)$ is replaced by the local Lebesgue space $L^p_{\mathrm{loc}}(\Omega)$. Elements in a Sobolev space can be approximated by smooth functions, i.e., these spaces allow **mollification**. In details one has the following result.

Theorem 1.2.2 *Let Ω be an open subset of $\mathbb{R}^n$, $k \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and $1 \leq p < \infty$. Then the following holds:*

a) For $u \in W^{k,p}_{\mathrm{loc}}(\Omega)$ there exists a sequence $u_m \in \mathcal{C}^\infty_c(\Omega)$ of $\mathcal{C}^\infty$ functions on Ω which have a compact support such that $u_m \to u$ in $W^{k,p}_{\mathrm{loc}}(\Omega)$.

b) $\mathcal{C}^\infty(\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$.

c) $\mathcal{C}^\infty_c(\mathbb{R}^n)$ is dense in $W^{k,p}(\mathbb{R}^n)$.

Proof. Here we have to refer to the literature, for instance 2, 1. □

Naturally, the space $\mathcal{C}^\infty_c(\Omega)$ is contained in $W^{k,p}(\Omega)$ for all $k = 0,1,2,\ldots$ and all $1 \leq p < \infty$. The closure of this space in $W^{k,p}(\Omega)$ is denoted by $W^{k,p}_0(\Omega)$.

In general $W_0^{k,p}(\Omega)$ is a proper subspace of $W^{k,p}(\Omega)$. For $\Omega = \mathbb{R}^n$ however equality holds.

The fact that $W_0^{k,p}(\Omega)$ is, in general, a proper subspace of $W^{k,p}(\Omega)$ plays a decisive role in the formulation of boundary value problems. Roughly one can say the following: If the boundary $\Gamma = \partial\Omega$ is sufficiently smooth, then elements $u \in W^{k,p}(\Omega)$ together with their normal derivatives of order $\leq k-1$ can be restricted to Γ. And elements in $W_0^{k,p}(\Omega)$ can then be characterized by the fact that this restriction vanishes. (There is a fairly technical theory involved here 1). A concrete example of a result of this type is the following theorem.

Theorem 1.2.3 *Let $\Omega \subset \mathbb{R}^n$ be a bounded open subset whose boundary $\Gamma = \partial\Omega$ is piecewise $\mathcal{C}^1$. Then the following holds:*

(a) every $u \in H^1(\Omega)$ has a restriction $\gamma_0 u = u|\Gamma$ to the boundary;

(b) $H_0^1(\Omega) = \ker\gamma_0 = \{u \in H^1(\Omega) : \gamma_0(u) = 0\}$.

Obviously, the Sobolev space $W^{k,p}(\Omega)$ embeds naturally into the Lebesgue space $L^p(\Omega)$. Depending on the value of the exponent p in relation to the dimension n of the underlying space $\mathbb{R}^n$ it embeds also into various other functions spaces, expressing various degrees of smoothness of elements in $W^{k,p}(\Omega)$. The following few sections present a number of (classical) estimates for elements in $W^{k,p}(\Omega)$ which then allow to prove the main results concerning **Sobolev embeddings**, i.e., embeddings of the Soboloev spaces into various other function spaces.

A simple example indicates what can be expected. Take $\psi \in C_c^\infty(\mathbb{R}^n)$ such that $\psi(x) = 1$ for all $|x| \leq 1$ and define $f(x) = |x|^q \psi(x)$ for $x \in \mathbb{R}^n$, for some $q \in \mathbb{R}$. Then $\nabla f \in L^p(\mathbb{R}^n)^n$ requires $n + (q-1)p \geq 0$, or

$$q \geq 1 - \frac{n}{p}.$$

Therefore, if $1 \leq p < n$ then $q < 0$ is allowed and thus f can have a singularity (at $x = 0$). If however $p \geq n$, then only exponents $q \geq 0$ are allowed, and then f is continuous at $x = 0$. The following estimates give a much more accurate picture. These estimates imply first that we get continuous embeddings and at a later stage we will show that for exponents $1 \leq p < n$ these embeddings are actually compact, if Ω is bounded.

We start with the case $n < p \leq +\infty$.

1.3 Morrey's inequality

Denote the unit sphere in $\mathbb{R}^n$ by S and introduce for a Borel measurable set $\Gamma \subset S$ with $\sigma(\Gamma) > 0$ ($\sigma(\Gamma)$ denotes the surface measure of Γ) the sets

$$\Gamma_{x,r} = \{x + t\omega : \omega \in \Gamma, 0 \leq t \leq r\}, \quad x \in \mathbb{R}^n, \quad r > 0.$$

$\Gamma_{x,r}$ is the set of all lines of length r from x in the direction $\omega \in \Gamma$. Note that for measurable functions f one has

$$\int_{\Gamma_{x,r}} f(y) \boldsymbol{D}y = \int_0^r \boldsymbol{D}t t^{n-1} \int_\Gamma f(x + t\omega) \boldsymbol{D}\sigma(\omega). \tag{1.5}$$

Choosing $f = 1$ we find for the Lebesgue measure of $\Gamma_{x,r}$:

$$|\Gamma_{x,r}| = r^n \sigma(\Gamma)/n. \tag{1.6}$$

Lemma 1.3.1 *If S, x, r are as above and $u \in C^1(\overline{\Gamma_{x,r}})$ then*

$$\int_{\Gamma_{x,r}} |u(y) - u(x)| \boldsymbol{D}y \leq \frac{r^n}{n} \int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|x - y|^{n-1}} \boldsymbol{D}y. \tag{1.7}$$

Proof. For $y = x + t\omega$, $0 \leq t \leq r$ and $\omega \in \Gamma$ one has

$$u(x+t\omega) - u(x) = \int_0^t \omega \cdot \nabla u(x+s\omega) \boldsymbol{D}s,$$

thus integration over Γ yields

$$\begin{aligned}
\int_\Gamma |u(x+t\omega) - u(x)| \boldsymbol{D}\sigma(\omega) &\leq \int_0^t \int_\Gamma |\nabla u(x+s\omega)| \boldsymbol{D}\sigma(\omega) \boldsymbol{D}s \\
&= \int_0^t s^{n-1} \int_\Gamma \frac{|\nabla u(x+s\omega)|}{|x+s\omega-x|^{n-1}} \boldsymbol{D}\sigma(\omega) \boldsymbol{D}s \\
&= \int_{\Gamma_{x,t}} \frac{|\nabla u(y)|}{|y-x|^{n-1}} \boldsymbol{D}y \leq \int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|y-x|^{n-1}} \boldsymbol{D}y.
\end{aligned}$$

If we multiply this inequality with t^{n-1} and integrate from 0 to r and observe Equation (1.5) we get (1.7). □

Corollary 1.3.2 *For any $n < p \leq +\infty$, any $0 < r < \infty$, any $x \in \mathbb{R}^n$ and any Borel measurable subset $\Gamma \subset S$ such that $\sigma(\Gamma) > 0$, one has, for all $u \in \mathcal{C}^1(\overline{\Gamma}_{x,r})$*

$$|u(x)| \leq C(\sigma(\Gamma), r, n, p) \, \|u\|_{W^{1,p}(\Gamma_{x,r})} \tag{1.8}$$

with

$$C(\sigma(\Gamma), r, n, p) = \frac{r^{1-n/p}}{\sigma(\Gamma)^{1/p}} \max\left\{ \frac{n^{-1/p}}{r}, \left(\frac{p-1}{p-n}\right)^{1-1/p} \right\}.$$

Proof. Clearly, $|u(x)| \leq |u(y)| + |u(x) - u(y)|$, for any $y \in \Gamma_{x,r}$; integration over $\Gamma_{x,r}$ and application of (1.7) gives

$$|\Gamma_{x,r}||u(x)| = \int_{\Gamma_{x,r}} |u(x)| \boldsymbol{D}y \leq \int_{\Gamma_{x,r}} |u(y)| \boldsymbol{D}y + \frac{r^n}{n} \int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|x-y|^{n-1}} \boldsymbol{D}y.$$

Now apply Hölder's inequality to continue this estimate by

$$\leq \|u\|_{L^p(\Gamma_{x,r})} \|1\|_{L^q(\Gamma_{x,r})} + \frac{r^n}{n} \|\nabla u\|_{L^p(\Gamma_{x,r})} \left\| \frac{1}{|x-\cdot|^{n-1}} \right\|_{L^q(\Gamma_{x,r})} \tag{1.9}$$

where q is the Hölder conjugate exponent of p, i.e., $q = \frac{p}{p-1}$. Calculate

$$\left\| \frac{1}{|\cdot|^{n-1}} \right\|_{L^q(\Gamma_{0,r})} = r^{1-n/p} \left(\sigma(\Gamma) \frac{p-1}{p-n} \right)^{\frac{p-1}{p}} \tag{1.10}$$

and insert the result into (1.9). A rearrangement and a simple estimate finally gives (1.8). □

Corollary 1.3.3 *Consider $n \in \mathbb{N}$ and $p \in (n, +\infty]$. There are constants $A = A_n$ and $B = B_n^{-1}$ (B_n given by (1.12)) such that for any $u \in C^1(\mathbb{R}^n)$ and any $x, y \in \mathbb{R}^n$ one has ($r = |x-y|$, $B(x,r)$ is the ball with center x and radius r)*

$$|u(y) - u(x)| \leq 2BA^{1/p} \left(\frac{p-1}{p-n} \right)^{\frac{p-1}{p}} \|\nabla u\|_{L^p(B(x,r)\cap B(y,r))} |x-y|^{1-\frac{n}{p}}. \tag{1.11}$$

Proof. Certainly, the intersection $V = B(x,r) \cap B(y,r)$ of the two balls is not empty. Introduce the following subsets Γ, Λ of the unit sphere in $\mathbb{R}^n$ by the requirement that $x + r\Gamma = (\partial B(x,r)) \cap B(y,r)$ and $y + r\Lambda = (\partial B(y,r)) \cap B(x,r)$, i.e., $\Gamma = \frac{1}{r}(\partial B(x,r) \cap B(y,r) - x)$ and $\Lambda = \frac{1}{r}(\partial B(y,r) \cap B(x,r) - y) = -\Gamma$. It is instructive to draw a picture of the sets introduced above.

Since $\Gamma_{x,r} = r\Gamma_{x,1}$ and $\Lambda_{y,r} = r\Lambda_{y,1}$ we find that

$$B_n = \frac{|\Gamma_{x,r} \cap \Lambda_{y,r}|}{|\Gamma_{x,r}|} = \frac{|\Gamma_{x,1} \cap \Lambda_{y,1}|}{|\Gamma_{x,1}|} \tag{1.12}$$

is a number between 0 and 1 which only depends on the dimension n. It follows $|\Gamma_{x,r}| = |\Lambda_{y,r}| = B_n^{-1}|W|$, $W = \Gamma_{x,r} \cap \Lambda_{y,r}$.

Now we estimate, using Lemma 1.3.1 and Hölder's inequality

$$\begin{aligned}|u(x)-u(y)||W| &\le \int_W |u(x)-u(z)|\boldsymbol{D}z + \int_W |u(z)-u(y)|\boldsymbol{D}z\\ &\le \int_{\Gamma_{x,r}} |u(x)-u(z)|\boldsymbol{D}z + \int_{\Lambda_{y,r}} |u(z)-u(y)|\boldsymbol{D}z\\ &\le \frac{r^n}{n}\int_{\Gamma_{x,r}} \frac{|\nabla u(z)|}{|x-y|^{n-1}}\boldsymbol{D}z + \frac{r^n}{n}\int_{\Lambda_{y,r}} \frac{|\nabla u(z)|}{|z-y|^{n-1}}\boldsymbol{D}z\\ &\le \frac{r^n}{n}\left(\|\nabla u\|_{L^p(\Gamma_{x,r})}\left\|\frac{1}{|x-\cdot|^{n-1}}\right\|_{L^q(\Gamma_{x,r})} + \|\nabla u\|_{L^p(\Lambda_{y,r})}\left\|\frac{1}{|y-\cdot|^{n-1}}\right\|_{L^q(\Lambda_{y,r})}\right)\\ &\le 2\frac{r^n}{n}\|\nabla u\|_{L^p(V)}\left\|\frac{1}{|\cdot|^{n-1}}\right\|_{L^q(\Gamma_{0,r})}.\end{aligned}$$

Taking (1.10), (1.12) and (1.6) into account and recalling $r = |x-y|$, estimate (1.11) follows with $A = \sigma(\Gamma)^{-1}$. □

Theorem 1.3.4 (Morrey's inequality) *Suppose $n < p \le +\infty$ and $u \in W^{1,p}(\mathbb{R}^n)$. Then there is a unique version u^* of u (i.e., $u^* = u$ almost everywhere) which is Hölder continuous of exponent $1-\frac{n}{p}$, i.e., $u^* \in \mathcal{C}^{0,1-\frac{n}{p}}(\mathbb{R}^n)$ and satisfies*

$$\|u^*\|_{\mathcal{C}^{0,1-\frac{n}{p}}(\mathbb{R}^n)} \le C\|u\|_{W^{1,p}(\mathbb{R}^n)} \tag{1.13}$$

where $C = C(n,p)$ is a universal constant. In addition the estimates in (1.7), (1.8) and (1.11) hold when u is replaced by u^*.

Proof. At first consider the case $n < p < \infty$. For $u \in \mathcal{C}^1_c(\mathbb{R}^n)$ Corollaries 1.3.2 and 1.3.3 imply ($\mathcal{C}_b(\mathbb{R}^n)$ denotes the space of bounded continuous functions on $\mathbb{R}^n$)

$$\|u\|_{\mathcal{C}_b(\mathbb{R}^n)} \le C\|u\|_{W^{1,p}(\mathbb{R}^n)} \quad \text{and} \quad \frac{|\mathrm{u(y)}-\mathrm{u(x)}|}{|\mathrm{y}-\mathrm{x}|^{1-\frac{\mathrm{n}}{\mathrm{p}}}} \le \mathrm{C}\|\nabla \mathrm{u}\|_{\mathrm{L^p(\mathbb{R}^n)}}.$$

This implies

$$\|u\|_{\mathcal{C}^{0,1-\frac{n}{p}}(\mathbb{R}^n)} \leq C\,\|u\|_{W^{1,p}(\mathbb{R}^n)}\,. \tag{1.14}$$

If $u \in W^{1,p}(\mathbb{R}^n)$ is given, there is a sequence of functions $u_j \in \mathcal{C}_c^1(\mathbb{R}^n)$ such that $u_j \to u$ in $W^{1,p}(\mathbb{R}^n)$. Estimate (1.14) implies that this sequence is also a Cauchy sequence in $\mathcal{C}^{0,1-\frac{n}{p}}(\mathbb{R}^n)$ and thus converges to a unique element u^* in this space. Clearly Estimate (1.13) holds for this limit element u^* and $u^* = u$ almost everywhere.

The case $p = \infty$ and $u \in W^{1,p}(\mathbb{R}^n)$ can be proven via a similar approximation argument. □

Corollary 1.3.5 (Morrey's inequality) *Let Ω be an open bounded subset of $\mathbb{R}^n$ with smooth boundary ($\mathcal{C}^1$) and $n < p \leq \infty$. Then for every $u \in W^{1,p}(\Omega)$ there exists a unique version u^* in $\mathcal{C}^{0,1-\frac{n}{p}}(\Omega)$ satisfying*

$$\|u^*\|_{\mathcal{C}^{0,1-\frac{n}{p}}(\Omega)} \leq C\,\|u\|_{W^{1,p}(\Omega)}\,. \tag{1.15}$$

with a universal constant $C = C(n,p,\Omega)$.

Proof. Under the assumptions of the corollary the extension theorem for Sobolev spaces applies according to which elements in $W^{1,p}(\Omega)$ are extended to all of $\mathbb{R}^n$ by zero such that there exists a continuous extension operator $J : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ (see for instance Theorem 48.35 of 10). Then, given $u \in W^{1,p}(\Omega)$, Theorem 1.3.4 implies that there is a continuous version $U^* \in \mathcal{C}^{0,1-\frac{n}{p}}(\mathbb{R}^n)$ of Ju which satisfies (1.13). Now define $u^* = U^*|_\Omega$. It follows

$$\|u^*\|_{\mathcal{C}^{0,1-\frac{n}{p}}(\Omega)} \leq \|U^*\|_{\mathcal{C}^{0,1-\frac{n}{p}}(\mathbb{R}^n)} \leq C\,\|Ju\|_{W^{1,p}(\mathbb{R}^n)} \leq C\,\|u\|_{W^{1,p}(\Omega)}\,.$$

□

1.4 Gagliardo-Nirenberg-Sobolev inequality

This important inequality is of the form

$$\|u\|_{L^q} \leq C \|\nabla u\|_{L^p}, \quad u \in \mathcal{C}_c^1(\mathbb{R}^n) \tag{1.16}$$

for a suitable exponent q for a given exponent p, $1 \leq p \leq n$. This exponent is easily determined through the scale covariance of the quantities in this inequality. For $\lambda > 0$ introduce u_λ by setting $u_\lambda(x) = u(\lambda x)$. A simple calculation shows $\|u_\lambda\|_{L^q} = \lambda^{-n/q} \|u\|_{L^q}$ and $\|\nabla u_\lambda\|_{L^p} = \lambda^{1-n/p} \|\nabla u\|_{L^p}$. Thus inserting u_λ into (1.16) gives

$$\lambda^{-n/q} \|u\|_{L^q} \leq C\lambda^{1-n/p} \|\nabla u\|_{L^p}$$

for all $\lambda > 0$. This is possible for all $u \in \mathcal{C}_c^1(\mathbb{R}^n)$ only if

$$1 - n/p + n/q = 0, \quad \text{i.e.,} \quad \frac{1}{\mathrm{p}} = \frac{1}{\mathrm{n}} + \frac{1}{\mathrm{q}}. \tag{1.17}$$

It is a standard notation to denote the exponent q which solves (1.17) by p^*, i.e.,

$$p^* = \frac{np}{n-p}$$

with the understanding that $p^* = \infty$ if $p = n$.

As we will show later the case $1 < p < n$ can easily be reduced to the case $p = 1$, thus we prove this inequality for $p = 1$, i.e., $p^* = 1^* = \frac{n}{n-1}$.

Theorem 1.4.1 *For all $u \in W^{1,1}(\mathbb{R}^n)$ one has*

$$\|u\|_{1^*} = \|u\|_{\frac{n}{n-1}} \le \prod_{i=1}^{n} \left(\int_{\mathbb{R}^n} |\partial_i u(x)| \boldsymbol{D}x \right)^{\frac{1}{n}} \le n^{-\frac{1}{2}} \|\nabla u\|_1 \tag{1.18}$$

Proof. According to Theorem 1.2.2 every element $u \in W^{1,1}(\mathbb{R}^n)$ is the limit of a sequence of elements $u_j \in \mathcal{C}_c^1(\mathbb{R}^n)$. Hence is suffices to prove this inequality for $u \in \mathcal{C}_c^1(\mathbb{R}^n)$, and this is done by induction on the dimension n.

We suggest that the reader proves this inequality for $n = 1$ and $n = 2$. Here we present first the case $n = 3$ before we come to the general case.

Suppose that $u \in \mathcal{C}_c^1(\mathbb{R}^3)$ is given. Observe that now $1^* = 3/2$. Introduce the notation $x^1 = (y_1, x_2, x_3)$, $x^2 = (x_1, y_2, x_3)$, and $x^3 = (x_1, x_2, y_3)$. The fundamental theorem of calculus implies for $i = 1,2,3$

$$|u(x)| \le \int_{-\infty}^{x_i} |\partial_i u(x^i)| \boldsymbol{D}y_i \le \int_{-\infty}^{\infty} |\partial_i u(x^i)| \boldsymbol{D}y_i,$$

hence multiplication of these three inequalities gives

$$|u(x)|^{\frac{3}{2}} \le \prod_{i=1}^{3} \left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| \boldsymbol{D}y_i \right)^{\frac{1}{2}}.$$

Now integrate this inequality with respect to x_1 and note that the first factor on the right does not depend on x_1:

$$\int_{\mathbb{R}} |u(x)|^{\frac{3}{2}} \boldsymbol{D}x_1 \le \left(\int_{-\infty}^{\infty} |\partial_1 u(x^1)| \boldsymbol{D}y_1 \right)^{\frac{1}{2}} \int_{\mathbb{R}} \prod_{i=2}^{3} \left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| \boldsymbol{D}y_i \right)^{\frac{1}{2}} \boldsymbol{D}x_1$$

Apply Hölder's inequality (for $p = q = 2$) to the second integral, this gives the estimate

$$\le \left(\int_{-\infty}^{\infty} |\partial_1 u(x^1)| \boldsymbol{D}y_1 \right)^{\frac{1}{2}} \prod_{i=2}^{3} \left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| \boldsymbol{D}x_1 \boldsymbol{D}y_i \right)^{\frac{1}{2}}.$$

Next we integrate this inequality with respect to x_2 and apply again Hölder's inequality to get

$$\begin{aligned}
&\int_{\mathbb{R}^2} |u(x)|^{\frac{3}{2}} \boldsymbol{D}x_1 \boldsymbol{D}x_2 \\
&\leq \left(\int_{\mathbb{R}^2} |\partial_2 u(x^2)| \boldsymbol{D}x_1 \boldsymbol{D}y_2\right)^{\frac{1}{2}} \int_{\mathbb{R}} \left(\int_{-\infty}^{\infty} |\partial_1 u(x^1)| \boldsymbol{D}y_1\right)^{\frac{1}{2}} \left(\int_{-\infty}^{\infty} |\partial_3 u(x^3)| \boldsymbol{D}x_1 \boldsymbol{D}y_3\right)^{\frac{1}{2}} \boldsymbol{D}x_2 \\
&\leq \left(\int_{\mathbb{R}^2} |\partial_2 u(x^2)| \boldsymbol{D}x_1 \boldsymbol{D}y_2\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^2} |\partial_1 u(x^1)| \boldsymbol{D}y_1 \boldsymbol{D}x_2\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^3} |\partial_3 u(x^3)| \boldsymbol{D}x_1 \boldsymbol{D}x_2 \boldsymbol{D}y_3\right)^{\frac{1}{2}}.
\end{aligned}$$

A final integration with respect to x_3 and applying Hölder's inequality as above implies

$$\begin{aligned}
&\int_{\mathbb{R}^3} |u(x)|^{\frac{3}{2}} \boldsymbol{D}x_1 \boldsymbol{D}x_2 \boldsymbol{D}x_3 \leq \left(\int_{\mathbb{R}^3} |\partial_1 u(x^1)| \boldsymbol{D}y_1 \boldsymbol{D}x_2 \boldsymbol{D}x_3\right)^{\frac{1}{2}} \times \\
&\left(\int_{\mathbb{R}^3} |\partial_2 u(x^2)| \boldsymbol{D}x_1 \boldsymbol{D}y_2 \boldsymbol{D}x_3\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^3} |\partial_3 u(x^3)| \boldsymbol{D}x_1 \boldsymbol{D}x_2 \boldsymbol{D}y_3\right)^{\frac{1}{2}} = \\
&\prod_{i=1}^{3} \left(\int_{\mathbb{R}^3} |\partial_i u(x)| \boldsymbol{D}x_1 \boldsymbol{D}x_2 \boldsymbol{D}x_3\right)^{\frac{1}{2}} \leq \left(\int_{\mathbb{R}^3} |\nabla u(x)| \boldsymbol{D}x_1 \boldsymbol{D}x_2 \boldsymbol{D}x_3\right)^{\frac{1}{2}}
\end{aligned}$$

which is the claimed inequality for $n = 3$.

The general case uses the same strategy. Naturally some more steps are necessary. Now we have $1^* = \frac{n}{n-1}$. For $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ introduce the variables $x^i = (x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n)$. The fundamental theorem of calculus implies for $i = 1, \ldots, n$

$$|u(x)| \leq \int_{\mathbb{R}} |\partial_i u(x^i)| \boldsymbol{D}y_i$$

and thus

$$|u(x)|^{\frac{n}{n-1}} \leq \prod_{i=1}^{n} \left(\int_{\mathbb{R}} |\partial_i u(x^i)| \boldsymbol{D}y_i\right)^{\frac{1}{n-1}}. \tag{1.19}$$

Recall Hölder's inequality for the product of $n-1$ functions in the form

$$\left\| \prod_{i=2}^{n} f_i \right\|_1 \leq \prod_{i=2}^{n} \|f_i\|_{n-1} \tag{1.20}$$

and integrate (1.19) with respect to x_1 to get

$$\begin{aligned}\int_{\mathbb{R}} |u(x)|^{\frac{n}{n-1}} \boldsymbol{D}x_1 &\leq \left(\int_{\mathbb{R}} |\partial_1 u(x^1)| \boldsymbol{D}y_1\right)^{\frac{1}{n-1}} \int_{\mathbb{R}} \prod_{i=2}^{n} \left(\int_{\mathbb{R}} |\partial_i u(x^i)| dy_i\right)^{\frac{1}{n-1}} \boldsymbol{D}x_1 \\ &\leq \left(\int_{\mathbb{R}} |\partial_1 u(x^1)| \boldsymbol{D}y_1\right)^{\frac{1}{n-1}} \prod_{i=2}^{n} \left(\int_{\mathbb{R}^2} |\partial_i u(x^i)| \boldsymbol{D}x_1 \boldsymbol{D}y_i\right)^{\frac{1}{n-1}} \\ &= \left(\int_{\mathbb{R}} |\partial_1 u(x^1)| \boldsymbol{D}y_1\right)^{\frac{1}{n-1}} \left(\int_{\mathbb{R}^2} |\partial_2 u(x^2)| \boldsymbol{D}x_1 \boldsymbol{D}y_2\right)^{\frac{1}{n-1}} \\ &\times \prod_{i=3}^{n} \left(\int_{\mathbb{R}^2} |\partial_i u(x^i)| \boldsymbol{D}x_1 \boldsymbol{D}y_i\right)^{\frac{1}{n-1}}\end{aligned}$$

where in the last step we isolated the x_2 independent term from the product. Now integrate this inequality with respect to x_2 and apply (1.20) again. This implies, after renaming the integration variable y_2,

$$\begin{aligned}&\int_{\mathbb{R}} |u(x)|^{\frac{n}{n-1}} \boldsymbol{D}x_1 \boldsymbol{D}x_2 \leq \left(\int_{\mathbb{R}^2} |\partial_2 u(x)| \boldsymbol{D}x_1 \boldsymbol{D}x_2\right)^{\frac{1}{n-1}} \\ &\times \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |\partial_1 u(x)| \boldsymbol{D}x_1\right)^{\frac{1}{n-1}} \prod_{i=3}^{n} \left(\int_{\mathbb{R}^2} |\partial_i u(x^i)| \boldsymbol{D}x_1 \boldsymbol{D}y_i\right)^{\frac{1}{n-1}} \boldsymbol{D}x_2 \leq \\ &\left(\int_{\mathbb{R}^2} |\partial_2 u(x)| \boldsymbol{D}x_1 \boldsymbol{D}x_2\right)^{\frac{1}{n-1}} \left(\int_{\mathbb{R}^2} |\partial_1 u(x)| \boldsymbol{D}x_1 \boldsymbol{D}x_2\right)^{\frac{1}{n-1}} \\ &\times \prod_{i=3}^{n} \left(\int_{\mathbb{R}^3} |\partial_i u(x^i)| \boldsymbol{D}x_1 \boldsymbol{D}x_2 \boldsymbol{D}y_i\right)^{\frac{1}{n-1}} \\ &= \prod_{i=1}^{2} \left(\int_{\mathbb{R}^2} |\partial_i u(x)| \boldsymbol{D}x_1 \boldsymbol{D}x_2\right)^{\frac{1}{n-1}} \times \prod_{i=3}^{n} \left(\int_{\mathbb{R}^3} |\partial_i u(x^i)| \boldsymbol{D}x_1 \boldsymbol{D}x_2 \boldsymbol{D}y_i\right)^{\frac{1}{n-1}}.\end{aligned}$$

Obviously one can repeat these steps successively for $x_3, \dots, x_n$ and one proves by induction that for $k \in \{1, \dots, n\}$ we get the estimate

$$\int_{\mathbb{R}^k} |u(x)|^{\frac{n}{n-1}} \boldsymbol{D}x_1 \boldsymbol{D}x_2 \cdots \boldsymbol{D}x_k \le \prod_{i=1}^{k} \left(\int_{\mathbb{R}^k} |\partial_i u(x)| \boldsymbol{D}x_1 \boldsymbol{D}x_2 \cdots \boldsymbol{D}x_k \right)^{\frac{n}{n-1}} \times \prod_{i=k+1}^{n} \left(\int_{\mathbb{R}^{k+1}} |\partial_i u(x^i)| \boldsymbol{D}x_1 \boldsymbol{D}x_2 \cdots \boldsymbol{D}x_k \boldsymbol{D}y_i \right)^{\frac{n}{n-1}}$$

where naturally for $k = n$ the second product does not occur. Thus for $k = n$ one has

$$\int_{\mathbb{R}^n} |u(x)|^{\frac{n}{n-1}} \boldsymbol{D}x_1 \boldsymbol{D}x_2 \cdots \boldsymbol{D}x_n \le \prod_{i=1}^{n} \left(\int_{\mathbb{R}^n} |\partial_i u(x)| \boldsymbol{D}x_1 \boldsymbol{D}x_2 \cdots \boldsymbol{D}x_n \right)^{\frac{n}{n-1}}$$

In order to improve this estimate recall Young's inequality in the elementary form $\prod_{i=1}^n A_i \le \frac{1}{n} \sum_{i=1}^n A_i^n$, where $A_i \ge 0$. Thus we get

$$\|u\|_{\frac{n}{n-1}} \le \prod_{i=1}^{n} \left(\int_{\mathbb{R}^n} |\partial_i u(x)| \boldsymbol{D}x \right)^{\frac{1}{n}} \le \frac{1}{n} \sum_{i=1}^{n} \int_{\mathbb{R}^n} |\partial_i u(x)| \boldsymbol{D}x$$

and by Hölder's inequality one knows $\sum_{i=1}^n |\partial_i u(x)| \le \sqrt{n} |\nabla u(x)|$, hence $\|u\|_{\frac{n}{n-1}} \le \frac{1}{\sqrt{n}} \|\nabla u\|_1$. □

Remark 1.4.2 *The starting point of our estimates was the identity $u(x) = \int_{-\infty}^{x_i} \partial_i u(x^i) \boldsymbol{D}y_i$ and the resulting estimate*

$$|u(x)| \le \int_{\mathbb{R}} |\partial_i u(x^i)| \boldsymbol{D}y_i, \quad i = 1, \dots, n.$$

If we write

$$u(x) = \frac{1}{2} \left(\int_{-\infty}^{x_i} \partial_i u(x^i) dy_i - \int_{x_i}^{\infty} \partial_i u(x^i) \boldsymbol{D}y_i \right),$$

we can improve this estimate to

$$|u(x)| \leq \frac{1}{2} \int_{\mathbb{R}} |\partial_i u(x^i)| \boldsymbol{D} y_i, \quad i = 1, \ldots, n.$$

Next we look at the case $1 < p < n$. As we will see it can easily be reduced to the case $p = 1$.

Theorem 1.4.3 (Gagliardo-Nirenberg-Sobolev inequality) *If $1 \leq p < n$ then, for all $u \in W^{1,p}(\mathbb{R}^n)$, with $p^* = \frac{np}{n-p}$,*

$$\|u\|_{p^*} \leq \frac{1}{\sqrt{n}} \frac{p(n-1)}{n-p} \|\nabla u\|_p . \tag{1.21}$$

Proof. Since elements in $W^{1,p}(\mathbb{R}^n)$ can be approximated by elements in $\mathcal{C}_c^1(\mathbb{R}^n)$ it suffices to prove Estimate (1.21) for $u \in \mathcal{C}_c^1(\mathbb{R}^n)$. For such a function u consider the function $v = |u|^s \in \mathcal{C}_c^1(\mathbb{R}^n)$ for an exponent $s > 1$ to be determined later. We have $\nabla v = s|u|^{s-1}\mathrm{sgn(u)}\nabla \mathrm{u}$ and thus by applying (1.18) to v we get

$$\| |u|^s \|_{1^*} \leq \frac{1}{\sqrt{n}} \|\nabla |u|^s\|_1 = \frac{s}{\sqrt{n}} \left\| |u|^{s-1} \nabla u \right\|_1 \leq \frac{s}{\sqrt{n}} \left\| |u|^{s-1} \right\|_q \|\nabla u\|_p \tag{1.22}$$

where q is the Hölder conjugate exponent of p. Note that this estimate can be written as

$$\|u\|_{s1^*}^s \leq \frac{s}{\sqrt{n}} \|u\|_{(s-1)q}^{s-1} \|\nabla u\|_p .$$

Now choose s such that $s1^* = (s-1)q$. This gives $s = \frac{q}{q-1^*} = \frac{p^*}{1^*}$ and accordingly the last estimate can be written as

$$\|u\|_{p^*}^s \leq \frac{s}{\sqrt{n}} \|u\|_{p^*}^{s-1} \|\nabla u\|_p .$$

Inserting the value $s = \frac{p(n-1}{n-p}$ of s now yields (1.21). □

Corollary 1.4.4 *Suppose that $\Omega \subset \mathbb{R}^n$ is a bounded open set with $\mathcal{C}^1$-boundary. Then for all $p \in [1,n)$ and $1 \leq q \leq p^*$ there is a constant $C = C(\Omega, p, q)$ such that for all $u \in W^{1,p}(\Omega)$*

$$\|u\|_q \leq C \|u\|_{1,p}.$$

Proof. Under the given conditions on Ω one can show that every $u \in W^{1,p}(\Omega)$ has an extension to $Ju \in W^{1,p}(\mathbb{R}^n)$ (i.e., $Ju|\Omega = u$ and $J : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ is continuous). Then for $u \in \mathcal{C}^1(\bar{\Omega}) \cap W^{1,p}(\Omega)$

$$\|u\|_{L^{p^*}(\Omega)} \leq C \|Ju\|_{L^{p^*}(\mathbb{R}^n)} \leq C \|\nabla(Ju)\|_{L^p(\mathbb{R}^n)} \leq C \|u\|_{W^{1,p}(\Omega)}. \tag{1.23}$$

Since $\mathcal{C}^1(\Omega)$ is dense in $W^{1,p}(\Omega)$, this estimate holds for all $u \in W^{1,p}(\Omega)$. If now $1 \leq q < p^*$ a simple application of Hölder's inequality gives

$$\|u\|_{L^q(\Omega)} \leq \|u\|_{L^{p^*}(\Omega)} \|1\|_{L^s(\Omega)} = \|u\|_{L^{p^*}(\Omega)} |\Omega|^{1/s} \leq C|\Omega|^{1/s} \|u\|_{W^{1,p}(\Omega)}$$

where $\frac{1}{s} + \frac{1}{p^*} = \frac{1}{q}$. □

1.4.1 Continuous Embeddings of Sobolev spaces

In this short review of the classical theory of Sobolev spaces we can only discuss the main embeddings results. In the literature one finds many additional cases.

For convenience of notation let us introduce, for a given number $r \geq 0$,

$$r_+ = \begin{cases} r & \text{if r} \notin \mathbb{N}_0 \\ r + \delta & \text{if r} \in \mathbb{N}_0 \end{cases}$$

where $\delta > 0$ is some arbitrary small number. For a number $r = k + \alpha$ with $k \in \mathbb{N}_0$ and $0 \leq \alpha < 1$ we write $\mathcal{C}^r(\Omega)$ for $\mathcal{C}^{k,\alpha}(\Omega)$.

Lemma 1.4.5 *For $i \in \mathbb{N}$ and $p \geq n$ and $i > n/p$ (i.e., $i \geq 1$ if $p > n$ and $i \geq 2$ if $p = n$) one has*

$$W^{i,p}(\Omega) \hookrightarrow \mathcal{C}^{i-(n/p)+}(\Omega)$$

and there is a constant $C > 0$ such that for all $u \in W^{i,p}(\Omega)$

$$\|u\|_{\mathcal{C}^{i-(n/p)+}(\Omega)} \leq C\,\|u\|_{i,p} \tag{1.24}$$

Proof. As earlier it suffices to prove (1.24) for $u \in \mathcal{C}^\infty(\Omega)$. For such u and $p > n$ and $|\alpha| \leq i-1$ apply Morrey's inequality to get

$$\|D^\alpha u\|_{\mathcal{C}^{0,1-n/p}(\Omega)} \leq C\,\|D^\alpha u\|_{i,p}$$

and therefore with $\mathcal{C}^{i-n/p}(\Omega) \equiv \mathcal{C}^{i-1,1-n/p}(\Omega)$, we get (1.24).

If $p = n$ (and thus $i \geq 2$) choose $q \in (1,n)$ close to n so that $i > n/q$ and $q^* = \frac{qn}{n-q} > n$. Then, by the first part of Theorem (1.4.6) and what we have just shown

$$W^{i,n}(\Omega) \hookrightarrow W^{i,q}(\Omega) \hookrightarrow W^{i-1,q^*}(\Omega) \hookrightarrow \mathcal{C}^{i-2,1-n/q^*}(\Omega).$$

As $q \uparrow n$ implies $n/q^* \downarrow 0$, we conclude $W^{i,n}(\Omega) \hookrightarrow \mathcal{C}^{i-2,\alpha}(\Omega)$ for any $\alpha \in (0,1)$ which is written as

$$W^{i,n}(\Omega) \hookrightarrow \mathcal{C}^{i-(n/n)+}(\Omega).$$

□

Theorem 1.4.6 (Sobolev Embedding Theorems) *Assume that $\Omega = \mathbb{R}^n$ or that Ω is a bounded open subset of $\mathbb{R}^n$ with a $\mathcal{C}^1$-boundary; furthermore assume that $1 \leq p < \infty$ and $k, m \in \mathbb{N}$ with $m \leq k$. Then one has:*

(1) If $p < n/m$, then $W^{k,p}(\Omega) \hookrightarrow W^{k-m,q}(\Omega)$ for $q = \frac{np}{n-pm}$ or $\frac{1}{q} = \frac{1}{p} - \frac{m}{n} > 0$, and there is a constant $C > 0$ such that

$$\|u\|_{k-m,q} \leq C\,\|u\|_{k,p} \quad \text{for all } u \in W^{k,p}(\Omega). \tag{1.25}$$

(2) If $p > n/k$, then $W^{k,p}(\Omega) \hookrightarrow \mathcal{C}^{k-(n/p)_+}(\Omega)$ and there is a constant $C > 0$ such that

$$\|u\|_{\mathcal{C}^{k-(n/p)_+}(\Omega)} \leq C\,\|u\|_{k,p} \quad \text{for all } u \in W^{k,p}(\Omega). \tag{1.26}$$

Proof. Suppose $p < n/m$ and $u \in W^{k,p}(\Omega)$; then $D^\alpha u \in W^{1,p}(\Omega)$ for all $|\alpha| \leq k-1$. Corollary 1.4.4 implies $D^\alpha u \in L^{p^*}(\Omega)$ for all $|\alpha| \leq k-1$ and therefore $W^{k,p}(\Omega) \hookrightarrow W^{k-1,p^*}(\Omega)$ and there is a constant $C_1 > 0$ such that

$$\|u\|_{k-1,p_1} \leq C_1\,\|u\|_{k,p} \tag{1.27}$$

for all $u \in W^{k,p}(\Omega)$, with $p_1 = p^*$. Next define p_j, $j \geq 2$, inductively by $p_j = p^*_{j-1}$. Thus $\frac{1}{p_j} = \frac{1}{p_{j-1}} - \frac{1}{n}$ and since $p < n/m$ we have $\frac{1}{p_m} = \frac{1}{p} - \frac{m}{n} > 0$. Therefore we can apply (1.27) repeatedly and find that the following inclusion maps are all bounded:

$$W^{k,p}(\Omega) \hookrightarrow W^{k-1,p_1}(\Omega) \hookrightarrow W^{k-2,p_2}(\Omega) \cdots \hookrightarrow W^{k-m,p_m}(\Omega)$$

and part (1) follows.

In order to prove part (2) consider $p > n/k$. For $p \geq n$ the statement follows from Lemma 1.4.5. Now consider the case $n > p > n/k$ and choose the largest m such that $1 \leq m < k$ and $n/m > p$. Define $q \geq n$ by $q = \frac{np}{n-mp}$ (i.e., $\frac{1}{q} = \frac{1}{p} - \frac{m}{n} > 0$). Then, by what we have established above, the following inclusion maps are all bounded:

$$W^{k,p}(\Omega) \hookrightarrow W^{k-m,q}(\Omega) \hookrightarrow \mathcal{C}^{k-m-(n/q)_+}(\Omega) = \mathcal{C}^{k-m-(\frac{n}{p}-m)_+}(\Omega) = \mathcal{C}^{k-(n/p)_+}(\Omega)$$

which is the estimate of Part (2). □

In the case $p = 2$ and $\Omega = \mathbb{R}^n$ one has the Fourier transform $\mathcal{F}$ available as a unitary operator on $L^2(\mathbb{R}^n)$. This allows to give a convenient characterization of the Sobolev space $H^k(\mathbb{R}^n) = W^{k,2}(\mathbb{R}^n)$ and to prove a useful embedding result.

Recall that for $u \in H^k(\mathbb{R}^n)$ one has $\mathcal{F}(D^\alpha u)(p) = \mathrm{i}^{|\mathrm{ff}|}\mathrm{p}^{\mathrm{ff}}\mathcal{F}(\mathrm{u})(\mathrm{p})$. Hence we can characterize this space as

$$\begin{aligned} H^k(\mathbb{R}^n) &= \{u \in L^2(\mathbb{R}^n) : p^\alpha \mathcal{F}(u) \in L^2(\mathbb{R}^n), |\alpha| \le k\} \\ &= \left\{u \in L^2(\mathbb{R}^n) : (1+p^2)^{k/2}\mathcal{F}(u) \in L^2(\mathbb{R}^n)\right\}. \end{aligned}$$

This definition can be extended to arbitrary $s \in \mathbb{R}$ and thus we can introduce the spaces

$$H^s(\mathbb{R}^n) = \left\{u \in L^2(\mathbb{R}^n) : (1+p^2)^{s/2}\mathcal{F}(u) \in L^2(\mathbb{R}^n)\right\}.$$

As we are going to show this space can be continuously embedded into the space

$$\mathcal{C}_b^k(\mathbb{R}^n) = \left\{ f \in \mathcal{C}^k(\mathbb{R}^n) : \|f\|_{k,\infty} = \sup_{|\alpha| \le k} \sup_{x \in \mathbb{R}^n} |D^\alpha f(x)| < \infty \right\}.$$

Theorem 1.4.7 *For $k \in \mathbb{N}$ and $s > k + n/2$ the Sobolev space $H^s(\mathbb{R}^n)$ is continuously embedded into the space $C_b^k(\mathbb{R}^n)$ and one has for all $u \in H^s(\mathbb{R}^n)$*

$$\|u\|_{k,\infty} \leq C\,\|u\|_{s,2}, \qquad \lim_{|x|\to\infty} |D^\alpha u(x)| = 0,\ |\alpha| \leq k.$$

Proof. Recall that the Lemma of Riemann-Lebesgue says that the Fourier transform of an $L^1(\mathbb{R}^n)$ function is continuous and vanishes at infinity. For $|\alpha| \leq k$ and $s > k + n/2$ one knows

$$\int_{\mathbb{R}^n} \frac{|p^{2\alpha}|}{(1+p^2)^s} \boldsymbol{D}p = C_\alpha^2 < \infty.$$

Thus, for $u \in H^s(\mathbb{R}^n)$ we can estimate

$$\int_{\mathbb{R}^n} |p^\alpha (\mathcal{F}u)(p)| \boldsymbol{D}p \leq C_\alpha \left(\int_{\mathbb{R}^n} (1+p^2)^s |\mathcal{F}u(p)|^2 \boldsymbol{D}p \right)^{1/2} = C_\alpha \|u\|_{s,2}$$

and therefore for all $x \in \mathbb{R}^n$

$$|D^\alpha u(x)| = \left| \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{ipx}} \mathrm{p}^{\mathrm{ff}} (\mathcal{F}\mathrm{u})(\mathrm{p}) \boldsymbol{D}\mathrm{p} \right| \leq C_\alpha \|u\|_{s,2}.$$

It follows $\|u\|_{k,\infty} \leq \|u\|_{s,2}$. By applying the Lemma of Riemann-Lebesgue we conclude. □

1.4.2 Compact Embeddings of Sobolev spaces

Here we show that some of the continuous embeddings established above are actually compact, that is they map bounded subsets into precompact sets. There are various ways to prove these compactness results. We present a proof which is based on the characterization of compact subsets $M \subset L^q(\mathbb{R}^n)$, due to Kolmogorov and Riesz 15, 21.

Theorem 1.4.8 (Kolmogorov-Riesz compactness criterion) *Suppose* $1 \leq q < \infty$. *Then a subset* $M \subset L^q(\mathbb{R}^n)$ *is precompact if, and only if M satisfies the following three conditions:*

(a) M is bounded, i.e.,

$$\exists_{C<\infty} \forall_{f \in M} \; \|f\| \leq C;$$

(b)

$$\forall_{\epsilon>0} \exists_{R<\infty} \forall_{f \in M} \left\| \pi_R^{\perp} f \right\|_q < \epsilon;$$

(c)

$$\forall_{\epsilon>0} \exists_{r>0} \forall_{f \in M} \forall_{\substack{y \in \mathbb{R}^n \\ |y|<r}} \left\| \tau_y(f) - f \right\|_q < \epsilon.$$

Here the following notation is used: $\pi_R^{\perp}$ is the operator of multiplication with the characteristic function of the set $\{x \in \mathbb{R}^n : |x| > R\}$ and τ_y denotes the operator of translation by $y \in \mathbb{R}^n$, i.e., $\tau_y(f)(x) = f(x+y)$.

Remark 1.4.9 *If* $\Omega \subset \mathbb{R}^n$ *is an open bounded subset we can consider* $L^q(\Omega)$ *as a subset of* $L^q(\mathbb{R}^n)$ *by extending all elements* $f \in L^q(\Omega)$ *by* 0 *to all of* $\mathbb{R}^n$. *Then the above* *characterization provides also a characterization of precompact subset $M \subset L^q(\Omega)$ where naturally condition (b) is satisfied always and where in condition (c) we have to use these extensions.*

There are several versions of compact embedding results depending on the assumptions on the domain $\Omega \subset \mathbb{R}^n$ which are used. The following version is already quite comprehensive though there are several newer results of this type.

Theorem 1.4.10 (Rellich-Kondrachov compactness theorem) *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Assume that the boundary of Ω is sufficiently smooth and that $1 \leq p < \infty$ and $k = 1, 2, \ldots$. Then the following holds:*

(a) The following embeddings are compact:

(i) $kp < n$: $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$, $1 \leq q < p^* = \frac{np}{n-kp}$;

(ii) $kp = n$: $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$, $1 \leq q < \infty$;

(iii) $kp > n$: $W^{k,p}(\Omega) \hookrightarrow C^0_b(\Omega)$.

(b) For the subspaces $W^{k,p}_0(\Omega)$ the embeddings (i) - (iii) are compact for arbitrary open sets Ω.

Proof. In some detail we present here only the proof of embedding (i) of part (a) for $k = 1$. For the remaining proofs we refer to the specialized literature 2, 1.

According to Corollary 1.4.4 the inclusion mapping $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$ is continuous for $1 \le q \le p^*$. We have to show that every bounded subset $M \subset W^{k,p}(\Omega)$ is precompact in $L^q(\Omega)$ for $1 \le q < p^*$. This is done by the Kolmogorov-Riesz compactness criterion. By Remark 1.4.9 only conditions (a) and (c) have to be verified for M considered as a subset of $L^q(\Omega)$. Since we know that this inclusion map is continuous, it follows that M is bounded in $L^q(\Omega)$ too and thus Condition (a) of the Kolmogorov-Riesz criterion is verified and we are left with verifying Condition (c).

Observe that for $1 \le q < p^*$ Hölder's inequality implies

$$\|u\|_q \le \|u\|_1^{\alpha} \|u\|_{p^*}^{1-\alpha}, \quad \alpha = \frac{1}{q}\frac{p^* - q}{p^* - 1} \in (0,1).$$

Now let $M \subset W^{1,p}(\Omega)$ be bounded; then this set is bounded in $L^{p^*}(\Omega)$ and hence there is a constant $C < \infty$ such that for all $u \in M$ we have

$$\|u\|_q \le C \|u\|_1^{\alpha}$$

and it follows

$$\left\|\tau_y u - u\right\|_q \le 2C \left\|\tau_y u - u\right\|_1^{\alpha}, \quad \forall\, u \in M \tag{1.28}$$

where we assume that for $u \in W^{1,p}(\Omega)$ the translated element $\tau_y u$ is extended by zero outside Ω. Therefore it suffices to verify condition (c) of Theorem 1.4.8 for the norm $\|\cdot\|_1$. For $i = 1,2,\ldots$ introduce the sets

$$\Omega_i = \{x \in \Omega : d(x,\partial\Omega) > 2/i\},$$

where $d(x,\partial\Omega)$ denotes the distance of the point x from the boundary $\partial\Omega$ of Ω. Another application of Hölder's inequality gives, for all $u \in M$,

$$\begin{aligned}\int_{\Omega\setminus\Omega_i} |u(x)| \boldsymbol{D}x &\le \left(\int_{\Omega\setminus\Omega_i} |u(x)|^{p^*} \boldsymbol{D}x\right)^{1/p^*} \left(\int_{\Omega\setminus\Omega_i} \boldsymbol{D}x\right)^{1-\frac{1}{p^*}} \\ &\le \|u\|_{p^*} |\Omega\setminus\Omega_i|^{1-\frac{1}{p^*}} \le C_M |\Omega\setminus\Omega_i|^{1-\frac{1}{p^*}}\end{aligned}$$

where C_M is a bound for M in $L^{p^*}(\Omega)$. Given $\epsilon > 0$ we can therefore find $i_0 = i_0(\epsilon)$ such that for $i \ge i_0$

$$\int_{\Omega\setminus\Omega_i} |u(x)| \boldsymbol{D}x < \epsilon/4$$

holds for all $u \in M$. Extend $u \in M$ outside Ω by 0 to get

$$\hat{u}(x) = \begin{cases} u(x), & x \in \Omega, \\ 0, & \text{otherwise.} \end{cases}$$

For a fixed $i \geq i_0$ and $y \in \mathbb{R}^n$, $|y| < 1/i$, we estimate

$$\begin{aligned} \|\tau_y u - u\|_1 &= \int_{\Omega_i} |u(x+y) - u(x)| \boldsymbol{D}x + \int_{\Omega \setminus \Omega_i} |\hat{u}(x+y) - \hat{u}(x)| \boldsymbol{D}x \\ &\leq \int_{\Omega_i} |u(x+y) - u(x)| \boldsymbol{D}x + \epsilon/2 \end{aligned}$$

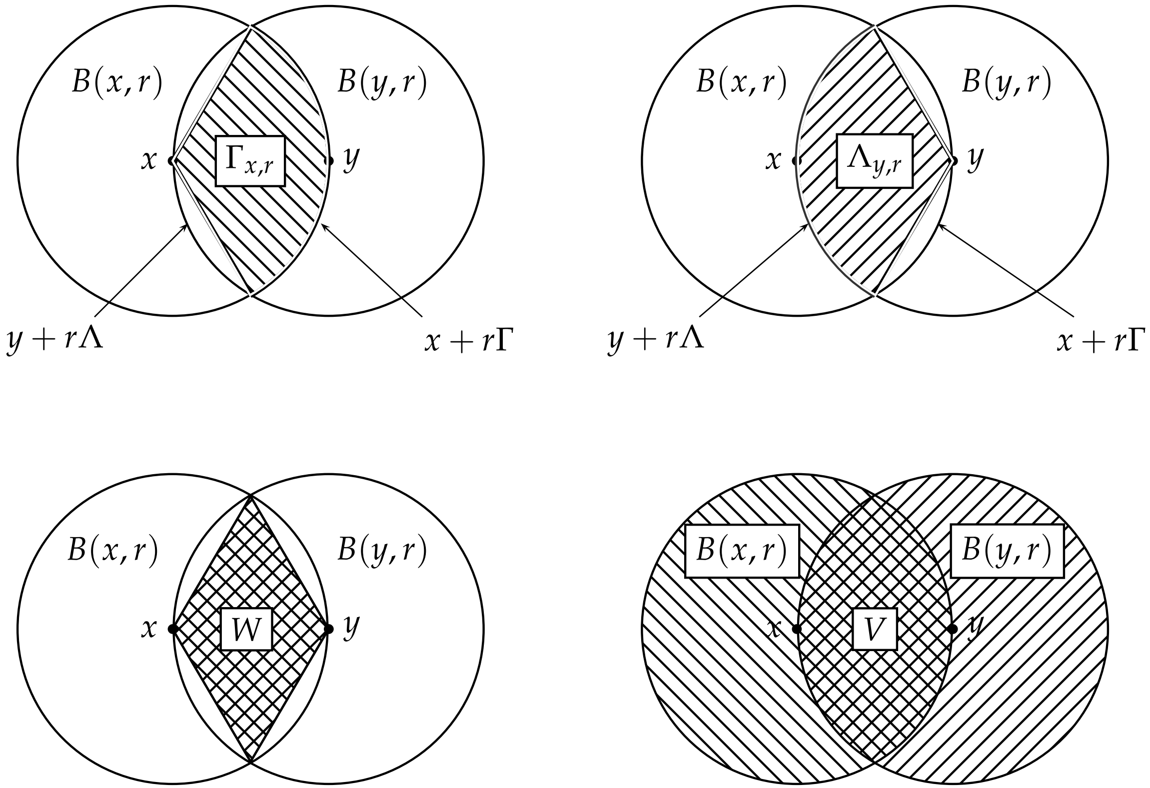
And the integral is estimated as follows (p' denotes the Hölder conjugate exponent of p):

$$\begin{aligned} &= \int_{\Omega_i} \left| \int_0^1 \frac{\boldsymbol{D}}{\boldsymbol{D}t} u(x+ty) \boldsymbol{D}t \right| \boldsymbol{D}x = \int_{\Omega_i} \left| \int_0^1 y \cdot \nabla u(x+ty) \boldsymbol{D}t \right| \boldsymbol{D}x \leq |y| \int_{\Omega_{2i}} |\nabla u(x)| \boldsymbol{D}x \\ &\leq |y| |\Omega_{2i}|^{\frac{1}{p'}} \|\nabla u\|_{L^p(\Omega_{2i})} \leq |y| |\Omega_{2i}|^{\frac{1}{p'}} C \leq |y| |\Omega|^{\frac{1}{p'}} C \end{aligned}$$

It follows that there is $r_0 > 0$ such that $\|\tau_y u - u\|_1 < \epsilon$ for all $|y| < r_0$. By estimate (1.28) we conclude that Condition (c) of Theorem 1.4.8 holds and therefore by this theorem $M \subset W^{1,p}(\Omega)$ is precompact in $L^q(\Omega)$. □

Remark 1.4.11 *The general case of $W^{k,p}(\Omega)$ with $k > 1$ follows from the following observation which can be proven similarly.*

> *For $m \geq 1$ and $\frac{1}{q} > \frac{1}{p} - \frac{m}{n} > 0$ the inclusion of $W^{k,p}(\Omega)$ into $W^{k-m,q}(\Omega)$ is compact.*



$W=\Gamma_{x,r}\cap\Gamma_{y,r}$ $\qquad$ $r=|x-y|$ $\qquad$ $V=B(x,r)\cap B(y,r)$

Figure 1.1: Intersecting balls and the related sets $\Gamma_{x,r}$ and $\Lambda_{y,r}$

Chapter 2

Hilbert-Schmidt and trace class operators

Abstract: This chapter introduces two subspaces of the space of compact operators and presents their theory in substantial detail. These spaces of operators are important in various areas of functional analysis and in applications of operator theory to quantum physics. Accordingly, after the characterization of Hilbert-Schmidt and trace class operators has been presented, the spectral representation for these operators is proven. Furthermore the dual spaces (spaces of continuous linear functionals) of these two spaces of operators are determined and their rôle in the description of locally convex topologies on the space $\mathcal{B}(\mathcal{H})$ of all bounded linear operators on a Hilbert space $\mathcal{H}$ is explained. Finally two results are included which are mainly used in quantum physics:

Partial trace for trace class operators on tensor products of separable Hilbert spaces and Schmidt decomposition.

2.1 Basic theory

Since they are closely related we discuss Hilbert-Schmidt and trace class operators together.

Definition 2.1.1 *A bounded linear operator A on a separable Hilbert space $\mathcal{H}$ is called a* **Hilbert-Schmidt operator** *respectively a* **trace class operator** *if, and only if, for some orthonormal basis $\{e_n : n \in \mathbb{N}\}$ the sum*

$$\sum_{n=1}^{\infty} \|Ae_n\|^2 = \sum_{n=1}^{\infty} \langle e_n, A^* Ae_n \rangle$$

respectively the sum

$$\sum_{n=1}^{\infty} \langle e_n, |A| e_n \rangle = \sum_{n=1}^{\infty} \left\| |A|^{1/2} e_n \right\|^2$$

is finite, where $|A|$ is the modulus of A (Definition 21.5.1).

The set of all Hilbert-Schmidt operators (trace class operators) on $\mathcal{H}$ is denoted by $\mathcal{B}_2(\mathcal{H})$ $(\mathcal{B}_1(\mathcal{H}))$.

Lemma 2.1.2 *The two sums in Definition 2.1.1 do not dependent on the particular basis and thus one defines the* **trace norm** $\|\cdot\|_1$ *of a trace class operator* A *by*

$$\|A\|_1 = \sum_{n=1}^{\infty} \langle e_n, |A| e_n \rangle \tag{2.1}$$

and the **Hilbert-Schmidt norm** $\|\cdot\|_2$ *of a Hilbert-Schmidt operator* A *by*

$$\|A\|_2 = \|A^* A\|_1^{1/2} = \Big(\sum_{n=1}^{\infty} \|Ae_n\|^2 \Big)^{1/2} . \tag{2.2}$$

Proof. Parseval's identity implies for any two orthonormal bases $\{e_n : n \in \mathbb{N}\}$ and $\{f_n : n \in \mathbb{N}\}$ of $\mathcal{H}$ and any bounded linear operator B

$$\sum_{n=1}^{\infty} \|Be_n\|^2 = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |\langle Be_n, f_m \rangle|^2 = \sum_{m=1}^{\infty} \|B^* f_m\|^2 .$$

Take another orthonormal basis $\{h_n : n \in \mathbb{N}\}$, the same calculation then shows that we can continue the above identity by

$$\sum_{n=1}^{\infty} \|B^{**} h_n\|^2 = \sum_{m=1}^{\infty} \|Be_n\|^2 ,$$

since $B^{**} = B$, and hence this sum is independent of the particular basis. If we apply this identity for $B = |A|^{1/2}$ we see that the defining sum for trace class is independent of the particular choice of the basis. □

Corollary 2.1.3 *For every* $A \in \mathcal{B}_2(\mathcal{H})$ *one has* $\|A\|_2 = \|A^*\|_2$.

Proof. This is immediate from the proof of Lemma 2.1.2.

Basic properties of the set of all Hilbert-Schmidt operators and of the Hilbert-Schmidt norm are collected in the following theorem.

Theorem 2.1.4 *a) $\mathcal{B}_2(\mathcal{H})$ is a vector space which is invariant under taking adjoints, i.e., $A \in \mathcal{B}_2(\mathcal{H})$ if, and only if, $A^* \in \mathcal{B}_2(\mathcal{H})$; furthermore for all $A \in \mathcal{B}_2(\mathcal{H})$,*

$$\|A^*\|_2 = \|A\|_2 \;;$$

b) The Hilbert-Schmidt norm $\|\cdot\|_2$ dominates the operator norm $\|\cdot\|$, i.e.,

$$\|A\| \leq \|A\|_2$$

for all $A \in \mathcal{B}_2(\mathcal{H})$.

c) For all $A \in \mathcal{B}_2(\mathcal{H})$ and all $B \in \mathcal{B}(\mathcal{H})$ one has $AB \in \mathcal{B}_2(\mathcal{H})$ and $BA \in \mathcal{B}_2(\mathcal{H})$ with the estimates

$$\|AB\|_2 \leq \|A\|_2 \|B\|, \quad \|BA\|_2 \leq \|B\| \, \|A\|_2$$

i.e., $\mathcal{B}_2(\mathcal{H})$ is a two-sided ideal in $\mathcal{B}(\mathcal{H})$.

d) The vector space $\mathcal{B}_2(\mathcal{H})$ is a Hilbert space with the inner product

$$\langle A, B \rangle_{HS} = \sum_{n=1}^{\infty} \langle Ae_n, Be_n \rangle = \mathrm{Tr}(A^*B), \quad A, B \in \mathcal{B}_2(\mathcal{H}) \tag{2.3}$$

and the Hilbert-Schmidt norm is defined by this inner product by $\|A\|_2 = \sqrt{\langle A, A\rangle_{HS}}$.

Proof. a) It is obvious that scalar multiples λA of elements $A \in \mathcal{B}_2(\mathcal{H})$ again belong to $\mathcal{B}_2(\mathcal{H})$. If $A, B \in \mathcal{B}_2(\mathcal{H})$ and if $\{e_n\}$ is an ONB then the estimate

$$\|(A+B)e_n\|^2 \leq 2(\|Ae_n\|^2 + \|Be_n\|^2)$$

immediately implies $A + B \in \mathcal{B}_2(\mathcal{H})$. Thus $\mathcal{B}_2(\mathcal{H})$ is a vector space. Corollary 2.1.3 now implies that for $A \in \mathcal{B}_2(\mathcal{H})$ also $A^* \in \mathcal{B}_2(\mathcal{H})$ and $\|A^*\|_2 = \|A\|_2$.

b) A given unit vector $h \in \mathcal{H}$ can be considered as an element of an ONB $\{e_n\}$, therefore we can estimate for $A \in \mathcal{B}_2(\mathcal{H})$

$$\|Ah\|^2 \leq \sum_n \|Ae_n\|^2 = \|A\|_2^2,$$

and it follows

$$\|A\| = \sup\{\|Ah\| : h \in \mathcal{H}, \|h\| = 1\} \leq \|A\|_2.$$

c) For $A \in \mathcal{B}_2(\mathcal{H})$ and $B \in \mathcal{B}(\mathcal{H})$ and every basis vector e_n one has $\|BAe_n\|^2 \leq \|B\|^2 \|Ae_n\|^2$ and thus (2.2) implies $\|BA\|_2 \leq \|B\| \|A\|_2$. Next part a) says $\|AB\|_2 = \|(AB)^*\|_2 = \|B^*A^*\|_2 \leq \|B^*\| \|A^*\|_2 = \|B\| \|A\|_2$. And it follows $AB, BA \in \mathcal{B}_2(\mathcal{H})$.

d) For an ONB $\{e_n\}$ and any $A, B \in \mathcal{B}_2(\mathcal{H})$ one has, using Schwarz' inequality twice,

$$\sum_n |\langle Ae_n, Be_n\rangle| \leq \|A\|_2 \|B\|_2.$$

We conclude that (2.3) is well defined on $\mathcal{B}_2(\mathcal{H})$ and then that it is a anti-linear in the first and linear in the second argument. Obviously $\langle A, A\rangle_{HS} \geq 0$ for all $A \in \mathcal{B}_2(\mathcal{H})$ and $\langle A, A\rangle_{HS} = 0$ if, and only if, $Ae_n = 0$ for all elements e_n of an ONB of $\mathcal{H}$, hence $A = 0$. Therefore (2.3) is an inner product on $\mathcal{B}_2(\mathcal{H})$ and clearly this inner product defines the Hilbert-Schmidt norm.

Finally we show completeness of this inner product space. Suppose that $\{A_n\}$ is a Cauchy sequence in $\mathcal{B}_2(\mathcal{H})$. Then, given $\epsilon > 0$, there is n_0 such that $\|A_m - A_n\|_2 \leq \epsilon$ for all $m, n \geq n_0$. Since $\|A\| \leq \|A\|_2$ this sequence is also a Cauchy sequence in $\mathcal{B}(\mathcal{H})$ and hence it converges to a unique $A \in \mathcal{B}(\mathcal{H})$, by Theorem 21.3.3. For an ONB $\{e_j\}$, $n \geq n_0$, and all $N \in \mathbb{N}$ we have

$$\sum_{j=1}^{N} \left\|(A - A_n)e_j\right\|^2 = \lim_{m\to\infty} \sum_{j=1}^{N} \left\|(A_m - A_n)e_j\right\|^2 \leq \lim_{m\to\infty} \|A_m - A_n\|_2^2 \leq \epsilon^2$$

and conclude

$$\sum_{j=1}^{\infty} \left\|(A - A_n)e_j\right\|^2 \leq \epsilon^2.$$

This shows $A - A_n \in \mathcal{B}_2(\mathcal{H})$, hence $A = A_n + (A - A_n) \in \mathcal{B}_2(\mathcal{H})$ and $\|A - A_n\|_2 \leq \epsilon$ for $n \geq n_0$ and A is the limit of the sequence $\{A_n\}$ in the Hilbert-Schmidt norm. □

Though trace class operators share many properties with Hilbert-Schmidt operators, some of the proofs are more complicated. For instance the fact that the modulus of a bounded linear operator is not subadditive does not allow such a simple proof of the fact that the set of all trace class operators is closed under addition of operators, as in the case of Hilbert-Schmidt operators. For this and some other important properties the following proposition will provide substantial simplifications in the proofs (see also 9).

Proposition 2.1.5 *For a bounded linear operator A on $\mathcal{H}$ the following statements are equivalent:*

(a) For some (and then for every) ONB $\{e_n\}$ one has $S_1(A) = \sum_n \langle e_n, |A| e_n \rangle < \infty$.

(b) $S_2(A) = \inf\{\|B\|_2 \|C\|_2 : B, C \in \mathcal{B}_2(\mathcal{H}), A = BC\} < \infty$.

(c) $S_3(A) = \sup\{\sum_n |\langle e_n, A f_n \rangle| : \{e_n\}, \{f_n\}$ are ONS in $\mathcal{H}\} < \infty$

For $A \in \mathcal{B}_1(\mathcal{H})$ one has $S_1(A) = S_2(A) = S_3(A) = \|A\|_1$.

Proof. Suppose $S_1(A) < \infty$. Write the polar decomposition $A = U|A|$ as $A = BC$ with $B = U|A|^{1/2}$ and $C = |A|^{1/2}$. Then

$$\|B\|_2^2 = \sum_n \left\| U|A|^{1/2} e_n \right\|^2 \leq \sum_n \left\| |A|^{1/2} e_n \right\|^2 = S_1(A) < \infty$$

and $\|C\|_2^2 = S_1(A) < \infty$ and thus $S_2(A) \le S_1(A) < \infty$.

Next suppose that $S_2(A) < \infty$ and take any ONS $\{e_n\}$ and $\{f_n\}$ in $\mathcal{H}$. Write $A = BC$ with $B, C \in \mathcal{B}_2(\mathcal{H})$ and estimate

$$\begin{aligned}\sum_n |\langle e_n, Af_n\rangle| = \sum_n |\langle B^* e_n, Cf_n\rangle| &\le \big(\sum_n \|B^* e_n\|^2\big)^{1/2}\big(\sum_n \|Cf_n\|^2\big)^{1/2} \\ &\le \|B^*\|_2 \|C\|_2 = \|B\|_2 \|C\|_2 .\end{aligned}$$

It follows $S_3(A) \le S_2(A) < \infty$.

Finally assume that $S_3(A) < \infty$ and take an ONB $\{e_n\}$ for $\overline{\mathrm{Ran}(|A|)}$. Then $f_n = Ue_n$ is an ONB for $\overline{\mathrm{Ran(A)}}$ and thus we can estimate

$$S_1(A) = \sum_n \langle e_n, |A| e_n\rangle = \sum_n \langle e_n, U^* A e_n\rangle = \sum_n \langle f_n, Ae_n\rangle \le S_3(A),$$

and therefore $S_1(A) \le S_3(A) < \infty$.

If $A \in \mathcal{B}_1(\mathcal{H})$, then by definition $S_1(A) < \infty$. The above chain of estimates shows $S_3(A) \le S_2(A) \le S_1(A) \le S_3(A)$ and thus we have equality. □

Theorem 2.1.6 *a) $\mathcal{B}_1(\mathcal{H})$ is a vector space which is invariant under taking adjoints, i.e., $A \in \mathcal{B}_1(\mathcal{H})$ if, and only if, $A^* \in \mathcal{B}_1(\mathcal{H})$; furthermore for all $A \in \mathcal{B}_1(\mathcal{H})$,*

$$\|A^*\|_1 = \|A\|_1 \ ;$$

b) The trace norm $\|\cdot\|_1$ dominates the operator norm $\|\cdot\|$, i.e.,

$$\|A\| \le \|A\|_1$$

for all $A \in \mathcal{B}_1(\mathcal{H})$.

c) For all $A \in \mathcal{B}_1(\mathcal{H})$ and all $B \in \mathcal{B}(\mathcal{H})$ one has $AB \in \mathcal{B}_1(\mathcal{H})$ and $BA \in \mathcal{B}_1(\mathcal{H})$ with the estimates

$$\|AB\|_1 \leq \|A\|_1 \|B\|, \quad \|BA\|_1 \leq \|B\| \|A\|_1$$

i.e., $\mathcal{B}_1(\mathcal{H})$ is a two-sided ideal in $\mathcal{B}(\mathcal{H})$.

d) The vector space $\mathcal{B}_1(\mathcal{H})$ is a Banach space under the trace norm $\|\cdot\|_1$.

Proof. a) For a scalar multiple λA of $A \in \mathcal{B}_1(\mathcal{H})$ one obviously has $S_1(\lambda A) = |\lambda| S_1(A)$ and thus $\lambda A \in \mathcal{B}_1(\mathcal{H})$ and $\|\lambda A\|_1 = |\lambda| \|A\|_1$. A simple calculation shows $S_3(A^*) = S_3(A)$ and therefore by Proposition 2.1.5, $A^* \in \mathcal{B}_1(\mathcal{H})$ whenever $A \in \mathcal{B}_1(\mathcal{H})$ and $\|A^*\|_1 = \|A\|_1$.

For $A, B \in \mathcal{B}_1(\mathcal{H})$ we know by Proposition 2.1.5 that $S_3(A)$ and $S_3(B)$ are finite. From the definition of $S_3(\cdot)$ we read off $S_3(A+B) \leq S_3(A) + S_3(B)$, thus $S_3(A+B)$ is finite, i.e., $A + B \in \mathcal{B}_1(\mathcal{H})$.

If for $A \in \mathcal{B}_1(\mathcal{H})$ one has $\|A\|_1 = 0$, then $|A|^{1/2} e_n = 0$ for all elements of an ONB $\{e_n\}$ of $\mathcal{H}$, hence $|A|^{1/2} = 0$ and thus $A = 0$. Therefore the trace norm $\|\cdot\|_1$ is indeed a norm on the vector space $\mathcal{B}_1(\mathcal{H})$.

b) Given unit vectors $e, f \in \mathcal{H}$ we can consider them as being an element of an ONS $\{e_n\}$ respectively of an ONS $\{f_n\}$; then

$$|\langle e, Af \rangle| \leq \sum_n |\langle e_n, Af_n \rangle| \leq S_3(A),$$

it follows

$$\|A\| = \sup\{|\langle e, Af \rangle| : e, f \in \mathcal{H}, \|e\| = \|f\| = 1\} \leq S_3(A) = \|A\|_1.$$

c) If $A \in \mathcal{B}_1(\mathcal{H})$ has a decomposition $A = CD$ with $C, D \in \mathcal{B}_2(\mathcal{H})$, then for any $B \in \mathcal{B}(\mathcal{H})$, BA has a decomposition $BA = BCD$ with $BC, D \in \mathcal{B}_2(\mathcal{H})$, by Theorem 2.1.4, part c). We conclude

$$S_2(BA) \leq \|BC\|_2 \|D\|_2 \leq \|B\| \|C\|_2 \|D\|_2$$

and thus $S_2(BA) \leq \|B\| S_2(A) < \infty$. It follows $\|BA\|_1 \leq \|B\| \|A\|_1$. Since we have established that $\mathcal{B}_1(\mathcal{H})$ is invariant under taking adjoints, we can prove $AB \in \mathcal{B}_1(\mathcal{H})$ as in the case of Hilbert-Schmidt operators.

d) Finally we show completeness of the normed space $\mathcal{B}_1(\mathcal{H})$. Suppose that $\{A_n\}$ is a Cauchy sequence in $\mathcal{B}_1(\mathcal{H})$. Then, given $\epsilon > 0$, there is n_0 such that $\|A_m - A_n\|_1 \leq \epsilon$ for all $m, n \geq n_0$. Since $\|A\| \leq \|A\|_1$ this sequence is also a Cauchy sequence in $\mathcal{B}(\mathcal{H})$ and hence it converges to a unique $A \in \mathcal{B}(\mathcal{H})$, by Theorem 21.3.3.

Fix $n \geq n_0$; for any ONS $\{e_j\}$ and $\{f_j\}$ in $\mathcal{H}$ and any $N \in \mathbb{N}$ we have

$$\sum_{j=1}^{N} |\langle e_j, (A - A_n) f_j \rangle| = \lim_{m\to\infty} \sum_{j=1}^{N} |\langle e_j, (A_m - A_n) f_j \rangle| \leq \lim_{m\to\infty} S_3(A_m - A_n) \leq \epsilon$$

and conclude

$$\sum_{j=1}^{\infty} |\langle e_j, (A - A_n) f_j \rangle| \leq \epsilon.$$

This shows $A - A_n \in \mathcal{B}_1(\mathcal{H})$, hence $A = A_n + (A - A_n) \in \mathcal{B}_1(\mathcal{H})$ and $\|A - A_n\|_1 \leq \epsilon$. □

Corollary 2.1.7 *The space of trace class operators is continuously embedded into the space of Hilbert-Schmidt operators:*

$$\mathcal{B}_1(\mathcal{H}) \hookrightarrow \mathcal{B}_2(\mathcal{H}).$$

Proof. According to the definitions one has $\|A\|_2^2 = \|A^*A\|_1$. When we apply parts c), b), and a) in this order we can estimate

$$\|A^*A\|_1 \leq \|A^*\| \, \|A\|_1 \leq \|A^*\|_1 \, \|A\|_1 = \|A\|_1^2$$

and thus $\|A\|_2 \leq \|A\|_1$ which implies our claim. □

Corollary 2.1.8 *On the space of all trace class operators the* **trace** *is well defined by ($\{e_n\}$ is any ONB of $\mathcal{H}$)*

$$\mathrm{Tr}(A) = \sum_n \langle e_n, A e_n \rangle, \quad A \in \mathcal{B}_1(\mathcal{H}) \,. \tag{2.4}$$

This function $\mathrm{Tr} : \mathcal{B}_1(\mathcal{H}) \longrightarrow \mathbb{K}$ *is linear and satisfies for all* $A \in \mathcal{B}_1(\mathcal{H})$

a) $|\mathrm{Tr}(A)| \leq \|A\|_1$;

b) $\mathrm{Tr}(A^*) = \overline{\mathrm{Tr}(A)}$;

c) $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$; *for all* $B \in \mathcal{B}_1(\mathcal{H})$;

d) $\mathrm{Tr}(UAU^*) = \mathrm{Tr}(A)$; *for all unitary operators* U *on* $\mathcal{H}$.

Proof. We know that $A \in \mathcal{B}_1(\mathcal{H})$ can be written as $A = BC$ with $B, C \in \mathcal{B}_2(\mathcal{H})$. For any ONB $\{e_n\}$ we estimate

$$\sum_n |\langle e_n, Ae_n\rangle| \leq \sum_n \|B^* e_n\| \, \|Ce_n\| \leq \|B^*\|_2 \, \|C\|_2 = \|B\|_2 \, \|C\|_2 < \infty$$

and conclude

$$\sum_n |\langle e_n, Ae_n\rangle| \leq S_2(A) = \|A\|_1 .$$

As earlier one proves that the sum in (2.4) does not depend on the choice of the particular basis. Linearity in $A \in \mathcal{B}_1(\mathcal{H})$ is obvious. Thus a) holds. The proof of b) is an elementary calculation.

If $A, B \in \mathcal{B}_1(\mathcal{H})$ choose another ONB $\{f_m\}$ and use the completeness relation to calculate

$$\sum_n \langle e_n, ABe_n\rangle = \sum_n \langle A^* e_n, Be_n\rangle = \sum_n \sum_m \langle A^* e_n, f_m\rangle \langle f_m, Be_n\rangle =$$
$$= \sum_m \sum_n \langle B^* f_m, e_n\rangle \langle e_n, Af_m\rangle = \sum_m \langle B^* f_m, Af_m\rangle = \sum_m \langle f_m, BAf_m\rangle,$$

hence c) holds, since we know that the above series converge absolutely so that the order of summation can be exchanged. Part d) is just a reformulation of the fact that the trace is independent of the basis which is used to calculate it. □

Theorem 2.1.9 (Spectral representation of Hilbert-Schmidt and trace class operators) *Let* $\mathcal{H}$ *be a separable Hilbert space and denote by* $\mathcal{B}_c(\mathcal{H})$ *the space of all compact operators on* $\mathcal{H}$ *(see Theorem 22.3.1). Then* *a) $\mathcal{B}_2(\mathcal{H}) \subset \mathcal{B}_c(\mathcal{H})$, i.e., Hilbert-Schmidt and thus trace class operators are compact.*

b) A bounded operator A on $\mathcal{H}$ is a Hilbert-Schmidt operator respectively a trace class operator if, and only if, there are two orthonormal bases $\{e_n\}$ and $\{x_n\}$ of $\mathcal{H}$ and there is a sequence $\{\lambda_n\}$ in $\mathbb{K}$ with

$$\sum_n |\lambda_n|^2 < \infty \quad \textit{respectively} \quad \sum_n |\lambda_n| < \infty$$

such that

$$Ax = \sum_n \lambda_n \langle e_n, x\rangle x_n \quad \textit{for all} \quad x \in \mathcal{H} \tag{2.5}$$

and then one has

$$\|A\|_2 = \Big(\sum_n |\lambda_n|^2\Big)^{1/2} \quad \textit{respectively} \quad \|A\|_1 = \sum_n |\lambda_n|.$$

Proof. a) Suppose that $\{e_n\}$ is an ONB of $\mathcal{H}$; denote by P_N the orthogonal projector onto the closed subspace $[e_1, \dots, e_N]$ spanned by $e_1, \dots, e_N$. Then for $A \in \mathcal{B}_2(\mathcal{H})$ one has

$$\|A - AP_N\|_2^2 = \sum_n \|(A - AP_N)e_n\|^2 = \sum_{n=N+1}^{\infty} \|Ae_n\|^2,$$

hence $\|A - AP_N\| \le \|A - AP_N\|_2 \longrightarrow 0$ as $N \to \infty$. Therefore A is the norm limit of the sequence of finite rank operators AP_N and thus compact by Theorem 22.3.2.

By Corollary 2.1.7 we know $\mathcal{B}_1(\mathcal{H}) \subset \mathcal{B}_2(\mathcal{H})$, hence trace class operators are compact.

b) Suppose that $A \in \mathcal{B}_j(\mathcal{H})$, $j = 1$ or $j = 2$, is given. By part a) we know that A and its modulus $|A|$ are compact. The polar decomposition (Theorem 21.5.2) relates A and $|A|$ by $A = U|A|$ where U is a partial isometry from $\operatorname{ran}|A|$ to $\overline{\operatorname{ran} A}$.

According to the Riesz-Schauder Theorem (**??**) the compact operator $|A|$ has the following spectral representation:

$$|A|x = \sum_j \lambda_j \langle e_j, x\rangle e_j \quad \text{for all} \quad x \in \mathcal{H} \tag{2.6}$$

with the specifications:

(i) $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n-1} \geq \lambda_n \geq 0$ are the eigen-values of $|A|$ enumerated in decreasing order and repeated in this list according to their multiplicity;

(ii) $\{e_j\}$ are the normalized eigen-vectors for the eigen-values λ_j;

(iii) the multiplicity of every eigen-value $\lambda_j > 0$ is finite, and if there are infinitely many eigen-values then $\lambda_j \longrightarrow 0$ as $j \to \infty$.

If the ONS $\{e_j\}$ is not complete we can extend it to an ONB $\{e'_n\}$ of $\mathcal{H}$ and calculate, using (2.6),

$$\|A\|_1 = \sum_n \langle e'_n, |A| e'_n\rangle = \sum_j \lambda_j < \infty \quad \text{for} \quad \mathrm{A} \in \mathcal{B}_1(\mathcal{H}),$$

respectively

$$\|A\|_2^2 = \sum_n \big\| |A| e'_n \big\|^2 = \sum_j \lambda_j^2 < \infty \quad \text{for} \quad \mathrm{A} \in \mathcal{B}_2(\mathcal{H}).$$

If we apply the partial isometry U to the representation (2.6) we get (2.5) with $x_n = Ue_n$.

Conversely suppose that an operator A has the representation (2.5). If $\sum_n |\lambda_n|^2 < \infty$ holds one has

$$\sum_j \|Ax_j\|^2 = \sum_j \sum_n |\lambda_n \langle e_n, x_j\rangle|^2 = \sum_n |\lambda_n|^2 \sum_j |\langle e_n, x_j\rangle|^2 = \sum_n |\lambda_n|^2,$$

thus $\|A\|_2^2 = \sum_n |\lambda_n|^2 < \infty$ and therefore $A \in \mathcal{B}_2(\mathcal{H})$.

Suppose that $\sum_n |\lambda_n| < \infty$ holds. In the Exercises we show that (2.5) implies

$$|A|x = \sum_n |\lambda_n| \langle e_n, x\rangle e_n, \quad x \in \mathcal{H}.$$

It follows $\|A\|_1 = \mathrm{Tr}(|A|) = \sum_n |\lambda_n| < \infty$, hence $A \in \mathcal{B}_1(\mathcal{H})$ and we conclude.

□

Remark 2.1.10 *One can define Hilbert-Schmidt and trace class operator also for the case of operators between two different Hilbert spaces as follows: For two separable Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ a bounded linear operator $A : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$ is called a* **Hilbert-Schmidt operator** *if, and only if, there is an orthonormal basis $\{e_n\}$ of $\mathcal{H}_1$ such that*

$$\sum_{n=1}^{\infty} \|Ae_n\|_2^2 < \infty$$

where $\|\cdot\|_2$ is the norm of $\mathcal{H}_2$.

A bounded linear operator $A : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$ is called a **trace class operator** *or a* **nuclear operator** *if, and only if, its modulus $|A| = \sqrt{A^*A}$ is a trace class operator on $\mathcal{H}_1$.*

With this slightly more general definitions the results presented above still hold with obvious modifications. We mention the spectral representation of Theorem 2.1.9.

A bounded linear operator $A : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$ is a Hilbert-Schmidt respectively a trace class operator if, and only if, it is of the form

$$Ax = \sum_{n=1}^{\infty} \lambda_n \langle e_n^1, x \rangle_1 e_n^2, \quad \text{for all } x \in \mathcal{H}_1 \tag{2.7}$$

where $\{e_n^i\}$ is an orthonormal system in $\mathcal{H}_i$, $i = 1,2$ and where the sequence of num-

bers $\lambda_n \neq 0$ *satisfies* $\sum_n |\lambda_n|^2 < \infty$ *respectively* $\sum_n |\lambda_n| < \infty$.

2.2 Dual spaces for the spaces of compact and trace class operators

The space of linear operators on $\mathcal{H}$ which have a finite rank is denoted by $\mathcal{B}_f(\mathcal{H})$. The following corollary highlights important results which in essence have been proven already in the last few theorems.

Corollary 2.2.1 *For any separable Hilbert space* $\mathcal{H}$ *one has*

$$\mathcal{B}_f(\mathcal{H}) \subset \mathcal{B}_1(\mathcal{H}) \subset \mathcal{B}_2(\mathcal{H}) \subset \mathcal{B}_c(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$$

and all the embeddings are continuous and dense. $\mathcal{B}_f(\mathcal{H})$ *is dense in* $\mathcal{B}_j(\mathcal{H})$, $j = 1,2$ *and in* $\mathcal{B}_c(\mathcal{H})$.

Proof. In the proofs of the last two results it was shown in particular that

$$\mathcal{B}_f(\mathcal{H}) \subset \mathcal{B}_j(\mathcal{H}) \subset \mathcal{B}_c(\mathcal{H}), \quad j = 1,2$$

holds and that the finite rank operators are dense in $\mathcal{B}_c(\mathcal{H})$ and in $\mathcal{B}_j(\mathcal{H})$ for $j = 1,2$. Parts b) of Theorem 2.1.4 respectively Theorem 2.1.6 imply that the embeddings $\mathcal{B}_j(\mathcal{H}) \hookrightarrow \mathcal{B}_c(\mathcal{H})$, $j = 1,2$, are continuous when $\mathcal{B}_c(\mathcal{H})$ is equipped with the operator norm. By Corollary 2.1.7 we conclude. □

According to Theorem 2.1.4 the space of Hilbert-Schmidt operators $\mathcal{B}_2(\mathcal{H})$ is a Hilbert space. Hence, according to the definition of the inner product (2.3) the continuous linear functionals f on this space are given by

$$f(A) = \mathrm{Tr}(BA) \quad \text{for all} \quad A \in \mathcal{B}_2(\mathcal{H}), \quad \text{for some} \quad B = B_f \in \mathcal{B}_2(\mathcal{H})\,. \tag{2.8}$$

Part c) of Theorem 2.1.6 says that the space of the trace class operators is a two-sided ideal in $\mathcal{B}(\mathcal{H})$, hence $\mathrm{Tr}(BA)$ is well defined for all $B \in \mathcal{B}(\mathcal{H})$ and all $A \in \mathcal{B}_1(\mathcal{H})$ and Part a) of Corollary 2.1.8 allows to estimate this trace by

$$|\mathrm{Tr}(BA)| \leq \|B\| \, \|A\|_1 \,. \tag{2.9}$$

Therefore, for fixed $B \in \mathcal{B}(\mathcal{H})$, $f_B(A) = \mathrm{Tr}(BA)$ is a continuous linear functional on $\mathcal{B}_1(\mathcal{H})$, and for fixed $A \in \mathcal{B}_1(\mathcal{H})$, $g_A(B) = \mathrm{Tr}(BA)$ is a continuous linear functional on $\mathcal{B}(\mathcal{H})$. Here we are interested in the space $\mathcal{B}_1(\mathcal{H})'$ of all continuous linear functionals on $\mathcal{B}_1(\mathcal{H})$ and in the space $\mathcal{B}_c(\mathcal{H})'$ of all continuous linear functionals on $\mathcal{B}_c(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$. Note that according to Corollary 2.2.1 the restriction of a continuous linear functional on $\mathcal{B}_c(\mathcal{H})$ to $\mathcal{B}_2(\mathcal{H})$ is a continuous linear functional on $\mathcal{B}_2(\mathcal{H})$ and thus given by the trace, i.e., formula (2.8).

Theorem 2.2.2 *For a separable Hilbert $\mathcal{H}$ the space of all continuous linear functionals $\mathcal{B}_c(\mathcal{H})'$ on the space $\mathcal{B}_c(\mathcal{H})$ of all compact operators on $\mathcal{H}$ and the space $\mathcal{B}_1(\mathcal{H})$ of all trace class operators are (isometrically) isomorphic, i.e.,*

$$\mathcal{B}_c(\mathcal{H})' \cong \mathcal{B}_1(\mathcal{H})\,.$$

The isomorphism $\mathcal{B}_1(\mathcal{H}) \longrightarrow \mathcal{B}_c(\mathcal{H})'$ *is given by* $B \longrightarrow \phi_B$ *with*

$$\phi_B(A) = \mathrm{Tr}(BA) \text{ for all } A \in \mathcal{B}_c(\mathcal{H}). \tag{2.10}$$

Proof. As mentioned above, given $F \in \mathcal{B}_c(\mathcal{H})'$, we know $F|_{\mathcal{B}_2(\mathcal{H})} \in \mathcal{B}_2(\mathcal{H})'$ and thus there is a unique $B \in \mathcal{B}_2(\mathcal{H})$ such that

$$F(A) = \mathrm{Tr}(BA) \quad \text{for all} \quad A \in \mathcal{B}_2(\mathcal{H}).$$

In order to show that actually $B \in \mathcal{B}_1(\mathcal{H})$ we use the characterization of trace class operators as given in Proposition 2.1.5 and estimate $S_3(B)$. To this end take any two ONS $\{e_n\}$ and $\{f_n\}$ in $\mathcal{H}$ and observe that there is $\alpha_n \in \mathbb{R}$ such that

$$e^{\mathrm{iff_n}} \langle f_n, Be_n \rangle = |\langle f_n, Be_n \rangle|.$$

Introduce the finite rank operators $[e_n, f_n]$ defined by $[e_n, f_n]x = \langle f_n, x\rangle e_n$ and then the finite rank operators

$$A_m = \sum_{n=1}^{m} e^{\mathrm{iff_n}} [e_n, f_n].$$

Since $\|A_m x\|^2 = \sum_{n=1}^m |\langle f_n, x\rangle|^2 \leq \|x\|^2$, one has $\|A_m\| \leq 1$. Thus we write, for any $m \in \mathbb{N}$,

$$\sum_{n=1}^{m} |\langle f_n, Be_n\rangle| = \sum_{n=1}^{m} e^{\mathrm{iff_n}} \langle f_n, Be_n\rangle = \mathrm{Tr}(BA_m) = F(A_m)$$

since $\langle f_n, Be_n\rangle = \mathrm{Tr}(B[e_n, f_n])$. We conclude

$$\sum_{n=1}^{m} |\langle f_n, Be_n\rangle| \leq \|F\|',$$

thus $S_3(B) \leq \|F\|'$ and hence $B \in \mathcal{B}_1(\mathcal{H})$. Introduce the continuous linear functional $\phi_B : \mathcal{B}_c(\mathcal{H}) \longrightarrow \mathbb{K}$ by

$$\phi_B(A) = \mathrm{Tr}(BA) \quad \text{for all} \quad A \in \mathcal{B}_c(\mathcal{H}).$$

We conclude that every $F \in \mathcal{B}_c(\mathcal{H})'$ is of the form $F = \phi_B$ with a unique $B \in \mathcal{B}_1(\mathcal{H})$. Now by (2.9) it follows

$$\|F\|' = \sup\{|F(A)| : A \in \mathcal{B}_c(\mathcal{H}), \|A\| \leq 1\} \leq \|B\|_1 .$$

In order to show $\|F\|' = \|\phi_B\|' = \|B\|_1$ recall that $\|B\|_1 = \mathrm{Tr}(|B|)$ when B has the polar decomposition $B = U|B|$ with a partial isometry U. For an ONB $\{e_n\}$ of $\mathcal{H}$ form the finite rank operator $A_m = \sum_{n=1}^{m}[e_n, e_n]U^*$ and calculate

$$\mathrm{Tr}(BA_m) = \mathrm{Tr}(A_m B) = \mathrm{Tr}(\sum_{n=1}^{m}[e_n, e_n]U^* B) = \sum_{n=1}^{m}\mathrm{Tr}([e_n, e_n]|B|) = \sum_{n=1}^{m}\langle e_n, |B|e_n\rangle.$$

It follows

$$\|\phi_B\|' \geq |\mathrm{Tr}(BA_m)| \geq \sum_{n=1}^{m}\langle e_n, |B|e_n\rangle$$

for all $m \in \mathbb{N}$ and thus $\|\phi_B\|' \geq \|B\|_1$, and we conclude. Basic properties of the trace show that the map $B \longrightarrow \phi_B$ is linear. Hence this map is an isometric isomorphism form $\mathcal{B}_1(\mathcal{H})$ to $\mathcal{B}_c(\mathcal{H})'$. □

In a similar way one can determine the dual space of the space of all trace class operators.

Theorem 2.2.3 *For a separable Hilbert $\mathcal{H}$ the space of all continuous linear functionals $\mathcal{B}_1(\mathcal{H})'$ on the space $\mathcal{B}_1(\mathcal{H})$ of all trace class operators on $\mathcal{H}$ and the space $\mathcal{B}(\mathcal{H})$ of all bounded linear operators are (isometrically) isomorphic, i.e.,*

$$\mathcal{B}_1(\mathcal{H})' \cong \mathcal{B}(\mathcal{H})\,.$$

The isomorphism $\mathcal{B}(\mathcal{H}) \longrightarrow \mathcal{B}_1(\mathcal{H})'$ is given by $B \longrightarrow \Psi_B$ where

$$\Psi_B(A) = \mathrm{Tr}(BA) \text{ for all } A \in \mathcal{B}_1(\mathcal{H}). \tag{2.11}$$

Proof. In a first step we show that given $f \in \mathcal{B}_1(\mathcal{H})'$ there is a unique $B = B_f \in \mathcal{B}(\mathcal{H})$ such that $f = \psi_B$ where again ψ_B is defined by the trace, i.e., $\Psi_B(A) = \mathrm{Tr}(BA)$ for all $A \in \mathcal{B}_1(\mathcal{H})$. For all $x, y \in \mathcal{H}$ define

$$b_f(x, y) = f([y, x])$$

where the operator $[y,x]$ is defined as above. Since $[y,x] \in \mathcal{B}_f(\mathcal{H}) \subset \mathcal{B}_1(\mathcal{H})$, b_f is well-defined on $\mathcal{H} \times \mathcal{H}$. Linearity of f implies immediately that b_f is a sesquilinear form on $\mathcal{H}$. This form is continuous: For all $x,y \in \mathcal{H}$ the estimate

$$|b_f(x,y)| = |f([y,x])| \leq \|f\|' \, \|[y,x]\|_1 \leq \|f\|' \, \|x\| \, \|y\|$$

holds, since by Proposition 2.1.5 one has, using Schwarz' and Bessel's inequality, $\|[y,x]\|_1 = S_3([y,x]) \leq \|x\| \, \|y\|$. Therefore by Theorem 20.2.1 there is a unique bounded linear operator B such that $b_f(x,y) = \langle x, By\rangle$, i.e.,

$$f([y,x]) = \langle x, By\rangle = \mathrm{Tr}(B[y,x]) \quad \text{for all } x,y \in \mathcal{H}.$$

The last identity follows from the completeness relation for an ONB $\{e_n\}$ of $\mathcal{H}$: $\mathrm{Tr}(B[y,x]) = \sum_n \langle e_n, B[y,x]e_n\rangle = \sum_n \langle e_n, By\langle x,e_n\rangle\rangle = \sum_n \langle e_n, By\rangle\langle x,e_n\rangle = \langle x, By\rangle$. By linearity this representation of f is extended to $\mathcal{B}_f(\mathcal{H}) \subset \mathcal{B}_1(\mathcal{H})$. And since both f and Tr are continuous with respect to the trace norm this representation has a unique extension to all of $\mathcal{B}_1(\mathcal{H})$ ($\mathcal{B}_f(\mathcal{H})$ is dense in $\mathcal{B}_1(\mathcal{H})$):

$$f(A) = \mathrm{Tr}(BA) = \Psi_B(A) \quad \text{for all } A \in \mathcal{B}_1(\mathcal{H}).$$

Linearity of Tr implies easily that $B \longrightarrow \mathcal{B}_1(\mathcal{H})'$ is a linear map from $\mathcal{B}(\mathcal{H})$ to $\mathcal{B}_1(\mathcal{H})'$. Finally we show that this map is isometric.

The continuity estimate (2.9) for the trace gives for $B \in \mathcal{B}(\mathcal{H})$

$$\|\Psi_B\|' = \sup\{|\Psi_B(A)| : A \in \mathcal{B}_1(\mathcal{H}), \|A\|_1 \leq 1\} \leq \|B\| \, .$$

We can assume $B \neq 0$. Then $\|B\| > 0$ and there is $x \in \mathcal{H}$, $\|x\| \leq 1$ such that $\|Bx\| \geq \|B\| - \epsilon$ for any $\epsilon \in (0, \|B\|)$. Introduce $\xi = \frac{Bx}{\|Bx\|}$ and calculate as above

$$\Psi_B([x,\xi]) = \mathrm{Tr}(B[x,\xi]) = \langle \xi, Bx\rangle = \|Bx\| \geq \|B\| - \epsilon,$$

hence $\|\Psi_B\|' \geq \|B\| - \epsilon$. This holds for any $0 < \epsilon < \|B\|$. We conclude

$$\|\Psi_B\|' \geq \|B\|$$

and thus $\|\Psi_B\|' = \|B\|$ and $B \longrightarrow \Psi_B$ is an isometric map from $\mathcal{B}(\mathcal{H})$ onto $\mathcal{B}_1(\mathcal{H})'$.

□

Remark 2.2.4 *According to this result one has the following useful expressions for the trace norm and the operator norm: The trace norm of $T \in \mathcal{B}_1(\mathcal{H})$ is given by*

$$\|T\|_1 = \sup |\mathrm{Tr}(BT)| \tag{2.12}$$

where the sup *is taken over all* $B \in \mathcal{B}(\mathcal{H})$ *with* $\|B\| = 1$ *and similarly the norm of* $B \in \mathcal{B}(\mathcal{H})$ *is*

$$\|B\| = \sup |\mathrm{Tr}(BT)| \tag{2.13}$$

where the sup *is taken over all* $T \in \mathcal{B}_1(\mathcal{H})$ *with* $\|T\|_1 = 1$*. Since* $\mathcal{B}_1(\mathcal{H})$ *is generated by the cone of its positive elements one also has*

$$\|B\| = \sup |\mathrm{Tr}(BW)| \tag{2.14}$$

where the sup *is taken over all density matrices* W*, i.e.,* $W \in \mathcal{B}_1(\mathcal{H})$*,* $W \geq 0$*, and* $\|W\|_1 = \mathrm{Tr}(W) = 1$*.*

Remark 2.2.5 *It is instructive to compare the chain of continuous dense embeddings*

$$\mathcal{B}_f(\mathcal{H}) \hookrightarrow \mathcal{B}_1(\mathcal{H}) \hookrightarrow \mathcal{B}_2(\mathcal{H}) \hookrightarrow \mathcal{B}_c(\mathcal{H}) \hookrightarrow \mathcal{B}(\mathcal{H}) \tag{2.15}$$

for the spaces of bounded linear operators on a separable Hilbert space $\mathcal{H}$ *over the field* $\mathbb{K}$ *with the chain of embeddings for the corresponding sequence spaces*

$$\ell_f(\mathbb{K}) \hookrightarrow \ell^1(\mathbb{K}) \hookrightarrow \ell^2(\mathbb{K}) \hookrightarrow c_0(\mathbb{K}) \hookrightarrow \ell^\infty(\mathbb{K}) \tag{2.16}$$

where $\ell_f(\mathbb{K})$ *denotes the space of terminating sequences and* $c_0(\mathbb{K})$ *the space of null sequences. And our results on the spectral representations of operators in* $\mathcal{B}_1(\mathcal{H})$*,*

$\mathcal{B}_2(\mathcal{H})$, *and* $\mathcal{B}_c(\mathcal{H})$ *indicate how these spaces are related to the sequence spaces* $\ell^1(\mathbb{K})$, $\ell^2(\mathbb{K})$, *and* $c_0(\mathbb{K})$.

For these sequence spaces it is well known that $\ell^\infty(\mathbb{K})$ *is the topological dual of* $\ell^1(\mathbb{K})$, $\ell^1(\mathbb{K})' \cong \ell^\infty(\mathbb{K})$, *and that* $\ell^1(\mathbb{K})$ *is the dual of* $c_0(\mathbb{K})$, $c_0(\mathbb{K})' \cong \ell^1(\mathbb{K})$, *as the counterpart of the last two results:* $\mathcal{B}_1(\mathcal{H})' \cong \mathcal{B}(\mathcal{H})$ *and* $\mathcal{B}_c(\mathcal{H})' \cong \mathcal{B}_1(\mathcal{H})$.

2.3 Related locally convex topologies on $\mathcal{B}(\mathcal{H})$

Recall that in Section 21.4 we introduced the weak and the strong operator topologies on $\mathcal{B}(\mathcal{H})$ as the topology of pointwise weak respectively pointwise norm convergence. In the study of operator algebras some further topologies play an important rôle. Here we restrict our discussion to the operator algebra $\mathcal{B}(\mathcal{H})$, respectively subalgebras of it. Recall also that in the second chapter we had learned how to define locally convex topologies on vector spaces in terms of suitable systems of seminorms. This approach we use here again. We begin by recalling the defining seminorms for the strong and the weak topology.

The **strong topology** on $\mathcal{B}(\mathcal{H})$ is defined by the system of seminorms p_x, $x \in \mathcal{H}$, with

$$p_x(A) = \|Ax\|, \quad A \in \mathcal{B}(\mathcal{H}).$$

Sometimes it is important to have a topology on $\mathcal{B}(\mathcal{H})$ with respect to which the involution * on $\mathcal{B}(\mathcal{H})$ is continuous. This is the case for the **strong* topology** defined by the system of seminorms p_x^*, $x \in \mathcal{H}$, with

$$p_x^*(A) = \sqrt{\|Ax\|^2 + \|A^*x\|^2}, \quad A \in \mathcal{B}(\mathcal{H}).$$

The **weak topology** on $\mathcal{B}(\mathcal{H})$ is defined by the system for seminorms $p_{x,y}$, $x, y \in \mathcal{H}$, with

$$p_{x,y}(A) = |\langle x, Ay \rangle|, \quad A \in \mathcal{B}(\mathcal{H}).$$

Similarly one defines the σ-weak and σ-strong topologies on $\mathcal{B}(\mathcal{H})$. Often these topologies are also called **ultraweak** respectively **ultrastrong** topology.

The **σ-strong topology** on $\mathcal{B}(\mathcal{H})$ is defined in terms of a system of seminorms $q = q_{\{e_n\}}$, $\{e_n\} \subset \mathcal{H}$, $\sum_n \|e_n\|^2 < \infty$, with

$$q(A) = \left(\sum_n \|Ae_n\|^2 \right)^{1/2}, \quad A \in \mathcal{B}(\mathcal{H}).$$

And the **σ-strong*** topology is defined by the system of seminorm q^*, q as above, with

$$q^*(A) = (q(A)^2 + q(A^*)^2)^{1/2}, \quad A \in \mathcal{B}(\mathcal{H}).$$

Next suppose that $\{e_n\}$ and $\{g_n\}$ are two sequences in $\mathcal{H}$ which satisfy $\sum_n \|e_n\|^2 < \infty$ and $\sum_n \|g_n\|^2 < \infty$. Then a continous linear functional $T = T_{\{e_n\},\{g_n\}}$ is welldefined on $\mathcal{B}(\mathcal{H})$ by (see Exercises)

$$T(A) = \sum_n \langle g_n, Ae_n \rangle, \quad A \in \mathcal{B}(\mathcal{H}). \tag{2.17}$$

Now the σ**-weak topology** on $\mathcal{B}(\mathcal{H})$ is defined by the system of seminorms p_T, T as above, by

$$p_T(A) = |T(A)| = |\sum_n \langle g_n, Ae_n \rangle|.$$

Using the finite rank operators $[e_n, g_n]$ introduced earlier we can form the operator $\hat{T} = \sum_n [e_n, g_n]$. For any two orthonormal systems $\{x_j\}$ and $\{y_j\}$ in $\mathcal{H}$ we estimate by Schwarz' and Bessel's inequalities

$$\sum_j |\langle x_j, [e_n, g_n] y_j \rangle| = \sum_j |\langle x_j, e_n \rangle \langle g_n, y_j \rangle| \leq \|e_n\| \, \|g_n\|$$

and thus

$$\sum_j |\langle x_j, \hat{T} y_j \rangle| \leq \sum_n \|e_n\| \, \|g_n\| < \infty.$$

Proposition 2.1.5 implies that $\hat{T}$ is a trace class operator on $\mathcal{H}$. In the Exercises we show that for all $A \in \mathcal{B}(\mathcal{H})$

$$\mathrm{Tr}(A\hat{T}) = \sum_n \langle g_n, Ae_n \rangle = T(A), \tag{2.18}$$

hence the functional T of (2.17) is represented as the trace of the trace class operator $\hat{T}$ multiplied by the argument of T.

According to Theorem 2.2.2 the Banach space dual of the space of compact operators $\mathcal{B}_c(\mathcal{H})$ is isometrically isomorphic to the space $\mathcal{B}_1(\mathcal{H})$ of trace class operators on $\mathcal{H}$ and according to Theorem 2.2.3 the Banach space dual of $\mathcal{B}_1(\mathcal{H})$ is isometrically isomorphic to the space $\mathcal{B}(\mathcal{H})$ of all bounded linear operators on $\mathcal{H}$.

Thus we can state:

> The σ**-weak** or **ultraweak topology** on $\mathcal{B}(\mathcal{H})$ is the weak*-topology from the identification of $\mathcal{B}(\mathcal{H})$ with the dual of $\mathcal{B}_1(\mathcal{H})$, i.e., the topology generated by the family of semi-norms $\{p_{\hat{T}} : \hat{T} \in \mathcal{B}_1(\mathcal{H})\}$ defined by $p_{\hat{T}}(A) = |\mathrm{Tr}(\hat{T}A)| = |T(A)|$ for $A \in \mathcal{B}(\mathcal{H})$.

It is easy to verify that the weak topology on $\mathcal{B}(\mathcal{H})$ is the dual topology $\sigma(\mathcal{B}(\mathcal{H}), \mathcal{B}_f(\mathcal{H}))$.

Recall also that nonnegative trace class operators are of the form $\hat{T} = \sum_n [e_n, e_n]$ with $\sum_n \|e_n\|^2 < \infty$. For $A \in \mathcal{B}(\mathcal{H})$ we find

$$\mathrm{Tr}(A^* A \hat{T}) = \sum_n \|Ae_n\|^2 = T(A^* A)$$

where T is the functional on $\mathcal{B}(\mathcal{H})$ which corresponds to $\hat{T}$ according to (2.18). Hence the defining seminorms q for the σ-strong topology are actually of the form

$$q(A) = (T(A^* A))^{1/2} = (\mathrm{Tr}(A^* A \hat{T}))^{1/2}.$$

From these definitions it is quite obvious how to compare these topologies on $\mathcal{B}(\mathcal{H})$: The σ-strong* topology is finer than the σ-strong topology which in turn is finer than the σ-weak topology. And certainly the σ-strong* topology is finer than the strong* topology and the σ-strong topology is finer than the strong topology which is finer than the weak topology. Finally the σ-weak topology is finer than the weak topology. Obviously the uniform or norm topology is finer than the σ-strong* topology. Nevertheless one has the following convenient result:

Lemma 2.3.1 *On the closed unit ball* $B = \{A \in \mathcal{B}(\mathcal{H}) : \|A\| \leq 1\}$ *the following topologies are the same:*

a) the weak and the σ-weak;

b) the strong and the σ-strong;

c) the strong and the σ-strong*.*

Proof. Since the proofs of these three statements are very similar we offer explicitly only the proof of b).

Clearly it suffices to show that for every neighborhood U of the origin for the σ-strong topology there is a neighborhood V of the origin for the strong toplogy such that $V \cap B \subset U \cap B$. Such a neighborhood U is of the form $U = \{A \in \mathcal{B}(\mathcal{H}) : q(A) < r\}$ with $r > 0$ and $q(A)^2 = \sum_n \|Ae_n\|^2$ for some sequence $e_n \in \mathcal{H}$ with $\sum_n \|e_n\|^2 < \infty$. Thus there is $m \in \mathbb{N}$ such that $\sum_{n=m+1}^{\infty} \|e_n\|^2 < r^2/2$. Define a neighborhood V of the origin for the strong topology by $V = \left\{A \in \mathcal{B}(\mathcal{H}) : p(A) < r/\sqrt{2}\right\}$ with the norm p given by $p(A)^2 = \sum_{n=1}^{m} \|Ae_n\|^2$. Now for $A \in V \cap B$ we estimate

$$q(A)^2 = \sum_{n=1}^{m} \|Ae_n\|^2 + \sum_{n=m+1}^{\infty} \|Ae_n\|^2 \leq \sum_{n=1}^{m} \|Ae_n\|^2 + \sum_{n=m+1}^{\infty} \|e_n\|^2 < r^2/2 + r^2/2 = r^2$$

and conclude $A \in U \cap B$. □

In addition continuity of linear functionals are the same within two groups of these topologies as the following theorem shows.

Theorem 2.3.2 *Suppose that $\mathcal{K} \subset \mathcal{B}(\mathcal{H})$ is a linear subspace which is σ-weakly closed. Then for every bounded linear functional T on $\mathcal{K}$ the following groups of equivalence statements hold.*

1) The following statements about T are equivalent:

(a) *T is of the form* $T(\cdot) = \sum_{j=1}^{m} \langle y_j, \cdot x_j \rangle$ *for some points* $x_j, y_j \in \mathcal{H}$*;*

(b) *T is weakly continuous;*

(c) *T is strongly continuous;*

(d) *T is strongly* continuous.*

2) *The following statements about T are equivalent (B is the closed unit ball in* $\mathcal{B}(\mathcal{H})$*):*

(a) *T is of the form* $T(\cdot) = \sum_{j=1}^{\infty} \langle y_j, \cdot x_j \rangle$ *for some sequences* $x_j, y_j \in \mathcal{H}$ *with* $\sum_j \|x_j\|^2 < \infty$ *and* $\sum_j \|y_j\|^2 < \infty$*;*

(b) *T is* σ*-weakly continuous;*

(c) *T is* σ*-strongly continuous;*

(d) *T is* σ*-strongly* continuous;*

(e) *T is weakly continuous on* $\mathcal{K} \cap B$ *;*

(f) *T is strongly continuous on* $\mathcal{K} \cap B$*;*

(g) *T is strongly* continuous on* $\mathcal{K} \cap B$*.*

Proof. Since the first group of equivalence statements is just a 'finite' variant of the second we do not prove it explicitly.

For the proof of 2) we proceed in the order a. $\Rightarrow$ b.$\Rightarrow$ c. $\Rightarrow$ d. $\Rightarrow$ a. If a. is assumed then

$$A \mapsto |T(A)| = |\sum_j \langle y_j, Ax_j \rangle|$$

is obviously a defining seminorm for the σ-weak topology. If we apply the Cauchy-Schwarz inequality twice this seminorm is estimated by

$$(\sum_j \|y_j\|^2)^{1/2}(\sum_j \|Ax_j\|^2)^{1/2}$$

which is a continuous seminorm for the σ-strong topology and thus T is also σ-strongly continuous. Another elementary estimate now proves d.

The only non-tivial part of the proof is the implication d. $\Rightarrow$ a. . If T is σ-strongly* continuous there is a sequences $\{x_j\}$ with $\sum_j \|x_j\|^2 < \infty$ such that for all $A \in \mathcal{K}$

$$|T(A)|^2 \leq \sum_{j=1}^{\infty}(\|Ax_j\|^2 + \|A^*x_j\| |^2) . \tag{2.19}$$

Form the direct sum Hilbert space

$$\tilde{\mathcal{H}} = \bigoplus_{j=1}^{\infty}(\mathcal{H}_j \oplus \mathcal{H}'_j) = \ell^2(\mathcal{H} \oplus \mathcal{H}')$$

where $\mathcal{H}'_j$ is the dual of $\mathcal{H}_j = \mathcal{H}$ for all $j \in \mathbb{N}$. For $A \in \mathcal{B}(\mathcal{H})$ define an operator $\tilde{A}$ on $\tilde{\mathcal{H}}$ by setting for $\tilde{y} \in \tilde{\mathcal{H}}$ with components $y_j \oplus y'_j \in \mathcal{H}_j \oplus \mathcal{H}'_j$

$$(\tilde{A}\tilde{y})_j = Ay_j \oplus (A^*y_j)', \quad j \in \mathbb{N}.$$

A straightforward estimate shows

$$\|\tilde{A}\tilde{y}\|_{\tilde{\mathcal{H}}} \leq (\|A\|^2 + \|A^*\|^2)^{1/2} \|\tilde{y}\|_{\tilde{\mathcal{H}}} .$$

On the subspace

$$\tilde{\mathcal{K}} = \{\tilde{A}\tilde{x} : A \in \mathcal{K}\}$$

of $\tilde{\mathcal{H}}$ where $\tilde{x}$ is defined by the sequence $\{x_j\}$ of the estimate (2.19), define the map $\tilde{T}$ by setting

$$\tilde{T}(\tilde{A}\tilde{x}) = T(A).$$

By (2.19) it follows that $\tilde{T}$ is a welldefined bounded linear map $\tilde{\mathcal{K}} \to \mathbb{K}$ (recall that $\mathcal{H}_j \to \mathcal{H}'_j$ is antilinear). Theorems 15.3.2 (extension theorem) and 15.3.1 (Riesz-Fréchet) imply that there is an element $\tilde{y}$ in the closure of the subspace $\tilde{\mathcal{K}}$ in $\tilde{\mathcal{H}}$ such that

$$\tilde{T}(\tilde{A}\tilde{x}) = \langle \tilde{y}, \tilde{A}\tilde{x} \rangle_{\tilde{\mathcal{H}}}$$

for all $A \in \mathcal{K}$, thus by expanding the inner product of $\tilde{\mathcal{H}}$

$$T(A) = \tilde{T}(\tilde{A}\tilde{x}) = \sum_j (\langle y_j, Ax_j \rangle_{\mathcal{H}} + \langle y_j', (A^* x_j)' \rangle_{\mathcal{H}'}) = \sum_j (\langle y_j, Ax_j \rangle_{\mathcal{H}} + \langle x_j, Ay_j \rangle_{\mathcal{H}})$$

and therefore T is of form given in statement a..

The remaining part of the proof follows with the help of Lemma 2.3.1 and Corollary **??** which says that a linear functional is continuous if, and only if, it is continuous at the origin (see also the Exercises). □

Remark 2.3.3 *A considerably more comprehensive list of conditions under which these various locally convex topologies on $\mathcal{B}(\mathcal{H})$ agree is available in Chapter II of 24.*

2.4 Partial Trace and Schmidt decomposition in separable Hilbert spaces

2.4.1 Partial Trace

The first guess for defining the partial trace in the case of infinite dimensional Hilbert spaces $\mathcal{H}_j$ would be, in analogy to the the case of finite dimensional Hilbert spaces, to start with the matrix representation of $A \in \mathcal{B}_1(\mathcal{H}_1 \otimes \mathcal{H}_2)$ with respect to an orthonormal basis $\{e_j \otimes f_k\}$ of $\mathcal{H}_1 \otimes \mathcal{H}_2$ and to calculate the usual sums with respect to one of the ONBs $\{e_j\}$ respectively $\{f_k\}$. However infinite sums might be divergent, and we have found no useful way to express the fact that A is of trace class in terms of properties of the matrix entries

$$A_{j_1k_1;j_2k_2} \quad j_1, j_2, k_1, k_2 \in \mathbb{N}.$$

But such a procedure can be imitated by introducing a suitable quadratic form and investigate its properties (see 3).

Theorem 2.4.1 (Existence, definition and basic properties of partial trace) *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two separable complex Hilbert spaces. Then there is a linear map*

$$T : \mathcal{B}_1(\mathcal{H}_1 \otimes \mathcal{H}_2) \longrightarrow \mathcal{B}_1(\mathcal{H}_1)$$

from the space of trace class operators on $\mathcal{H}_1 \otimes \mathcal{H}_2$ into the space of trace class operators on $\mathcal{H}_1$ which is continuous with respect to the trace norm. It has the following properties

$$T(A_1 \otimes A_2) = A_1 \mathrm{Tr}_{\mathcal{H}_2}(\mathrm{A}_2) \quad \textit{for all } \mathrm{A_i} \in \mathcal{B}_1(\mathcal{H}_\mathrm{i}), \mathrm{i} = 1,2; \tag{2.20}$$

$$\mathrm{Tr}_{\mathcal{H}_1}(\mathrm{T}(\mathrm{A})) = \mathrm{Tr}_{\mathcal{H}_1 \otimes \mathcal{H}_2}(\mathrm{A}) \quad \textit{for all } \mathrm{A} \in \mathcal{B}_1(\mathcal{H}_1 \otimes \mathcal{H}_2); \tag{2.21}$$

$$T((A_1 \otimes I_2)A) = A_1 T(A) \quad \textit{for all } A_1 \in \mathcal{B}(\mathcal{H}_1), \textit{and all } A \in \mathcal{B}_1(\mathcal{H}_1 \otimes \mathcal{H}_2); \tag{2.22}$$

where I_2 denotes the identity operator on $\mathcal{H}_2$ and $\mathcal{B}(\mathcal{H}_1)$ the space of bounded linear operators on $\mathcal{H}_1$.

On the basis of Property (2.20) the map T is usually denoted by $\mathrm{Tr}_{\mathcal{H}_2}$ and called the **partial trace with respect to $\mathcal{H}_2$**. Later in Proposition 2.4.3 an enhanced characterization of the partial trace will be offered. Actually this map T is surjective (in Formula (2.20) take any fixed $A_2 \in \mathcal{B}_1(\mathcal{H}_2)$ with $\mathrm{Tr}(A_2) = 1$.)

Proof. Let $\{f_j; j \in \mathbb{N}\}$ be an orthonormal basis of $\mathcal{H}_2$. For a given operator $A \in \mathcal{B}_1(\mathcal{H}_1 \otimes \mathcal{H}_2)$ define a sesquilinear form Q_A on $\mathcal{H}_1$ by setting for $u, v \in \mathcal{H}_1$

$$Q_A(u,v) = \sum_{j=1}^{\infty} \langle u \otimes f_j, A(v \otimes f_j)\rangle_{1\otimes 2}. \tag{2.23}$$

By inserting the spectral representation (2.5) for $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ we can write this as

$$Q_A(u,v) = \sum_{j=1}^{\infty} \sum_{n=1}^{\infty} \lambda_n \langle u \otimes f_j, x_n\rangle_{1\otimes 2} \langle e_n, v \otimes f_j\rangle_{1\otimes 2}$$

For $u, v \in \mathcal{H}_1$ with $\|u\| = \|v\| = 1$ we know that $\{u \otimes f_j\}$ and $\{v \otimes f_j\}$ are ONS in $\mathcal{H}_1 \otimes \mathcal{H}_2$ and thus we can estimate, using first Schwarz' inequality and then Bessel's inequality for these ONS,

$$|Q_A(u,v)| \leq \sum_n |\lambda_n| \sum_j |\langle u \otimes f_j, x_n\rangle_{1\otimes 2} \langle e_n, v \otimes f_j\rangle_{1\otimes 2}|$$

$$\leq \sum_n |\lambda_n| \, \|e_n\| \, \|x_n\| = \sum_n |\lambda_n| = \|A\|_1$$

This implies for general $u, v \in \mathcal{H}_1$

$$|Q_A(u,v)| \leq \|A\|_1 \, \|u\|_1 \, \|v\|_1 \, ,$$

and thus the sesquilinear form Q_A is well defined and continuous. Therefore the representation formula for continuous sesqulinear forms applies and assures the existence of a unique bounded linear operator $T(A)$ on $\mathcal{H}_1$ such that

$$Q_A(u,v) = \langle u, T(A)v\rangle_1 \quad \text{for all } u, v \in \mathcal{H}_1 \tag{2.24}$$

and $\|T(A)\| \leq \|A\|_1$.

In order to show $T(A) \in \mathcal{B}_1(\mathcal{H}_1)$ we use the characterization of trace class operators as given in Proposition 2.1.5 and estimate $S_3(T(A))$ by inserting the spectral representation (2.5) for $A \in \mathcal{B}_1(\mathcal{H}_1 \otimes \mathcal{H}_2)$. To this end take any orthonormal sequences $\{u_n\}$ and

$\{v_m\}$ in $\mathcal{H}_1$. Since then $\{u_n \otimes f_j\}$ and $\{v_m \otimes f_j\}$ are orthonormal sequences in $\mathcal{H}_1 \otimes \mathcal{H}_2$ we can estimate as follows, again applying first Schwarz' and then Bessel's inequality:

$$\begin{aligned}\sum_k |\langle u_n, T(A) v_k\rangle_1| &= \sum_k |\sum_{n,j} \lambda_n \langle u_k \otimes f_j, x_n\rangle_{1\otimes 2} \langle e_n, v_k \otimes f_j\rangle_{1\otimes 2}| \\ &\leq \sum_n |\lambda_n| \sum_{k,j} |\langle u_k \otimes f_j, x_n\rangle_{1\otimes 2} \langle e_n, v_k \otimes f_j\rangle_{1\otimes 2}| \\ &\leq \sum_k |\lambda_n| \, \|e_n\|_{1\otimes 2} \, \|x_n\|_{1\otimes 2} = \sum_k |\lambda_n| = \|A\|_1 < \infty.\end{aligned}$$

We conclude $T(A) \in \mathcal{B}_1(\mathcal{H}_1)$ and

$$\|T(A)\|_1 \leq \|A\|_1 . \tag{2.25}$$

The above definition of $T(A)$ is based on the choice of an orthonormal basis $\{f_j; j \in \mathbb{N}\}$. However, as in the case of a trace, the value of $T(A)$ does actually not depend on the basis which is used to calculate it. Suppose that $\{h_j; j \in \mathbb{N}\}$ is another orthonormal basis of $\mathcal{H}_2$. Express the f_j in terms of the new basis, i.e.,

$$f_j = \sum_\nu u_{j\nu} h_\nu, \quad u_{j\nu} \in \mathbb{C}.$$

Since the transition from one orthonormal basis to another is given by a unitary operator, one knows $\sum_j \overline{u}_{j\nu} u_{j\mu} = \delta_{\nu\mu}$. Now calculate for $u, v \in \mathcal{H}_1$

$$\begin{aligned}\sum_j \langle u \otimes f_j, A(v \otimes f_j)\rangle_{1\otimes 2} &= \sum_{j,\nu,\mu} \overline{u}_{j\nu} u_{j\mu} \langle u \otimes h_\nu, A(v \otimes h_\mu)\rangle_{1\otimes 2} \\ &= \sum_{\nu,\mu} \sum_j \overline{u}_{j\nu} u_{j\mu} \langle u \otimes h_\nu, A(v \otimes h_\mu)\rangle_{1\otimes 2} = \sum_\nu \langle u \otimes h_\nu, A(v \otimes h_\nu)\rangle_{1\otimes 2}.\end{aligned}$$

Therefore the sesquilinear form Q_A does not depend on the orthonormal basis which is used to calculate it. We conclude that the definition of $T(A)$ does not depend on the basis.

Equations (2.23) and (2.24) imply immediately that our map T is linear and thus by (2.25) continuity with respect to the trace norms follows.

Next we verify the basic properties (2.20), (2.21), and (2.22). For $A_i \in \mathcal{B}_1(\mathcal{H}_i)$, $i = 1, 2$ one finds by applying the definitions for all $u, v \in \mathcal{H}_1$

$$\begin{aligned}\langle u, T(A_1 \otimes A_2) v\rangle_1 &= \sum_j \langle u \otimes f_j, (A_1 \otimes A_2)(v \otimes f_j)\rangle_{1\otimes 2} \\ &= \sum_j \langle u, A_1 v\rangle_1 \langle f_j, A_2 f_j\rangle_2 = \langle u, A_1 v\rangle_1 \mathrm{Tr}_{\mathcal{H}_2}(\mathrm{A}_2),\end{aligned}$$

hence $T(A_1 \otimes A_2) = A_1 \mathrm{Tr}_{\mathcal{H}_2}(\mathrm{A}_2)$, i.e., (2.20) holds.

For $A \in \mathcal{B}_1(\mathcal{H}_1 \otimes \mathcal{H}_2)$ we calculate, using an orthonormal basis $\{e_i; i \in \mathbb{N}\}$ of $\mathcal{H}_1$

$$\mathrm{Tr}_{\mathcal{H}_1}(T(A)) = \sum_i \langle e_i, T(A)e_i\rangle_1 = \sum_i \sum_j \langle e_i \otimes f_j, Ae_i \otimes f_j\rangle_{1\otimes 2} = \mathrm{Tr}_{\mathcal{H}_1 \otimes \mathcal{H}_2}(A)$$

and find that (2.21) holds.

Finally take any bounded linear operator A_1 on $\mathcal{H}_1$, any $A \in \mathcal{B}_1(\mathcal{H}_1 \otimes \mathcal{H}_2)$, and any vectors $u, v \in \mathcal{H}_1$. Our definition gives

$$\begin{aligned}\langle u, T((A_1 \otimes I_2)A)v\rangle_1 &= \sum_j \langle u \otimes f_j, (A_1 \otimes I_2)A(v \otimes f_j)\rangle_{1\otimes 2} \\ &= \sum_j \langle A_1^* u \otimes f_j, A(v \otimes f_j)\rangle_{1\otimes 2} = \langle A_1^* u, T(A)v\rangle_1 = \langle u, A_1 T(A)v\rangle_1,\end{aligned}$$

and thus $T((A_1 \otimes I_2)A) = A_1 T(A)$, i.e., (2.22) is established. □

Corollary 2.4.2 *For all bounded linear operators A_1 on $\mathcal{H}_1$ and all $A \in \mathcal{B}_1(\mathcal{H}_1 \otimes \mathcal{H}_2)$ one has*

$$\mathrm{Tr}_{\mathcal{H}_1 \otimes \mathcal{H}_2}((A_1 \otimes I_2)A) = \mathrm{Tr}_{\mathcal{H}_1}(A_1 \mathrm{Tr}_{\mathcal{H}_2}(\mathrm{A})). \tag{2.26}$$

Proof. Apply first (2.21) and then (2.22) and observe that the product of a trace class operator with a bounded linear operator is again a trace class operator (see Theorem 2.1.6). □

Proposition 2.4.3 (partial trace characterization) *Suppose that $\mathcal{H}_1$ and $\mathcal{H}_2$ are two separable Hilbert spaces and that a linear map $L : \mathcal{B}_1(\mathcal{H}_1 \otimes \mathcal{H}_2) \longrightarrow \mathcal{B}_1(\mathcal{H}_1)$ satisfies*

$$\mathrm{Tr}_{\mathcal{H}_1}(PL(A)) = \mathrm{Tr}_{\mathcal{H}_1 \otimes \mathcal{H}_2}(P \otimes I_2 A) \tag{2.27}$$

for all finite rank orthogonal projectors P on $\mathcal{H}_1$ and all $A \in \mathcal{B}_1(\mathcal{H}_1 \otimes \mathcal{H}_2)$. Then L is the partial trace with respect to $\mathcal{H}_2$:

$$L(A) = \mathrm{Tr}_{\mathcal{H}_2}(A). \tag{2.28}$$

Proof. By taking linear combinations of (2.27) of finite rank projectors P_j we conclude that

$$\mathrm{Tr}_{\mathcal{H}_1}(BL(A)) = \mathrm{Tr}_{\mathcal{H}_1 \otimes \mathcal{H}_2}((B \otimes I_2)A)$$

holds for all $B \in \mathcal{B}_f(\mathcal{H}_1)$ and all $A \in \mathcal{B}_1(\mathcal{H}_1 \otimes \mathcal{H}_2)$. Hence by Equation 2.26 we find

$$\mathrm{Tr}_{\mathcal{H}_1}(BL(A)) = \mathrm{Tr}_{\mathcal{H}_1}(B\mathrm{Tr}_{\mathcal{H}_2}(A))$$

or, observing Corollary 2.2.1 and taking the definition of the inner product on $\mathcal{B}_2(\mathcal{H})$ in (2.3) into account,

$$\langle B, L(A)\rangle_2 = \langle B, \mathrm{Tr}_{\mathcal{H}_2}(\mathrm{A})\rangle_2$$

all $B \in \mathcal{B}_f(\mathcal{H}_1)$ and all $A \in \mathcal{B}_1(\mathcal{H}_1 \otimes \mathcal{H}_1)$. Since $\mathcal{B}_f(\mathcal{H}_1)$ is dense in $\mathcal{B}_2(\mathcal{H}_1)$ we conclude. □

2.4.2 Schmidt decomposition

The elements of $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ can be described explicitly in terms of orthonormal bases of $\mathcal{H}_i$: Suppose e_j, $j \in \mathbb{N}$ is an orthonormal basis of $\mathcal{H}_1$ and f_j, $j \in \mathbb{N}$ is an orthonormal basis of $\mathcal{H}_2$. Then every element $x \in \mathcal{H}$ is of the form (see for instance 4)

$$x = \sum_{i,j=1}^{\infty} c_{i,j} e_i \otimes f_j, \quad c_{i,j} \in \mathbb{C}, \quad \sum_{i,j=1}^{\infty} |c_{i,j}|^2 = \|x\|^2. \tag{2.29}$$

However in the discussion of entanglement in quantum physics and quantum information theory it has become the standard to use the *Schmidt representation* of vectors in $\mathcal{H}$ which reduces the double sum in (2.29) to a simple bi-orthogonal sum.

Theorem 2.4.4 (Schmidt decomposition) *For every $x \in \mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ there are nonnegative numbers p_n and orthonormal bases $g_n; n \in \mathbb{N}$, of $\mathcal{H}_1$ and h_n, $n \in \mathbb{N}$, of $\mathcal{H}_2$ such that*

$$x = \sum_{n=1}^{\infty} p_n g_n \otimes h_n, \quad \sum_{n=1}^{\infty} p_n^2 = \|x\|^2. \tag{2.30}$$

Proof. We use the standard isomorphism I between the Hilbert tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ and the space $\mathcal{L}_{HS}(\mathcal{H}_1; \mathcal{H}_2)$ of Hilbert-Schmidt operators $\mathcal{H}_1 \longrightarrow \mathcal{H}_2$. In the notation of physicists I is given by $I(x) = \sum_{i,j=1}^{\infty} c_{i,j} |f_j\rangle\langle e_i|$, i.e., for all $y \in \mathcal{H}_1$ we have

$$I(x)(y) = \sum_{i,j=1}^{\infty} c_{i,j} \langle e_i, y\rangle_1 f_j$$

where $\langle\cdot,\cdot\rangle_1$ denotes the inner product of $\mathcal{H}_1$ and where x is given by (2.29). It is easily seen that $I(x)$ is a well defined bounded linear operator $\mathcal{H}_1 \longrightarrow \mathcal{H}_2$. Hence $I(x)^* I(x)$ is a bounded linear operator $\mathcal{H}_1 \longrightarrow \mathcal{H}_1$ which is of trace class since

$$\mathrm{Tr}_{\mathcal{H}_1}(\mathrm{I}(\mathrm{x})^*\mathrm{I}(\mathrm{x})) = \sum_{\mathrm{i}=1}^{\infty} \langle \mathrm{I}(\mathrm{x})\mathrm{e_i}, \mathrm{I}(\mathrm{x})\mathrm{e_i}\rangle_2 = \sum_{\mathrm{i,j}=1}^{\infty} |\mathrm{c_{i,j}}|^2 = \|\mathrm{x}\|^2.$$

Thus $I(x)$ is a Hilbert Schmidt operator with norm

$$\|I(x)\|_2 = +\sqrt{\mathrm{Tr}_{\mathcal{H}_1}(\mathrm{I}(\mathrm{x})^*\mathrm{I}(\mathrm{x}))} = \|x\|. \tag{2.31}$$

Since $I(x)^*I(x)$ is a positive trace class operator on $\mathcal{H}_1$, it is of the form

$$I(x)^*I(x) = \sum_{n=1}^{\infty} \lambda_n P_{g_n}, \quad \sum_{n=1}^{\infty} \lambda_n = \|x\|^2, \tag{2.32}$$

where $P_{g_n} = |g_n\rangle\langle g_n|$ is the orthogonal projector onto the subspace spanned by the element g_n of an orthonormal basis g_n, $n \in \mathbb{N}$, of $\mathcal{H}_1$.

This spectral representation allows easily to calculate the square root of the operator $I(x)^*I(x)$:

$$|I(x)| \stackrel{\text{def}}{=} +\sqrt{I(x)^*I(x))} = \sum_{n=1}^{\infty} \sqrt{\lambda_n} P_{g_n}. \tag{2.33}$$

This prepares for the polar decomposition (see for instance 4) of the operator $I(x) : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$, according to which this operator can be written as

$$I(x) = U|I(x)|, \quad U = \text{ partial isometry } \mathcal{H}_1 \longrightarrow \mathcal{H}_2, \tag{2.34}$$

i.e., U is an isometry from $(\ker I(x))^\perp \subset \mathcal{H}_1$ onto $\overline{\mathrm{ran}\, \mathrm{I(x)}} \subset \mathcal{H}_2$.

Finally denote by $h_n, n \in \mathbb{N}$, the orthonormal system obtained from the basis g_n, $n \in \mathbb{N}$, under this partial isometry, $h_n = Ug_n$. Hence, from (2.33) and (2.34) we get

$$I(x) = \sum_{n=1}^{\infty} \sqrt{\lambda_n}\, |h_n\rangle\langle g_n|.$$

If we identify p_n with $\sqrt{\lambda_n}$ and if we apply I^{-1} to this identity, then Equation (2.30) follows. □

2.5 Some applications in Quantum Mechanics

Remark 2.5.1 *In the case of concrete Hilbert spaces the trace can often be evaluated explicitly without much effort, usually easier than for instance the operator norm.*

Consider the Hilbert–Schmidt integral operator K in $L^2(\mathbb{R}^n)$ discussed earlier. It is defined in terms of a kernel $k \in L^2(\mathbb{R}^n \times \mathbb{R}^n)$ by

$$K\psi(x) = \int_{\mathbb{R}^n} k(x,y)\psi(y)\boldsymbol{D}y \qquad \forall\, \psi \in L^2(\mathbb{R}^n).$$

In the Exercises we show that

$$Tr(K^*K) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \overline{k(x,y)}k(x,y)\boldsymbol{D}x\boldsymbol{D}y.$$

A special class of trace class operators is of great importance for quantum mechanics, which we briefly mention.

Definition 2.5.2 *A* **density matrix** *or* **statistical operator** *W on a separable Hilbert space $\mathcal{H}$ is a trace class operator which is symmetric ($W^* = W$), positive ($\langle x, Wx\rangle \geq 0$ for all $x \in \mathcal{H}$), and normalized ($\mathrm{Tr}\, W = 1$).*

Note that in a complex Hilbert space symmetry is implied by positivity. In quantum mechanics density matrices are usually denoted by ρ. Density matrices can be characterized explicitly.

Theorem 2.5.3 *A bounded linear operator W on a separable Hilbert space $\mathcal{H}$ is a density matrix if, and only if, there are a sequence of nonnegative numbers $\rho_n \geq 0$*

with $\sum_{n=1}^{\infty} \rho_n = 1$ *and an orthonormal basis* $\{e_n : n \in \mathbb{N}\}$ *of* $\mathcal{H}$ *such that for all* $x \in \mathcal{H}$,

$$Wx = \sum_{n=1}^{\infty} \rho_n \langle e_n, x \rangle e_n, \tag{2.35}$$

i.e., $W = \sum_{n=1}^{\infty} \rho_n P_{e_n}$, P_{e_n} *=projector onto the subspace* $\mathbb{K} e_n$.

Proof. Using the spectral representation (2.1.9) of trace class operators the proof is straight forward and is left as an exercise. □

The results of this chapter have important applications in quantum mechanics, but also in other areas. We mention, respectively sketch, some of these applications briefly.

We begin with a reminder of some of the basic principles of quantum mechanics (see for instance 14, 12).

1. The *states* of a quantum mechanical system are described in terms of density matrices on a separable complex Hilbert space $\mathcal{H}$.

2. The *observables* of the systems are represented by self-adjoint operators in $\mathcal{H}$.

3. The *mean value* or *expectation value* of an observable a in a state z is equal to the expectation value $E(A, W)$ of the corresponding operators in $\mathcal{H}$; if the

self-adjoint operator A represents the observable a and the density matrix W represents the state z, this means that

$$m(a,z) = E(A,W) = \mathrm{Tr}(AW).$$

Naturally, the mean value $m(a,z)$ is considered as the mean value of the results of a measurement procedure. Here we have to assume that AW is a trace class operator, reflecting the fact that not all observables can be measured in all states.

4. Examples of density matrices W are projectors P_e on $\mathcal{H}$, $e \in \mathcal{H}$, $\|e\| = 1$, i.e., $Wx = \langle e, x \rangle e$. Such states are called *vector states* and e the representing vector. Then clearly $E(A, P_e) = \langle e, Ae \rangle = \mathrm{Tr}(P_e A)$.

5. *Convex combinations* of states, i.e., $\sum_{j=1}^{n} \lambda_j W_j$ of states W_j are again states (here $\lambda_j \geq 0$ for all j and $\sum_{j=1}^{n} \lambda_j = 1$). Those states which can not be represented as nontrivial convex combinations of other states are called *extremal* or *pure* states. Under quite general conditions one can prove: There are extremal states and the set of all convex combinations of pure states is dense in the space of all states (Theorem of Krein–Milman, 22, not discussed here).

Thus we learn, that and how, projectors and density matrices enter in quantum mechanics.

Next we discuss a basic application of Stone's Theorem 5 on groups of unitary operators. As we had argued earlier, the Hilbert space of an elementary localizable particle in one dimension is the separable Hilbert space $L^2(\mathbb{R})$. The translation of elements $f \in L^2(\mathbb{R})$ is described by the unitary operators $U(a)$, $a \in \mathbb{R}$: $(U(a)f)(x) = f_a(x) = f(x-a)$. It is not difficult to show that this one-parameter group of unitary operators acts strongly continuous on $L^2(\mathbb{R})$: One shows $\lim_{a\to 0} \|f_a - f\|_2 = 0$. Now Stone's theorem applies. It says that this group is generated by a self-adjoint operator P which is defined on the domain

$$D = \left\{ f \in L^2(\mathbb{R}) : \lim_{a\to 0} \frac{1}{a}(f_a - f) \quad \text{exists in } L^2(\mathbb{R}) \right\}$$

by

$$Pf = \frac{1}{i} \lim_{a\to 0} \frac{1}{a}(f_a - f) \qquad \forall f \in D.$$

The domain D is known to be $D = W^1(\mathbb{R}) \equiv \{ f \in L^2(\mathbb{R}) : f' \in L^2(\mathbb{R}) \}$ and clearly $Pf = -\mathrm{i}f' = -\mathrm{i}\frac{Df}{Dx}$. This operator P represents the momentum of the particle which is consistent with the fact that P generates the translations:

$$U(a) = \mathrm{e}^{-\mathrm{iaP}}.$$

As an illustration of the use of trace class operators and the trace functional we discuss a general form of the *Heisenberg uncertainty principle*. Given a density matrix W on a separable Hilbert space $\mathcal{H}$, introduce the set

$$O_W = \{A \in \mathcal{B}(\mathcal{H}) : A^* AW \in \mathcal{B}_1(\mathcal{H})\}$$

and a functional on $O_W \times O_W$,

$$(A, B) \mapsto \langle A, B \rangle_W = \mathrm{Tr}(A^* BW).$$

One shows (see Exercises) that this is a sesquilinear form on O_W which is positive semi-definite ($\langle A, A \rangle_W \geq 0$), hence the Cauchy–Schwarz inequality applies, i.e.,

$$|\langle A, B \rangle_W| \leq \sqrt{\langle A, A \rangle_W} \sqrt{\langle B, B \rangle_W} \qquad \forall\, A, B \in O_W.$$

Now consider two self-adjoint operators such that all the operators AAW, BBW, AW, BW, ABW, BAW are of trace class. Then the following quantities are well defined:

$$\overline{A} = A - \langle A \rangle_W I, \qquad \overline{B} = B - \langle B \rangle_W I$$

and then

$$\triangle_W(A) = \sqrt{\mathrm{Tr}(\overline{A}\,\overline{A}W)} = \sqrt{\mathrm{Tr}(A^2W) - \langle A\rangle_W^2},$$
$$\triangle_W(B) = \sqrt{\mathrm{Tr}(\overline{B}\,\overline{B}W)} = \sqrt{\mathrm{Tr}(B^2W) - \langle B\rangle_W^2}.$$

The quantity $\triangle_W(A)$ is called the *uncertainty* of the observable 'A' in the state 'W'. Next calculate the expectation value of the commutator $[A,B] = AB - BA$. One finds

$$\mathrm{Tr}([A,B]W) = \mathrm{Tr}([\overline{A},\overline{B}]W) = \mathrm{Tr}(\overline{A}\,\overline{B}W) - \mathrm{Tr}(\overline{B}\,\overline{A}W) = \langle\overline{A},\overline{B}\rangle_W - \langle\overline{B},\overline{A}\rangle_W$$

and by the above inequality this expectation value is bounded by the product of the uncertainties:

$$|\mathrm{Tr}([A,B]W)| \le |\langle\overline{A},\overline{B}\rangle_W + |\langle\overline{B},\overline{A}\rangle_W| \le \triangle_W(A)\triangle_W(B) + \triangle_W(B)\triangle_W(A).$$

Usually this estimate of the expectation value of the commutator in terms of the uncertainties is written as

$$\frac{1}{2}|\mathrm{Tr}([A,B]W)| \le \triangle_W(A)\triangle_W(B)$$

and called the *Heisenberg uncertainty relations* (for the 'observables' A, B).

Actually in quantum mechanics many observables are represented by unbounded self-adjoint operators. Then the above calculations do not apply directly and thus typically they are not done for a general density matrix as above but for pure states only. Originally they were formulated by Heisenberg corresponding to the observables of the position and the momentum, represented by the self-adjoint operators Q and P with the commutator $[Q,P] \subseteq iI$ and thus on suitable pure states ψ the famous version

$$\frac{1}{2} \leq \triangle_{\psi}(Q)\triangle_{\psi}(P)$$

of these uncertainty relations follows.

2.6 Exercises

1. Using Theorem 2.1.9 determine the form of the adjoint of a trace class operator A on $\mathcal{H}$ explicitly.

2. For a Hilbert–Schmidt operator K with kernel $k \in L^2(\mathbb{R}^n \times \mathbb{R}^n)$ show that

$$Tr(K^*K) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \overline{k(x,y)} k(x,y) \boldsymbol{D}x\boldsymbol{D}y.$$

3. Prove that statements e. - g. in the second part of Theorem 2.3.2 are equivalent to the corresponding statements b. - d. .

4. Prove the characterization (2.35) of a density matrix W.

   Hints: One can use $W^* = W = |W| = \sqrt{W^*W}$ and the explicit representation of the adjoint of a trace class operator (see the previous problem).

5. Show: A density matrix W on a Hilbert space $\mathcal{H}$ represents a vector state, i.e., can be written as the projector P_ψ onto the subspace generated by a vector $\psi \in \mathcal{H}$ if, and only if, $W^2 = W$.

6. Show: If a bounded linear operator A has the representation (2.5), then its absolute value is given by

$$|A|x = \sum_n |\lambda_n| \langle e_n, x \rangle e_n, \quad x \in \mathcal{H}.$$

7. Prove that (2.17) defines a continuous linear functional on $\mathcal{B}(\mathcal{H})$, under the assumptioned stated with this formula.

8. Prove Formula (2.18).

Chapter 3

Operator algebras and positive mappings

Abstract: In various parts of quantum physics positive mappings play a fundamental rôle. These mappings are defined on some involutive algebra (often of operators on some Hilbert space). By definition a positive mapping sends positive elements of its domain to positive elements of the target space. Thus first several characterizations of positive elements in an involutive normed algebra are given. Two types of positive mappings are considered: Positive linear functionals which have values in $\mathbb{C}$ and completely positive mappings which have values in some other involutive algebra. For the structural analysis of positive mappings the concept of a representation of an involutive algebra is needed. This is introduced and in the case of the involutive al-

gebra $\mathcal{B}(\mathcal{H})$ of bounded linear operators on a Hilbert space $\mathcal{H}$ the general form of its representations is determined (Naimark's theorem). The structure of positive linear functionals on an involutive algebra with unit is presented in the Gelfand-Naimark-Segal (GNS)-construction. Positive linear functionals f on an involutive algebra $\mathcal{A}$ with unit I such that $f(I) = 1$ are called states. On a weakly closed subalgebra $\mathcal{A}$ of $\mathcal{B}(\mathcal{H})$ special states are of the form $f(A) = \mathrm{Tr}(AW)$ where W is a density matrix on $\mathcal{H}$. These states are characterized in terms of an additional continuity condition (normality, complete additivity). and are called normal states. The Stinespring factorization theorem gives the general form of a completely positive map between C*-algebras. When this result is combined with Naimark's theorem of representations of $\mathcal{B}(\mathcal{H})$ it allows to determine the general form of completely positive mappings on $\mathcal{B}(\mathcal{H})$ in more detail.

3.1 Representations of C*-algebras

In quantum mechanics, in local quantum field theory (Haag), in the functional approach to relativistic quantum field theory (Garding-Wightman), and in quantum information theory positive functionals and completely positive mappings play a fundamental rôle, mainly in connection with the mathematical description of ‘states’ of quantum systems and their manipulation. In this chapter we present the most important structural results for positive linear functionals and completely positive maps, namely the Gelfand-Naimark Segal representation for positive linear functionals (on C*-algebras) and the Stinespring factorization theorem. The natural mathematical framework for these results is the theory of abstract C*-algebras and we formulate these results in this framework, but in our proofs we consider only the cases of C*-algebras of operators on Hilbert spaces. This allows to use some simplification in the characterization of positive elements in these algebras. For the general case we refer to the literature.

Definition 3.1.1 *Let $\mathcal{A}$ be an algebra over the field $\mathbb{C}$. If $\mathcal{A}$ admits an* **involution** * *which is compatible with the algebraic structure of $\mathcal{A}$, i.e., a mapping $a \mapsto a^*$ such that for all $a, b \in \mathcal{A}$ the following holds:*

$$(a^*)^* = a, \qquad (\lambda a)^* = \overline{\lambda} a^*,\ \lambda \in \mathbb{C}$$
$$(a+b)^* = a^* + b^* \qquad (ab)^* = b^* a^*$$

$\mathcal{A}$ is called an **involutive algebra** *or a* ***-algebra**. *If $\mathcal{A}$ admits a norm $\|\cdot\|$ under*

which $\mathcal{A}$ is a Banach space such that

$$\|ab\| \leq \|a\| \, \|b\| \quad \textit{for all} \quad a,b \in \mathcal{A}$$

then $\mathcal{A}$ is called a **Banach algebra**. *If a Banach algebra $\mathcal{A}$ has an involution * for which the norm satifies $\|a^*\| = \|a\|$ for all $a \in \mathcal{A}$ such that $\mathcal{A}$ is *-algebra then $\mathcal{A}$ is called an* **involutive Banach algebra** *or a* **Banach *-algebra**.

If in addition the norm of an involutive Banach algebra satisfies the conditon

$$\|a^*a\| = \|aa^*\| = \|a\|^2 \quad \textit{for all } a \in \mathcal{A}$$

it is called a **C*-algebra**.

Often an involutive Banach algebra or a C*-algebra contains a **unit** I. Then we assume that $\|I\| = 1$.

Definition 3.1.2 *Suppose that $\mathcal{A}, \mathcal{B}$ are $*$-algebras. A map $\pi : \mathcal{A} \longrightarrow \mathcal{B}$ is called a* **homomorphism of $*$-algebras** *or a* **$*$-homomorphism** *if, and only if, it respects the structure of a $*$-algebra, i. e.,*

$$\pi(\alpha a + \beta b) = \alpha\pi(a) + \beta\pi(b), \quad \forall a,b \in \mathcal{A}, \forall \alpha,\beta \in \mathbb{C}; \tag{3.1}$$

$$\pi(ab) = \pi(a)\pi(b), \quad \forall a,b \in \mathcal{A}; \tag{3.2}$$

$$\pi(a^*) = \pi(a)^*, \forall a \in \mathcal{A}. \tag{3.3}$$

Many results about abstract C*-algebras are obtained by studying properties of their representations by operators on a Hilbert space where one defines

Definition 3.1.3 *Let $\mathcal{A}$ be a C*-algebra. A* **representation** $(\pi, \mathcal{H})$ *of $\mathcal{A}$ is a *-homomorphism π of $\mathcal{A}$ into the C*-algebra $\mathcal{B}(\mathcal{H})$ of all bounded linear operators on a Hilbert space $\mathcal{H}$.*

A representation $(\pi, \mathcal{H})$ of $\mathcal{A}$ is called **cyclic** *if there exists a* **cyclic vector**, *i.e., a vector $x \in \mathcal{H}$ such that the closed linear subspace $[\pi(\mathcal{A})x]$ generated by all $\pi(a)x \in \mathcal{H}$ equals the representation space $\mathcal{H}$:*

$$[\pi(\mathcal{A})x] = \mathcal{H}.$$

3.1.1 Representations of $\mathcal{B}(\mathcal{H})$

For the C*-algebra of all bounded linear operators on a Hilbert space the general form of its representations can be determined. This result will be used later in our analysis of completely positive maps.

By Theorem 22.3.4 we know that the set of all compact operators $\mathcal{B}_c(\mathcal{H})$ on a (separable) Hilbert space $\mathcal{H}$ is a C*-algebra (without unit, if $\mathcal{H}$ is infinite dimensional). The following result clarifies the structure of all its representations.

Theorem 3.1.4 *Every continuous representation $(\pi, \mathcal{H}_\pi)$ of the C^*-algebra $\mathcal{B}_c(\mathcal{H})$ is equivalent to the direct sum $\bigoplus_n(\pi_n, \mathcal{H}_n)$ of the identity representation π_n, $A \mapsto A$, $A \in \mathcal{B}_c(\mathcal{H})$, and the zero representation $A \mapsto 0$, $A \in \mathcal{B}_c(\mathcal{H})$.*

Proof. Let $\{e_n\}$ be an orthonormal basis of $\mathcal{H}$ and denote by P_n the orthogonal projector onto the one-dimensional subspace $[e_n] = \mathbb{C}e_n$ spanned by e_n. By Theorem 22.3.2 we know that for $A \in \mathcal{B}_c(\mathcal{H})$ the finite rank operators

$$A \sum_{n=1}^{N} P_n, \quad N \in \mathbb{N}$$

converge in (operator) norm to A. Since the sequence of projectors $\sum_{j=1}^{M} P_j$, $M \in \mathbb{N}$ is bounded in norm (by 1) it follows that

$$A = \lim_{M,N\to\infty} \sum_{j=1}^{M} \sum_{n=1}^{N} P_j A P_n \quad \text{in } \mathcal{B}(\mathcal{H}). \tag{3.4}$$

Now calculate for $x \in \mathcal{H}$

$$P_j A P_n x = \langle e_j, A e_n \rangle \langle e_n, x \rangle e_j$$

and define operators U_{jn} on $\mathcal{H}$ by

$$U_{jn} x = \langle e_n, x \rangle e_j.$$

Then we can write

$$P_j A P_n = a_{jn} U_{jn}, \quad a_{jn} = \langle e_j, A e_n \rangle. \tag{3.5}$$

The operators U_{jn} are partial isometries from $[e_n]$ to $[e_j]$ which satisfy for all $j, n, m \in \mathbb{N}$

$$U_{nn} = P_n, \; U_{jn}^* = U_{nj}, \; U_{jn} U_{nm} = U_{jm} \tag{3.6}$$

As finite rank operator all the operators P_n, U_{jn} belong to $\mathcal{B}_c(\mathcal{H})$. Note also that for any fixed $n \in \mathbb{N}$ the closure of the ranges of all the operators $\{U_{jn} : j \in \mathbb{N}\}$ is $\mathcal{H}$.

Now let $(\pi, \mathcal{H}_\pi)$ be a continuous representation of $\mathcal{B}_c(\mathcal{H})$. For every $A \in \mathcal{B}_c(\mathcal{H})$ it follows

$$\pi(A) = \lim_{M,N\to\infty} \sum_{j=1}^{M} \sum_{n=1}^{N} \pi(P_j)\pi(A)\pi(P_n) = \lim_{M,N\to\infty} \sum_{j=1}^{M} \sum_{n=1}^{N} a_{jn} \pi(U_{jn}). \tag{3.7}$$

Therefore the knowledge of all the $\pi(U_{jn})$ allows to find the representatives $\pi(A)$ for all $A \in \mathcal{B}_c(\mathcal{H})$. Since π is a representation the relations (3.6) also hold for the representing operators $\pi(P_n)$ and $\pi(U_{jn})$. For $n \in \mathbb{N}$ define

$$M_n = \mathrm{ran}\,(\text{ß}(\mathrm{P_n})) \subset \mathcal{H}_{\text{ß}}.$$

Since $\pi(P_n)\pi(P_j) = 0$ for $n \neq j$ the closed subspaces M_n of $\mathcal{H}_\pi$ are pairwise orthogonal for different indices. Relations (3.6) for $\pi(P_n)$, $\pi(U_{jn})$ imply furthermore that the operators $\pi(U_{jn})$ are partial isometries with initial domain M_n and terminal domain M_j. Hence all the subspaces $\{M_n\}$ have the same dimension.

If $\pi(P_n) = 0$ for all $n \in \mathbb{N}$, then by (3.7) one has $\pi(A) = 0$ for all $A \in \mathcal{B}_c(\mathcal{H})$ and π is the zero representation. If π is not the zero representation, there is $n \in \mathbb{N}$ such that $\pi(P_n) \neq 0$ and thus $M_n \neq \{0\}$. Hence there is a unit vector $f_n \in M_n$ and we can define an orthonormal system $\{f_j\}$ in $\mathcal{H}_\pi$ by setting

$$f_j = \pi(U_{jn})f_n, \quad \text{for all } j \in \mathbb{N}.$$

This orthonormal system generates a closed linear subspace $M_\pi \subset \mathcal{H}_\pi$:

$$M_\pi = [\{f_j\}].$$

We now show that this subspace is invariant under all $\pi(A)$, $A \in \mathcal{B}_c(\mathcal{H})$: Observe first that

$$\pi(U_{jm})f_k = \pi(U_{jm})\pi(U_{kn})f_n = 0 \text{ for m} \neq \text{k}$$
$$\pi(U_{jm})f_m = \pi(U_{jm})\pi(U_{mn})f_n = \pi(U_{jn})f_n = f_j$$

holds. Because of (3.5) this implies

$$\pi(P_j)\pi(A)\pi(P_m)f_k = 0 \quad \text{for m} \neq \text{k} \tag{3.8}$$
$$\pi(P_j)\pi(A)\pi(P_m)f_m = a_{jm}f_j. \tag{3.9}$$

Thus all the operators $\pi(P_j)\pi(A)\pi(P_m)$ are reduced by the subspace M_π; now (3.7) implies that all operators $\pi(A)$, $A \in \mathcal{B}_c(\mathcal{H})$, are reduced by this subspace too.

In the subspace M_π the operator $\pi(P_j)$ is the projection onto the subspace $\mathbb{C}f_j$, $j \in \mathbb{N}$, hence by (3.9) the matrix

$$[a_{jm}] = [\langle e_j, Ae_m\rangle] = [\langle f_j, \pi(A)f_m\rangle_{M_\pi}] \tag{3.10}$$

is the matrix of $\pi(A)$ with respect to the orthonormal basis $\{f_j\}$ of M_π.

Next define an isometric mapping V of $\mathcal{H}$ onto M_π by setting

$$Ve_j = f_j, \quad j \in \mathbb{N}$$

and extend it by linearity and continuity to all of $\mathcal{H}$. Relation (3.10) implies

$$V^* \pi(A) V = A \quad \text{for all } A \in \mathcal{B}_c(\mathcal{H})$$

and thus the representation π is unitarily equivalent to the identity representation.

If $M_\pi = \mathcal{H}_\pi$, we are done. Otherwise look at the orthogonal complement $M_\pi^\perp$ of M_π in $\mathcal{H}_\pi$. Certainly, $M_\pi^\perp$ is invariant under all $\pi(A)$, $A \in \mathcal{B}_c(\mathcal{H})$. For all $x \in M_\pi$ and all $y \in M_\pi^\perp$ we have

$$\langle x, \pi(A)y\rangle_{\mathcal{H}_\pi} = \langle \pi(A^*)x, y\rangle_{\mathcal{H}_\pi} = 0$$

since $\pi(A^*)x \in M_\pi$. Thus the restriction of $\pi(A)$ to $M_\pi^\perp$ defines a representation of $\mathcal{B}_c(\mathcal{H})$ in $M_\pi^\perp$ and we can proceed as above to find an invariant subspace M'_π of $M_\pi^\perp$ on which this representation is unitarily equivalent to the identity representation. Now by iteration of this argument we conclude. □

Theorem 3.1.5 (Naimark) *Every representation $(\pi, \mathcal{H}_\pi)$ of the C^*-algebra $\mathcal{B}(\mathcal{H})$ of all bounded operators on a separable Hilbert space $\mathcal{H}$ is the direct sum of identity representations $A \mapsto A$ and a representation of the quotient algebra $\mathcal{B}(\mathcal{H})/\mathcal{B}_c(\mathcal{H})$.*

If the representation of this quotient algebra is not the zero representation, then it is an isomorphism of $\mathcal{B}(\mathcal{H})/\mathcal{B}_c(\mathcal{H})$ into the algebra of bounded linear operators on a Hilbert space.

Proof. Since $\mathcal{B}(\mathcal{H})$ is a C^*-algebra with unit Theorem 3.2.3 implies that every representation of it is continuous. Therefore a representation $(\pi, \mathcal{H}_\pi)$ of $\mathcal{B}(\mathcal{H})$ is at the same time a continuous representation of $\mathcal{B}_c(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$. Thus Theorem 3.1.4 applies. Hence modulo a unitary map the representation space $\mathcal{H}_\pi$ is the direct sum $\bigoplus_n \mathcal{H}_n$ of copies $\mathcal{H}_n = \mathcal{H}$ of $\mathcal{H}$ and a space $\mathcal{H}_0$. In each of the spaces $\mathcal{H}_n$ the representation π is the identity representation of $\mathcal{B}_c(\mathcal{H})$ while in $\mathcal{H}_0$ it is the zero representation. We have to show that for arbitrary but fixed n the representation π of $\mathcal{B}(\mathcal{H})$ also reduces to the identity representation in $\mathcal{H}_n$.

With the orthogonal projector $Q_n : \mathcal{H}_\pi \longrightarrow \mathcal{H}_n$ introduce $\pi(A)_n = Q_n\pi(A)Q_n$, i.e.,

$$\pi(A)_n x = Q_n\pi(A)x \quad \text{for all } x \in \mathcal{H}_n.$$

Clearly $\pi(A)_n$ is a well defined bounded linear operator on $\mathcal{H}_n$. We show $\pi(A)_n = A$ by showing that their matrices coincide, calculated with respect to an ONB $\{e_j\}$ of $\mathcal{H}_n = \mathcal{H}$. For the projection operator P_j onto $\mathbb{C}e_j$ we have as earlier for all $i, j \in \mathbb{N}$,

$$P_i A P_j = a_{ij} U_{ij}$$

and thus

$$\pi(P_i)\pi(A)\pi(P_j) = \pi(P_i A P_j) = a_{ij} U_{ij}.$$

Recall $P_j, U_{ij}, P_i A P_j \in \mathcal{B}_c(\mathcal{H})$ and thus we can calculate as follows:

$$\begin{aligned} &\langle e_i, \pi(A)_n e_j\rangle = \langle Q_n e_i, \pi(A)e_j\rangle = \langle e_i, \pi(A)e_j\rangle = \\ &\langle P_i e_i, \pi(A)P_j e_j\rangle = \langle e_i, P_i\pi(A)P_j e_j\rangle = a_{ij}\langle e_i, U_{ij}e_j\rangle = a_{ij} \end{aligned}$$

and we conclude $\pi(A)_n = A$.

It follows, for all $x \in \mathcal{H}_n$, $A, B \in \mathcal{B}(\mathcal{H})$,

$$Q_n\pi(A)\pi(B)x = (AB)x, \quad Q_n\pi(A)Q_n\pi(B)x = A(Bx),$$

and therefore $Q_n\pi(A)(\pi(B)x) = Q_n\pi(A)Q_n(\pi(B)x)$. Denote the closed linear hull of the set $\{\pi(B)x : x \in \mathcal{H}_n, B \in \mathcal{B}(\mathcal{H})\}$ by $\mathcal{H}'_n$. Then the above argument shows that in $\mathcal{H}'_n$ one has

$$Q_n\pi(A) = Q_n\pi(A)Q_n. \tag{3.11}$$

Naturally $\mathcal{H}_n \subset \mathcal{H}'_n$ and therefore $Q_n = 0$ on ${\mathcal{H}'_n}^\perp \subset \mathcal{H}_\pi$. As earlier one shows that $\mathcal{H}'_n$ and ${\mathcal{H}'_n}^\perp$ are invariant under all $\pi(A)$, $A \in \mathcal{B}(\mathcal{H})$, hence $Q_n\pi(A) = 0$ in ${\mathcal{H}'_n}^\perp$ and (3.11) holds in all of $\mathcal{H}_\pi$.

Now apply the involution to (3.11) and in the result replace A^* by A. This gives $\pi(A)Q_n = Q_n\pi(A)Q_n$ and thus

$$Q_n\pi(A) = \pi(A)Q_n,$$

i.e., the space $\mathcal{H}_n$ reduces all the operators $\pi(A)$, $A \in \mathcal{B}(\mathcal{H})$. In the space $\mathcal{H}_n$ the identity $\pi(A)_n = A$ now takes the form

$$Ax = Q_n\pi(A)x = \pi(A)Q_n x = \pi(A)x, \quad x \in \mathcal{H}_n$$

and thus in the space $\mathcal{H}_n$ the representation of all of $\mathcal{B}(\mathcal{H})$ reduces to the identity representation. □

3.2 On positive elements and positive functionals

An element a of a *-algebra $\mathcal{A}$ is called **positive** if, and only if, one of the following equivalent conditions hold:

$$a = b^* b \qquad \text{for some} \quad b \in \mathcal{A}; \tag{3.12}$$
$$a = c^2 \qquad \text{for some} \quad c \in \mathcal{A}_h = \{a \in \mathcal{A} : a^* = a\}. \tag{3.13}$$

In the case that $\mathcal{A}$ is a C*-algebra of operators on a Hilbert space $\mathcal{H}$ or a subspace of such an algebra one has a third characterization of positive elements a, namely

$$\langle x, ax \rangle \geq 0 \quad \text{for all} \quad x \in \mathcal{H}. \tag{3.14}$$

In this case the proof of equivalence of these three conditions is straight forward by using the square root lemma (Theorem 21.5.1) and the polar decomposition of operators (Theorem 21.5.2). In the case of an abstract C*-algebra the characterization of positive elements one has to refer to spectral theory for these algebras (see Theorem 6.1 of 24).

Using the characterization (3.14) of positive elements it follows easily that the set $\mathcal{A}_+$ of all positive elements in $\mathcal{A}$ is a closed convex cone, i.e., if $a, b \in \mathcal{A}_+$ and $\alpha, \beta \geq 0$ then $\alpha a + \beta b \in \mathcal{A}_+$. This cone satisfies $\mathcal{A}_+ \cap (-\mathcal{A}_+) = \{0\}$. Hence

$\mathcal{A}_+$ induces an **order** in the real Banach space $\mathcal{A}_h$ of hermitian elements in $\mathcal{A}$. For $a, b \in \mathcal{A}_h$ we write $a \geq b$ if, and only if, $a - b \in \mathcal{A}_+$.

Definition 3.2.1 *A linear map $f : \mathcal{A} \longrightarrow \mathbb{C}$ on a C^*-algebra $\mathcal{A}$ is called a* **positive functional** *if its restriction to the cone $\mathcal{A}_+$ of positive elements has only nonnegative values, i.e., if $f(a) \geq 0$ for all $a \in \mathcal{A}_+$.*

The following proposition collects the basic facts about positive linear functionals.

Proposition 3.2.2 *For a positive linear functional f on a C^*-algebra $\mathcal{A}$ with unit one has:*

a) For all $a, b \in \mathcal{A}$

$$f(a^*b) = \overline{f(b^*a)} \tag{3.15}$$

$$|f(a^*b)|^2 \leq f(a^*a) f(b^*b) \tag{3.16}$$

b) f is continuous and $\|f\| = f(I)$.

Proof. For the proof of the first part of a) take arbitrary $a, b \in \mathcal{A}$ and apply f to the two polarization identities

$$4a^*b = \sum_{j=0}^{3} \mathrm{i}^j (\mathrm{b} + \mathrm{i}^j \mathrm{a})^* (\mathrm{b} + \mathrm{i}^j \mathrm{a}), \quad 4\mathrm{ba}^* = \sum_{\mathrm{j}=0}^{3} \mathrm{i}^j (\mathrm{b} + \mathrm{i}^j \mathrm{a})(\mathrm{b} + \mathrm{i}^j \mathrm{a})^*$$

and compare the results. For the proof of the estimate take arbitrary $a, b \in \mathcal{A}$ and arbitrary $\alpha, \beta \in \mathbb{C}$ and observe

$$\begin{aligned} 0 &\leq f((\alpha a + \beta b)^*(\alpha a + \beta b)) \\ &= |\alpha|^2 f(a^*a) + \overline{\alpha}\beta f(a^*b) + \alpha\overline{\beta} f(b^*a) + |\beta|^2 f(b^*b), \end{aligned}$$

thus, because of the first part, the quadratic form on $\mathbb{C}$

$$|\alpha|^2 f(a^*a) + 2\mathrm{Re}(\overline{\mathrm{ff}}\mathrm{fif}(\mathrm{a}^*\mathrm{b})) + |\mathrm{fi}|^2\mathrm{f}(\mathrm{b}^*\mathrm{b})$$

is nonnegative, hence its coefficients have to satisfy (3.16).

For the proof of b) observe for all $a \in \mathcal{A}$ and all $x \in \mathcal{H}$

$$\langle x, a^*ax \rangle = \|ax\|^2 \leq \|a\|^2 \|x\|^2,$$

thus

$$a^*a \leq \|a\|^2 I, \tag{3.17}$$

and hence for a positive functional f it follows $f(a^*a) \leq \|a\|^2 f(I)$. The Cauchy Schwarz inequality (3.16) implies $|f(a)|^2 \leq f(I^*I) f(a^*a)$ and we find for all $a \in \mathcal{A}$

$$|f(a)| \leq f(I) \|a\|$$

and this proves that f is continuous and that $\|f\| = \sup\{|f(a)|; a \in \mathcal{A}, \|a\| \leq 1\} \leq f(I)$. But clearly $f(I) \leq \|f\|$ and therefore $\|f\| = f(I)$.
□

Here are some simple **examples of positive functionals** on a C^*-algebra $\mathcal{A}$ of operators on a Hilbert space $\mathcal{H}$:

For $x \in \mathcal{H}$ define a function $f_x : \mathcal{A} \to \mathbb{C}$ by

$$f_x(a) = \langle x, ax \rangle \quad \text{for all} \quad a \in \mathcal{A}. \tag{3.18}$$

Clearly, by condition (3.14) this functional is positive when restricted to $\mathcal{A}_+$.

Theorem 3.2.3 *Every representation $(\pi, \mathcal{H})$ of a C^*-algebra $\mathcal{A}$ with unit I is continuous and for all $a \in \mathcal{A}$*

$$\|\pi(a)\| \leq \|a\| .$$

Proof. Given a representation $(\pi, \mathcal{H})$ of $\mathcal{A}$, take $x \in \mathcal{H}$ and define a functional $f_x : \mathcal{A} \longrightarrow \mathbb{C}$ by

$$f_x(a) = \langle x, \pi(a)x \rangle \quad \text{for all } a \in \mathcal{A}.$$

Since π is a *-homomorphism, f_x is a positive functional on $\mathcal{A}$: For all $a \in \mathcal{A}$ one has

$$f_x(a^*a) = \|\pi(a)x\|^2 \geq 0.$$

By Proposition 3.2.2 the norm of this functional is $\|f_x\| = f_x(I) = \|x\|^2$. It follows

$$\|\pi(a)x\|^2 = f_x(a^*a) \leq \|f_x\| \, \|a^*a\| = \|x\|^2 \, \|a\|^2$$

and thus $\|\pi(a)x\| \leq \|x\| \, \|a\|$, hence $\|\pi(a)\| = \sup\{\|\pi(a)x\| : x \in \mathcal{H}, \|x\| \leq 1\} \leq \|a\|$. □

3.2.1 The GNS-construction

In this section we provide the answer to the question *What is the general form of positive functionals on a C^*-algebra with unit?*

The answer is well known since many years and is given by the GNS-construction (Gelfand-Naimark-Segal) which we explain now. This construction can be done in a much more general setting.

Theorem 3.2.4 (GNS-construction for $\mathcal{A}$)) *Let f be a positive functional on a C^*-algebra $\mathcal{A}$ with unit I with $f(I) = 1$. Then there is a Hilbert space $\mathcal{H}_f$, a unit vector $\Omega_f \in \mathcal{H}_f$, and a mapping π_f on $\mathcal{A}$ with values in the space $\mathcal{B}(\mathcal{H}_f)$ of bounded linear operators on $\mathcal{H}_f$ with the following properties:*

1. *$\pi_f : \mathcal{A} \longrightarrow \mathcal{B}(\mathcal{H}_f)$ is linear;*
2. *$\pi_f(ab) = \pi_f(a)\pi_f(b)$ for all $a,b \in \mathcal{A}$;*
3. *$\pi_f(a^*) = \pi_f(a)^*$ for all $a \in \mathcal{A}$;*

such that

$$f(a) = \langle \Omega_f, \pi_f(a)\Omega_f \rangle \quad \text{for all } a \in \mathcal{A} \tag{3.19}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product of $\mathcal{H}_f$.

In addition one has $[\pi_f(\mathcal{A})\Omega_f] = \mathcal{H}_f$, i.e., Ω_f is cyclic so that π_f is a cyclic representation of $\mathcal{A}$.

The triple $(\mathcal{H}_f, \Omega_f, \pi_f)$ is unique up to unitary equivalence, i.e., if we also have $f(a) = \langle \Omega, \pi(a)\Omega \rangle$ for all $a \in \mathcal{A}$ where the triple $(\mathcal{H}, \Omega, \pi)$ has the properties specified above for $(\mathcal{H}_f, \Omega_f, \pi_f)$ then there is a unitary operator $U : \mathcal{H}_f \to \mathcal{H}$ such that $\Omega = U\Omega_f$ and $\pi(a) = U\pi_f(a)U^*$ for all $a \in \mathcal{A}$.

Proof. Construction of $\mathcal{H}_f$: Define

$$I_f = \{a \in \mathcal{A} : f(a^*a) = 0\}.$$

By Proposition 3.2.2 the given functional f is continuous on $\mathcal{A}$; since $a \longrightarrow a^*a$ is continuous on $\mathcal{A}$, I_f is a closed subset. If $a,b \in I_f$ then,

$$f((a+b)^*(a+b)) = f(a^*a) + 2\mathrm{Re}(\mathrm{f}(\mathrm{a}^*\mathrm{b})) + \mathrm{f}(\mathrm{b}^*\mathrm{b}) = 2\mathrm{Re}(\mathrm{f}(\mathrm{a}^*\mathrm{b})) = 0$$

since by (3.16) $2\mathrm{Re}(\mathrm{f}(\mathrm{a}^*\mathrm{b})) \leq \mathrm{f}(\mathrm{a}^*\mathrm{a})\mathrm{f}(\mathrm{b}^*\mathrm{b})$, hence $a+b \in I_f$. Similarly, for $a \in \mathcal{A}$ and $b \in I_f$ the Cauchy-Schwarz inequality (3.16) implies that

$$f((ab)^*(ab)) = f(b^*a^*ab)) \leq f(b^*b)^{1/2} f((a^*ab)^*(a^*ab))^{1/2} = 0,$$

hence $ab \in I_f$. Since I_f is obviously invariant under multiplication with scalars we conclude that I_f is a closed left ideal in $\mathcal{A}$.

$$\mathcal{A} \cdot I_f \subseteq I_f\,. \tag{3.20}$$

Form the quotient space $\mathcal{H}_f^0$ of $\mathcal{A}$ with respect to I_f, i.e., the space of all equivalence classes

$$[a]_f = a + I_f, \quad a \in \mathcal{A}; \tag{3.21}$$

$$\mathcal{H}_f^0 = \mathcal{A}/I_f = \left\{ [a]_f : a \in \mathcal{A} \right\}. \tag{3.22}$$

On $\mathcal{H}_f^0$ define addition and scalar multiplication of equivalence classes through their representatives, i.e.,

$$[a]_f + [b]_f = [a+b]_f,\ \lambda[a]_f = [\lambda a]_f, \quad \forall\, \lambda \in \mathbb{C}, a,b \in \mathcal{A}.$$

Thus $\mathcal{H}_f^0$ becomes a complex vector space.

Next one shows that the formula

$$\langle [a]_f, [b]_f \rangle = f(a^*b), \quad [a]_f, [b]_f \in \mathcal{H}_f^0 \tag{3.23}$$

defines a scalar product on the vector space $\mathcal{H}_f^0$. Finally define $\mathcal{H}_f$ as the completion of $\mathcal{H}_f^0$ with respect to the norm defined by this scalar product and extend the scalar product (3.23) by continuity to $\mathcal{H}_f$. Thus $\mathcal{H}_f$ is a complex Hilbert space.

Construction of Ω_f and π_f: First define

$$\Omega_f = [I]_f. \tag{3.24}$$

Clearly $\Omega_f \in \mathcal{H}_f$ satisfies $\langle \Omega_f, \Omega_f \rangle = f(I^*I) = f(I) = 1$. Hence this is a unit vector.

Next define

$$\pi_f^0(a)[b]_f = [ab]_f, \quad \forall\, a,b \in \mathcal{A}. \tag{3.25}$$

Because of property (3.20), π_f^0 is well defined. And it follows easily that $\pi_f^0(a)$ is a linear operator on $\mathcal{H}_f^0$ for all $a \in \mathcal{A}$. In order to prove boundedness of the linear operator $\pi_f^0(a)$ we estimate as follows. For all $b \in \mathcal{A}$ one has, using (3.17),

$$\left\|\pi_f^0(a)[b]_f\right\|_f^2 = \langle [ab]_f, [ab]_f \rangle = f((ab)^* ab) = f(b^* a^* ab)$$
$$\leq f(b^* \|a\|^2 Ib) = \|a\|^2 f(b^* b) = \|a\|^2 \langle [b]_f, [b]_f \rangle_f,$$

hence $\pi_f^0(A)$ is bounded and $\left\|\pi_f^0(a)\right\| \leq \|a\|$.

Thus by continuity and density of $\mathcal{H}_f^0$, π_f^0 extends uniquely to a mapping $\pi_f : \mathcal{A} \longrightarrow \mathcal{B}(\mathcal{H}_f)$. It is straight forward to see that this mapping is linear. This proves property 1.

Since the product in $\mathcal{A}$ is associative, property 2. of π_f follows easily from its definition:

$$\pi_f(ab)[c]_f = [(ab)c]_f = [a(bc)]_f = \pi_f(a)[bc]_f =$$
$$\pi_f(a)(\pi_f(b)[c]_f) = (\pi_f(a)\pi_f(b))[c]_f, \quad \forall a,b,c \in \mathcal{A}.$$

In order to establish property 3. we calculate as follows, for arbitrary $a,b,c \in \mathcal{A}$:

$$\langle \pi_f(a)^*[b]_f, [c]_f \rangle = \langle [b]_f, \pi_f(a)[c]_f \rangle = \langle [b]_f, [ac]_f \rangle = f(b^* ac)$$
$$= f((a^* b)^* c) = \langle [a^* b]_f, [c]_f \rangle = \langle \pi_f(a^*)[b]_f, [c]_f \rangle,$$

hence $\pi_f(a^*) = \pi_f(a)^*$ for all $a \in \mathcal{A}$.

Finally note that our construction gives for all $a \in \mathcal{A}$

$$\langle \Omega_f, \pi_f(a)\Omega_f \rangle = \langle [I]_f, \pi_f(a)[I]_f \rangle = f(I^* a I) = f(a),$$

therefore the representation formula (3.19) holds.

Uniqueness up to unitary equivalence: Suppose that we also have f represented as $f(a) = \langle \Omega, \pi(a)\Omega \rangle$. Define a linear mapping $U : \mathcal{H}_f \longrightarrow \mathcal{H}$ by

$$U^0[a]_f = \pi(A)\Omega, \quad \forall a \in \mathcal{A}.$$

It follows that U^0 is linear and satisfies, for all $a,b \in \mathcal{A}$

$$\langle U^0[a]_f, U^0[b]_f \rangle = \langle \pi(a)\Omega, \pi(b)\Omega \rangle = \langle \Omega, \pi(a^* b)\Omega \rangle = f(a^* b) = \langle [a]_f, [b]_f \rangle.$$

We conclude that U^0 is an isometry defined on the dense subspace $\mathcal{H}_f^0$ with dense range $\pi(\mathcal{A})\Omega$, hence it extends continuously to a unique unitary operator $U : \mathcal{H}_f \longrightarrow \mathcal{H}$. Next calculate

$$\begin{aligned}\pi(a)\pi(b)\Omega &= \pi(ab)\Omega = U[ab]_f = U\pi_f(a)[b]_f = \\ &= U\pi_f(a)U^*U[b]_f = U\pi_f(a)U^*\pi(b)\Omega\end{aligned}$$

for all $a, b \in \mathcal{A}$. Since $\pi(\mathcal{A})\Omega$ is dense in $\mathcal{H}$ we conclude $\pi(a) = U\pi_f(a)U^*$ for all $a \in \mathcal{A}$. □

3.3 Normal States

In the last section we considered positive linear functionals f on a C*-algebra $\mathcal{A}$ with unit I satisfying $f(I) = 1$. Such functionals are called **states of** $\mathcal{A}$. Here under an addtional continuity asssumption we determine the general form of states in the case where $\mathcal{A}$ is a weakly closed subalgebra of $\mathcal{B}(\mathcal{H})$ (see Section 2.3), i.e., if $\mathcal{A}$ is a **von Neumann algebra**.

By Propositon 3.2.2 states are continuous for the (operator) norm when $\mathcal{A}$ is a C*-algebra. Note that states generate the cone of positive linear functionals.

Simple examples of states are **vector states** μ_x, $x \in \mathcal{H}$, $\|x\| = 1$, defined by

$$\mu_x(A) = \langle x, Ax \rangle \quad A \in \mathcal{A}. \tag{3.26}$$

Another class of examples is obtained as follows. In Formula 2.18 choose $g_n = e_n$ with $\{e_n\} \in \ell^2(\mathcal{H})$ and $\sum_n \|e_n\|^2 = 1$. Then the operator $\hat{T} = \sum_n [e_n, e_n]$ is a positive trace class operator with $\mathrm{Tr}(\hat{T}) = \sum_n \|e_n\|^2 = 1$ and the formula

$$T(A) = \mathrm{Tr}(\hat{T}A) = \sum_n \langle e_n, Ae_n \rangle \tag{3.27}$$

defines a state on $\mathcal{B}(\mathcal{H})$.

It turns out that under the conditions we are considering every state on $\mathcal{A}$ will be of this form. This result is used quite often in quantum physics and naturally in the theory of operator algebras.

Definition 3.3.1 *A positive linear functional μ on a von Neumann algebra $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ is called*

1. **normal** *if for every bounded increasing net $\{A_i : i \in I\} \subset \mathcal{A}_h = \{A \in \mathcal{A} : A^* = A\}$ one has*

$$\mu(\sup_I A_i) = \sup_I \mu(A_i). \tag{3.28}$$

2. **completely additive** *if for every orthogonal family of projections p_i in $\mathcal{A}$ one has*

$$\mu(\sum_{i\in I} p_i) = \sum_{i\in I} \mu(p_i). \tag{3.29}$$

Note that a family of projections is called orthogonal if any two different projections are orthogonal, i.e., $p_i p_j = 0$ for $i \neq j$. In this context it is important to be aware of the following simple result.

Proposition 3.3.2 (Theorem of Vigier) *If $\{A_i : i \in I\} \subset \mathcal{B}(\mathcal{H})$ is a bounded increasing net of self-adjoint operators then there is a self-adjoint operator $A = \sup_I A_i$ such that*

$$A = \lim_I A_i$$

in the strong topology on $\mathcal{B}(\mathcal{H})$.

Proof. Every $x \in \mathcal{H}$ defines an increasing net $\langle x, A_i x\rangle$ in $\mathbb{R}$ which is bounded by $C\|x\|^2$ if C denotes the bound for the given net ($\|A_i\| \leq C$ for all $i \in I$), hence the net converges. The polarization identity (Proposition 14.1.2) implies that the net $\langle y, A_i x\rangle$ converges for any fixed $x, y \in \mathcal{H}$; denote the limit by $B(y,x)$. It follows that $B(y,x)$ is a symmetric sesquilinear form on $\mathcal{H}$ bounded by $C\|y\|\,\|x\|$. Such forms define a unique self-adjoint operator $A \in \mathcal{B}(\mathcal{H})$ by $B(y,x) = \langle y, Ax\rangle$ for all $x, y \in \mathcal{H}$. By construction $\langle y, Ax\rangle = \lim_I \langle y, A_i x\rangle$, hence $A = \lim_I A_i$ for the weak topology on $\mathcal{B}(\mathcal{H})$. Furthermore, for every $x \in \mathcal{H}$,

$$\langle x, Ax\rangle = \lim_I \langle x, A_i x\rangle = \sup_I \langle x, A_i x\rangle,$$

hence $A = \sup_I A_i$.

Since the net is bounded it follows that $A = \lim_I A_i$ for the strong topology on $\mathcal{B}(\mathcal{H})$: For every $x \in \mathcal{H}$ we can estimate $Ax - A_i x$, using $A - A_i \geq 0$,

$$\begin{aligned}\|(A - A_i)x\|^2 &= \left\|(A - A_i)^{1/2}(A - A_i)^{1/2}x\right\|^2 \leq \left\|(A - A_i)^{1/2}\right\|^2 \left\|(A - A_i)^{1/2}x\right\|^2 \\ &= \|A - A_i\|\,\langle x, (A - A_i)x\rangle \leq 2C\langle x, (A - A_i)x\rangle;\end{aligned}$$

thus weak convergence of the net implies strong convergence. □

Note that in our context this result implies that $\sup_I A_i \in \mathcal{A}$ so that (3.28) is meaningful. The main result of this section is

Theorem 3.3.3 (characterization of normal states) *For a state μ on a von Neumann algebra $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ the following statements are equivalent:*

a) μ is normal;

b) μ is completely additive;

c) μ is of the form

$$\mu(A) = \mathrm{Tr}(AW), \quad A \in \mathcal{A} \tag{3.30}$$

with a positive trace class operator W with $\mathrm{Tr}(W) = 1$.

Proof. For the proof we proceed in the order of the implications a) $\Rightarrow$ b) $\Rightarrow$ c) $\Rightarrow$ a).

a) $\Rightarrow$ b): Let $\{p_i : i \in I\}$ be any orthogonal family of projections in $\mathcal{A}$. For finite parts J of the index set I introduce the projection $p_J = \sum_{i \in J} p_i$. Then $\{p_J : J \subset I, J \text{ finite}\}$ is a monotone increasing net which is bounded by *id*. Thus by Proposition 3.3.2

$$\lim_J p_J = \sum_{i \in I} p_i.$$

Since μ is assumed to be normal it follows

$$\mu(\sum_{i \in I} p_i) = \lim_J \mu(p_J) = \lim_J \sum_{i \in J} \mu(p_i) = \sum_{i \in I} \mu(p_i),$$

hence μ is completely additive.

b) $\Rightarrow$ c): This is the core of the proof. The main technical part of the argument is formulated in the following Proposition 3.3.4. This proposition states that a completely additive state μ on $\mathcal{A}$ is strongly continuous when restricted to the unit ball $\mathcal{A}_1$ of $\mathcal{A}$. Then the second part of Theorem 2.3.2 implies that μ is of the form

$$\mu(A) = \sum_n \langle g_n, Ae_n \rangle, \quad A \in \mathcal{A}$$

with $\{g_n\}, \{e_n\} \in \ell^2(\mathcal{H})$ where we used (2.18). As in the introductory example the form (3.30) of μ follows.

c) $\Rightarrow$ a): Again according to the second part of Theorem 2.3.2 μ is σ-weakly continuous if c) is assumed. On bounded sets the weak and σ-weak topology agree according to Lemma 2.3.1. Thus by Proposition 3.3.2 we conclude. □

Proposition 3.3.4 *Every completely additive state μ on a von Neumann algebra $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ is strongly continous when restricted to the unit ball $\mathcal{A}_1$ of $\mathcal{A}$.*

Proof. For the proof we have to find suitable seminorms for the strong topology by which we can estimate the given state. This could be achieved by finding suitable vector states which dominate μ. This idea can be realised first on certain parts of $\mathcal{A}$ and then on all of $\mathcal{A}$.

Claim 1: There are a nonzero projection $p \in \mathcal{A}$ and a vector $x \in \mathcal{H}$ such that

$$\mu \leq \mu_x \quad \text{on} \quad \text{p}\mathcal{A}\text{p}.$$

For the proof of this claim choose $x \in \mathcal{H}$, $\|x\| = 1$. Then we have $\mu_x(I) = \langle x, Ix \rangle = 1 = \mu(I)$. Introduce $\mathcal{P}_0 = \{p \in \mathcal{A} : p = \text{projection}, {}^{-}{}_{\text{x}}(\text{p}) < {}^{-}(\text{p})\}$. If $\mathcal{P}_0$ is empty, then $\mu(p) \leq \mu_x(p)$ for all projections p in $\mathcal{A}$ and we are done.

If $\mathcal{P}_0$ is not empty consider the collection $\mathcal{P}$ of all subsets

$$P = \{p_i \in \mathcal{P}_0 : p_i \text{ mutually orthogonal}\} \subset \mathcal{P}_0.$$

By set inclusion $\mathcal{P}$ is a partially ordered set in which every chain has an upper bound (the union of the elements of this chain). Hence by Zorn's lemma $\mathcal{P}$ has a maximal element P. Then $p = \sum_{p_i \in P} p_i \leq I$ and thus, since μ is completely additive,

$$\mu_x(p) = \sum_{p_i \in P} \mu_x(p_i) < \sum_{p_i \in P} \mu(p_i) = \mu(\sum_{p_i \in P} p_i) = \mu(p) \leq \mu(I) = \mu_x(I),$$

hence $p < I$ and therefore $q = I - p \neq 0$. Since every projection $q' \in q\mathcal{A}q$ is orthogonal to each $p_i \in P$, by maximality of P we know $q' \notin \mathcal{P}_0$, and thus for all projections $q' \in q\mathcal{A}q$ it follows $\mu(q') \leq \mu_x(q')$.

According to the spectral theorem (Theorem **??**)[1] every positive $A \in q\mathcal{A}q$ is the norm limit of linear combinations $\sum_{j=1}^{n} \lambda_j q_j$ of projections $q_j \in q\mathcal{A}q$ with positive coefficients λ_j. The above estimate implies

$$\mu(\sum_{j=1}^{n} \lambda_j q_j) = \sum_{j=1}^{n} \lambda_j \mu(q_j) \leq \sum_{j=1}^{n} \lambda_j \mu_x(q_j) = \mu_x(\sum_{j=1}^{n} \lambda_j q_j).$$

Since μ and μ_x are continuous with respect to the norm topology we get in the limit

$$\mu(A) \leq \mu_x(A), \quad A \in q\mathcal{A}q, A \geq 0$$

and claim 1 follows.

Claim 2: There is a family $\{p_i\}$ of mutually orthogonal projections in $\mathcal{A}$ and of points $x_i \in \mathcal{H}$ such that

$$\mu \leq \mu_{x_i} \quad \text{on} \quad p_i \mathcal{A} p_i \quad \text{and} \quad \sum_i p_i = I. \tag{3.31}$$

The proof of this claim relies again on Zorn's lemma. According to the first claim we know that the set

$$\mathcal{S}_0 = \{(p, x) : p \text{ projection in } \mathcal{A}, x \in \mathcal{H}, \mu \leq \mu_x \text{ on pAp}\}$$

is not empty. Then consider the collection $\mathcal{S}$ of subsets

$$S = \{(p_i, x_i) \in \mathcal{S}_0 : \{p_i\} \text{ mutually orthogonal}\}.$$

$\mathcal{S}$ is partially ordered by set inclusion and then every chain in $\mathcal{S}$ has an upper bound. Zorn's lemma implies that $\mathcal{S}$ has a maximal element S_m. Define p as the sum of all projections p_i for which $(p_i, x_i) \in S_m$: $p = \sum_i p_i$. If $p < I$ then $q = I - p$ is a nontrivial projection. Apply the statement of the first claim to $q\mathcal{A}q \subseteq \mathcal{B}(q\mathcal{H})$. Hence there is $y \in q\mathcal{H}$ and a projection $p_0 \in q\mathcal{A}q$ such that $\mu \leq \mu_y$ on $p_0 q\mathcal{A}q p_0$. By construction p_0 is mutually orthogonal to all projections p_i with $(p_i, x_i) \in S_m$. This contradicts the maximality of S_m and therefore we get $p = \sum_i p_i = I$ and (3.31) follows.

Claim 3: Given $\epsilon > 0$ there is a neighborhood of zero U for the strong topology such that $|\mu(A)| \leq \epsilon$ for all $A \in U \cap \mathcal{A}_1$.

The proof of this claim follows now easily from (3.31): Since μ is a completely additive state we know

$$1 = \mu(I) = \mu(\sum_i p_i) = \sum_i \mu(p_i);$$

[1]Our version of the spectral theorem proves this claim only for the case $\mathcal{A} = \mathcal{B}(\mathcal{H})$. For the general case of $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ we have to refer to Theorem 5.2.2 of 19.

thus the index set of our maximal family S_m is countable and there is a finite subset J of the index set of our maximal family S_m such that $\sum_{i\in J}\mu(p_i)\leq 1-\frac{\epsilon^2}{4}$. For $q=I-\sum_{i\in J}p_i$ this gives $\mu(q)\leq\frac{\epsilon^2}{4}$. Define $U=\left\{A\in\mathcal{A}:\sum_{i\in J}\|Ap_ix_i\|\leq\epsilon/2\right\}$ and observe $\mu(A)=\mu(Aq)+\mu(A\sum_{i\in J}p_i)$. For all $A\in\mathcal{A}$ with $\|A\|\leq 1$ we estimate as follows:

$$|\mu(Aq)|\leq\mu(A^*A)^{1/2}\mu(q^*q)^{1/2}\leq\mu(q)^{1/2}\leq\epsilon/2$$

and similarly

$$\begin{aligned}|\mu(A\sum_{i\in J}p_i)|&\leq\sum_{i\in J}|\mu(Ap_i)|\leq\sum_{i\in J}^{n}\mu(I^*I)^{1/2}\mu((Ap_i)^*Ap_i)^{1/2}=\sum_{i\in J}\mu(p_iA^*Ap_i)^{1/2}\\&\leq\sum_{i\in J}\mu_{x_i}(p_iA^*Ap_i)^{1/2}=\sum_{i\in J}\|Ap_ix_i\|.\end{aligned}$$

Putting these estimates together gives our claim 3 for the neighborhood U introduced above. □

3.4 Completely positive maps

In Section 3.2.1 the general form of positive linear maps

$$f:\mathcal{A}\longrightarrow\mathbb{C}=M_1(\mathbb{C})$$

has been determined for C*-algebras. A natural extension of this problem is to look for the general form of positive linear maps with values in the space $M_k(\mathbb{C})$ of complex $k\times k$ matrices

$$F:\mathcal{A}\longrightarrow M_k(\mathbb{C}),\quad k>1,$$

or even more general for mappings with values in a C*-algebra of operators on a Hilbert space $\mathcal{H}$,

$$F : \mathcal{A} \longrightarrow \mathcal{B}, \quad \mathcal{B} \subset \mathcal{B}(\mathcal{H}) \tag{3.32}$$

extending the representation formula (3.19). This problem was investigated and solved by Stinespring in 1955 for the general case of C^*-algebras 23. It was found that one can arrive at a representation formula similar to (3.19) if one imposes on F a stronger positivity requirement, namely that of **complete positivity**.

3.4.1 Positive elements in $M_k(\mathcal{A})$

Let $\mathcal{A}$ be a C*-algebra. For $k = 1,2,\ldots$ introduce the space $M_k(\mathcal{A})$ of $k \times k$ matrices $[a_{ij}]$ with entries $a_{ij} \in \mathcal{A}$, $i,j = 1,\ldots,k$. In a natural way this space is a C*-algebra (with unit if $\mathcal{A}$ has a unit) (see 24).

According to our earlier discussion we call an element

$$a = [a_{ij}] = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \vdots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}, \quad a_{ij} \in \mathcal{A}$$

positive, $a \geq 0$ if, and only if, $a = c^*c$ for some $c \in M_k(\mathcal{A})$.

If $\mathcal{A}$ is a C*-algebra of operators on a Hilbert space $\mathcal{H}$ then $a = [a_{ij}] \in M_k(\mathcal{A})$ acts naturally on

$$\mathcal{H}^k = \left\{ \underline{\xi} = (\xi_1, \ldots, \xi_k) : \xi_j \in \mathcal{H}, j = 1, \ldots, k \right\} \tag{3.33}$$

according to the rule

$$([a_{ij}]\underline{\xi})_i = \sum_{j=1}^{k} a_{ij}\xi_j, \quad \underline{\xi} \in \mathcal{H}^k. \tag{3.34}$$

The space (3.33) is a Hilbert space with the scalar product

$$\langle \underline{\xi}, \underline{\eta} \rangle_{\mathcal{H}^k} = \sum_{j=1}^{k} \langle \xi_j, \eta_j \rangle_{\mathcal{H}}, \quad \forall\, \underline{\xi}, \underline{\eta} \in \mathcal{H}^k. \tag{3.35}$$

Positive elements in $M_k(\mathcal{A})$ are characterized by the following lemma (see 24):

Lemma 3.4.1 *The following conditions are equivalent for an element $a = [a_{ij}] \in M_k(\mathcal{A})$:*

(1) $a = b^* b$ for some $b \in M_k(\mathcal{A})$;

(2) $\langle \underline{\xi}, a\underline{\xi} \rangle_{\mathcal{H}^k} \geq 0$ for all $\underline{\xi} \in \mathcal{H}^k$;

(3) $a = [a_{ij}]$ is a sum of matrices of the form $[a_i^* a_j]$, $a_1, \dots, a_k \in M_k(\mathcal{A})$;

(4) For all $x_1, \dots, x_k \in \mathcal{A}$ one has

$$\sum_{i,j=1}^{k} x_i^* a_{ij} x_j \geq 0 \quad \text{in } \mathcal{A}.$$

Proof. (1) $\Rightarrow$ (3): If $a = b^*b$ for some $b \in M_k(\mathcal{A})$, then $a_{ij} = (b^*b)_{ij} = \sum_{m=1}^{k} b_{mi}^* b_{mj}$; $c_m = [b_{mi}^* b_{mj}] \in M_k(\mathcal{A})$ is of the claimed form and $a = \sum_{m=1}^{k} c_m$. Thus (3) holds.

(3) $\Rightarrow$ (4): If we know $a = [a_i^* a_j]$ for some $a_j \in \mathcal{A}$ and if any elements $x_1, \dots, x_k \in \mathcal{A}$ are given, then

$$\sum_{i,j=1}^{k} x_i^* a_{ij} x_j = \sum_{i,j=1}^{k} x_i^* a_i^* a_j x_j = \left(\sum_{i=1}^{k} a_i x_i\right)^* \left(\sum_{i=1}^{k} a_j x_j\right),$$

now $b = \sum_{i=1}^{k} a_i x_i \in \mathcal{A}$ and

$$\sum_{i,j=1}^{k} x_i^* a_{ij} x_j = b^* b,$$

hence this sum is positive.

(4) $\Rightarrow$ (2): If (4) holds, then by condition (3.14), for all $\boldsymbol{x} \in \mathcal{H}$ and all $x_i \in \mathcal{A}$,

$$\langle \boldsymbol{x}, \left(\sum_{i,j=1}^{k} x_i^* a_{ij} x_j\right) \boldsymbol{x} \rangle_{\mathcal{H}} \geq 0.$$

It follows

$$\langle \underline{\xi}, a\underline{\xi} \rangle_{\mathcal{H}^k} \geq 0$$

for all $\underline{\xi} \in \mathcal{H}^k$ which are of the form

$$\underline{\xi} = (x_1 \boldsymbol{x}, \dots, x_k \boldsymbol{x}), \quad \boldsymbol{x} \in \mathcal{H}\ x_j \in \mathcal{A}.$$

But this set equals $\mathcal{H}^k$, hence (2) holds.

(2) ⇒ (1): If (2) holds, the square root lemma (Theorem 21.5.1) implies that a has a positive square root $b = \sqrt{a} \in M_k(\mathcal{A})$ such that $a = b^2 = b^*b$ and (1) follows.

Note that (1) ⇒ (2) is trivial and also (3) ⇒ (1) is simple. If a is of the form $[a_i^* a_j]$, then

$$a = \begin{pmatrix} a_1^* a_1 & \cdots & a_1^* a_k \\ \vdots & \vdots & \vdots \\ a_k^* a_1 & \cdots & a_k^* a_k \end{pmatrix} = \begin{pmatrix} a_1^* & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_k^* & 0 & \cdots & 0 \end{pmatrix} \cdot \begin{pmatrix} a_1 & \cdots & a_k \\ 0 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} = b^* b,$$

hence (1). □

Elements in $M_k(\mathcal{A})$ which satisfy any of the 4 equivalent conditions of Lemma 3.4.1 are called **positive**.

Lemma 3.4.2 *Let $\mathcal{A}$ be a C^*-algebra of operators in a Hilbert space $\mathcal{H}$. Then, given any $a_1, \ldots, a_k \in \mathcal{A}$ one has for all $a \in \mathcal{A}$*

$$[(aa_i)^*(aa_j)] \leq \|a\|^2 [a_i^* a_j] \quad \text{in } \mathrm{M_k}(\mathcal{A}). \tag{3.36}$$

Proof. The matrix of operators $[(aa_i)^*(aa_j)]$ acts on the Hilbert space $\mathcal{H}^k$ according to (3.34) and for all $\underline{x} = (x_1, \ldots, x_k) \in \mathcal{H}^k$ we have

$$\langle \underline{x}, [(aa_i)^*(aa_j)]\underline{x} \rangle_{\mathcal{H}^k} = \left\| \sum_j a a_j x_j \right\|_{\mathcal{H}}^2 \leq \|a\|^2 \left\| \sum_j a_j x_j \right\|_{\mathcal{H}}^2 = \|a\|^2 \langle \underline{x}, [a_i^* a_j] \underline{x} \rangle_{\mathcal{H}^k},$$

thus (3.36) follows. □

3.4.2 Some basic properties of positive linear mappings

In the proof of the Stinespring factorization theorem for completely positive maps we need some basic properties of positive linear maps. These are briefly discussed here.

Let $\mathcal{A}$ and $\mathcal{B}$ be $*$-algebras. Recall: Elements a in $\mathcal{A}$ are called **positive** (more accurately, nonnegative), in symbols $a \geq 0$, if there is $b \in \mathcal{A}$ such that $a = b^* b$. A corresponding characterization of positive elements applies to $\mathcal{B}$. Furthermore, a linear mapping $T : \mathcal{A} \longrightarrow \mathcal{B}$ is called **positive** if, and only if, for all $a \in \mathcal{A}$ with $a \geq 0$ one has $T(a) \geq 0$ (in $\mathcal{B}$).

Knowing what positive elements are, we can define an **order** on $\mathcal{A}$ and $\mathcal{B}$: For $a_1, a_2 \in \mathcal{A}$ one says that a_1 is smaller or equal to a_2, in symbols $a_1 \leq a_2$, if, and only if, $a_2 - a_1 \geq 0$.

For any positive linear mapping $T : \mathcal{A} \longrightarrow \mathcal{B}$ the following holds:

$$a_1, a_2 \in \mathcal{A},\ a_1 \leq a_2 \ \Rightarrow\ T(a_1) \leq T(a_2)\,. \tag{3.37}$$

Positive linear maps $T : \mathcal{A} \longrightarrow \mathcal{B}$ satisfy the following important estimates:

Lemma 3.4.3 *Suppose that $\mathcal{A}$ is a C^*-algebra with unit I. Then any positive linear map $T : \mathcal{A} \longrightarrow \mathcal{B}$ satisfies*

$$T(x^* a^* a x) \leq \|a\|^2 \, T(x^* x) \quad \forall\, a, x \in \mathcal{A}\,. \tag{3.38}$$

In particular

$$T(a^*a) \leq \|a\|^2 T(I) \quad \forall\, a \in \mathcal{A}$$

and thus $T(I) = 0$ *implies* $T = 0$.

Proof. From (3.17) we know for all $a \in \mathcal{A}$

$$a^*a \leq \|a\|^2 I, \tag{3.39}$$

hence for all $x \in \mathcal{A}$,

$$x^*a^*ax \leq \|a\|^2 x^*x,$$

and thus for any positive linear mapping $T : \mathcal{A} \longrightarrow \mathcal{B}$ estimate (3.38) follows. If we choose $x = I$ we get the estimate for $T(a^*a)$. This estimate implies that T vanishes on all positive elements of $\mathcal{A}$ if $T(I) = 0$. Now observe that every $a \in \mathcal{A}$ can be written as

$$a = \frac{1}{2}(a + a^*) + \mathrm{i}\frac{1}{2\mathrm{i}}(\mathrm{a} - \mathrm{a}^*)$$

where the elements $a_r = \frac{1}{2}(a + a^*)$ and $a_i = \frac{1}{2\mathrm{i}}(a - a^*)$ are self-adjoint (Hermitian), i.e., $a^*_{r,i} = a_{r,i}$. From spectral theory it follows that every self-adjoint $b \in \mathcal{A}$ can be written as the difference of two positive elements in $\mathcal{A}$, $b = b_+ - b_-$ with $b_\pm \geq 0$. By linearity of T we conclude that T vanishes on all of $\mathcal{A}$.
□

3.4.3 Completely positive maps between C*-algebras

Suppose that a linear map $F : \mathcal{A} \longrightarrow \mathcal{B}$ between two C*-algebras $\mathcal{A}, \mathcal{B}$ is given. For $k = 1, 2, \ldots$ it induces a map

$$F_k : M_k(\mathcal{A}) \longrightarrow M_k(\mathcal{B}), \quad F_k([a_{ij}]) = [F(a_{ij})], \quad . \tag{3.40}$$

for all $a_{ij} \in \mathcal{A}$, $i, j = 1, 2, \ldots, k$.

Definition 3.4.4 *A linear map $F : \mathcal{A} \longrightarrow \mathcal{B}$ as above is called k-**positive** if, and only if, F_k is positive, i.e., if F_k maps positive elements of $M_k(\mathcal{A})$ to positive elements of $M_k(\mathcal{B})$, If F is k-positive for all $k \in \mathbb{N}$ then F is called **completely positive**.*

Remark 3.4.5 *In Physics literature the map F_k is usually written as*

$$F_k = I_k \otimes F : M_k(\mathbb{C}) \otimes \mathcal{A} \longrightarrow M_k(\mathbb{C}) \otimes \mathcal{B}. \tag{3.41}$$

Naturally, our characterization of positive elements in $M_k(\mathcal{A})$ of the previous subsection implies a characterization of k-positive and completely positive maps.

Corollary 3.4.6 *Let $F : \mathcal{A} \longrightarrow \mathcal{B}$ be as above. Then F is k-positive if, and only if,*

$$\forall_{x_i \in \mathcal{A}} \forall_{y_j \in \mathcal{B}} \sum_{i,j=1}^{k} y_i^* F(x_i^* x_j) y_j \geq 0 \quad \text{in } \mathcal{B}. \tag{3.42}$$

Proof. By condition (3) of Lemma 3.4.1 every positive element $[a_{ij}]$ in $M_k(\mathcal{A})$ is a sum elements of the form $[x_i^* x_j]$, $x_1, \ldots, x_k \in \mathcal{A}$; hence F is k-positive if, and only if, $[F(x_i^* x_j)]$ is positive in $M_k(\mathcal{B})$. According to Condition (4) of Lemma 3.4.1, this is the case if, and only if condition (3.42) holds. Thus we conclude. □

Corollary 3.4.7 *Let $F : \mathcal{A} \longrightarrow \mathcal{B}$ be as above with $\mathcal{B} = M_1(\mathbb{C}) = \mathbb{C}$. If F is positive, then F is completely positive.*

Proof. If $F : \mathcal{A} \longrightarrow \mathbb{C}$ is positive, then $F(b^*b) \geq 0$ for all $b \in \mathcal{A}$. Using the characterization (3.42) we show that F is k-positive for all k. For $y_j \in \mathbb{C}$ the sum in (3.42) can be written as

$$\sum_{i,j=1}^{k} y_i^* F(x_i^* x_j) y_j = F\left(\left(\sum_{i=1}^{k} y_i x_i\right)^* \left(\sum_{i=1}^{k} y_i x_i\right)\right) = F(b^*b).$$

Thus, by Corollary 3.4.6, F is k-positive and we conclude. □

Our first **example** of a completely positive map $F : \mathcal{A} \longrightarrow \mathcal{B}$: Any homomorphism F of C*-algebras is completely positive.

The proof is simple. Using Corollary 3.4.6 we show that a homomorphism of $*$-algebras is k-positive for every $k \in \mathbb{N}$. For all $x_i \in \mathcal{A}$ and all $y_j \in \mathcal{B}$ one has, using the properties of a homomorphism of $*$-algebras and Lemma 3.4.1,

$$\sum_{i,j=1}^{k} y_i^* F(x_i^* x_j) y_j = \sum_{i,j=1}^{k} y_i^* F(x_i)^* F(x_j) y_j = \left(\sum_{i=1}^{k} F(x_i) y_i\right)^* \left(\sum_{j=1}^{k} F(x_j) y_j\right)$$

which is certainly ≥ 0 in $\mathcal{B}$. Hence F is k-positive for every k and therefore completely positive.

Our next example is just a slight extension of the first. Let $\pi : \mathcal{A} \longrightarrow \mathcal{B}$ a $*$-homomorphism of $\mathcal{A}$ and V some element in $\mathcal{B}$; define $F : \mathcal{A} \longrightarrow \mathcal{B}$ by

$$F(a) = V^* \pi(a) V, \quad \forall\, a \in \mathcal{A}. \tag{3.43}$$

A similar calculation as above shows that F is a completely positive map. For all $x_i \in \mathcal{A}$ and all $y_j \in \mathcal{B}$ one has, using the properties of a homomorphism of $*$-algebras and Lemma 3.4.1,

$$\sum_{i,j=1}^{k} y_i^* F(x_i^* x_j) y_j = \sum_{i,j=1}^{k} y_i^* V^* \pi(x_i)^* \pi(x_j) V y_j =$$
$$\left(\sum_{i=1}^{k} \pi(x_i) V y_i \right)^* \left(\sum_{j=1}^{k} \pi(x_j) V y_j \right) \geq 0 \text{ in } \mathcal{B}.$$

Note that if V is not unitary then the map F of (3.43) is not a representation of $\mathcal{A}$.

3.4.4 Stinespring factorization theorem for completely positive maps

The Stinespring factorization theorem shows that essentially all completely positive maps are of the form (3.43). The proof is a straightforward extension of the proof for the GNS-construction.

We state and prove this result explicitly for the case where $\mathcal{A}$ and $\mathcal{B}$ are C^*-algebras of operators on a Hilbert space. The general case is given in 24.

Theorem 3.4.8 (Stinespring factorization theorem) *Let $\mathcal{A}$ be a C^*-algebra with unit I and $\mathcal{B} \subset \mathcal{B}(\mathcal{H})$ be a C^*-algebra of operators in a Hilbert space $\mathcal{H}$. Then for every completely positive map $f : \mathcal{A} \longrightarrow \mathcal{B}$ there exist a Hilbert space $\mathcal{K}_f$, a representation π_f of $\mathcal{A}$ in $\mathcal{K}_f$, and a bounded linear operator $V : \mathcal{H} \longrightarrow \mathcal{K}_f$ such that*

$$f(a) = V^* \pi_f(a) V \quad \forall\, a \in \mathcal{A}. \tag{3.44}$$

Furthermore, for all $\xi \in \mathcal{H}$, $\|V\xi\|_f = \left\|f(I)^{1/2}\xi\right\|_{\mathcal{H}}$.

Proof. Construction of $\mathcal{K}_f$: On the algebraic tensor product $\mathcal{A} \otimes \mathcal{H}$ define, for elements $\zeta = \sum_{i=1}^k a_i \otimes \xi_i, \chi = \sum_{j=1}^l b_j \otimes \eta_j$ in $\mathcal{A} \otimes \mathcal{H}$,

$$\langle \zeta, \chi \rangle_f = \sum_{i=1}^{k} \sum_{j=1}^{l} \langle \xi_i, f(a_i^* b_j) \eta_j \rangle_{\mathcal{H}}. \tag{3.45}$$

One verifies that this formula defines a sesquilinear form on $\mathcal{A} \otimes \mathcal{H}$. In particular, in the notation of Section 4.1,

$$\langle \zeta, \zeta \rangle_f = \sum_{i=1,j}^{k} \langle \xi_i, f(a_i^* a_j) \xi_j \rangle_{\mathcal{H}} = \langle \underline{\xi}, [f(a_i^* a_j)] \underline{\xi} \rangle_{\mathcal{H}^k} = \langle \underline{\xi}, f_k([a_i^* a_j]) \underline{\xi} \rangle_{\mathcal{H}^k}.$$

According to Lemma 3.4.1 the element $a = [a_i^* a_j] \in M_k(\mathcal{A})$ is positive and, since f is completely positive, f_k is a positive mapping from $M_k(\mathcal{A})$ into $M_k(\mathcal{B})$, hence $f_k([a_i^* a_j])$ is a positive matrix on $\mathcal{H}^k$ and we conclude $\langle \zeta, \zeta \rangle_f \geq 0$. Therefore the sesquilinear form (3.45) is positive semi-definite and hence it satisfies the Cauchy-Schwarz inequality

$$|\langle \zeta, \chi \rangle_f|^2 \leq \langle \zeta, \zeta \rangle_f \langle \chi, \chi \rangle_f.$$

We conclude that the kernel

$$I_f = \left\{ \zeta = \sum_{i=1}^{k} a_i \otimes \xi_i \in \mathcal{A} \otimes \mathcal{H} : \langle \zeta, \zeta \rangle_f = 0 \right\}$$

of this sesquilinear form is a linear subspace of $\mathcal{A} \otimes \mathcal{H}$. On the quotient space

$$\mathcal{K}_f^0 = \mathcal{A} \otimes \mathcal{H} / I_f = \left\{ [\zeta]_f = \zeta + I_f : \zeta \in \mathcal{A} \otimes \mathcal{H} \right\}$$

the formula

$$\langle [\zeta]_f, [\chi]_f \rangle = \langle \zeta, \chi \rangle_f$$

then defines an inner product and thus the completion $\mathcal{K}_f$ of $\mathcal{K}_f^0$ with respect to the norm defined by this inner product is a Hilbert space.

Construction of π_f: For $[\zeta]_f \in \mathcal{K}_f^0$, $\zeta = \sum_{i=1}^k a_i \otimes \xi_i \in \mathcal{A} \otimes \mathcal{H}$ define

$$\pi_f^0(a)[\zeta]_f = [\sum_{i=1}^k aa_i \otimes \xi_i]_f \quad \forall a \in \mathcal{A}. \tag{3.46}$$

At first we calculate

$$\langle \pi_f^0(a)[\zeta]_f, \pi_f^0(a)[\zeta]_f \rangle = \langle [\sum_{i=1}^k aa_i \otimes \xi_i]_f, [\sum_{j=1}^k aa_j \otimes \xi_j]_f \rangle = \langle \sum_{i=1}^k aa_i \otimes \xi_i,$$
$$\sum_{j=1}^k aa_j \otimes \xi_j \rangle_f = \sum_{i,j=1}^k \langle \xi_i, f((aa_i)^*(aa_j))\xi_j \rangle_{\mathcal{H}} = \langle \underline{\xi}, f_k([(aa_i)^*(aa_j)])\underline{\xi} \rangle_{\mathcal{H}^k}$$

where $\underline{\xi} = (\xi_1, \ldots, \xi_k) \in \mathcal{H}^k$. Lemma 3.4.2 says

$$[a_i^* a^* a a_j] \leq \|a\|^2 [a_i^* a_j].$$

Since f is completely positive, the map f_k is positive and therefore

$$f_k([a_i^* a^* a a_j]) \leq \|a\|^2 f_k([a_i^* a_j]).$$

We conclude

$$\langle \underline{\xi}, f_k([(aa_i)^*(aa_j)])\underline{\xi} \rangle_{\mathcal{H}^k} \leq \|a\|^2 \langle \underline{\xi}, f_k([a_i^* a_j])\underline{\xi} \rangle_{\mathcal{H}^k}$$

and hence

$$\langle \pi_f^0(a)[\zeta]_f, \pi_f^0(a)[\zeta]_f \rangle_f \leq \|a\|^2 \langle [\zeta]_f, [\zeta]_f \rangle_f . \tag{3.47}$$

This estimate shows first that $\pi_f^0(a)$ is well defined (i.e., $\pi_f^0(a)$ is indeed a map between equivalence classes and does not depend on the representatives of the equivalence classes which are used in its definition).

Now, using (3.46), it is a straightforward calculation to show that $\pi_f^0 : \mathcal{K}_f^0 \longrightarrow \mathcal{K}_f^0$ is linear. Then (3.47) implies that this map is bounded and $\left\| \pi_f^0(a) \right\| \leq \|a\|$.

From the definition (3.46) it is immediate that π_f^0 satisfies

$$\pi_f^0(ab) = \pi_f^0(a)\pi_f^0(b) \quad \forall\, a, b \in \mathcal{A}.$$

In order to show

$$\pi_f^0(a^*) = \pi_f^0(a)^* \quad \forall\, a \in \mathcal{A}$$

we calculate as follows: For ζ, χ as in (3.45) and $a \in \mathcal{A}$, using (3.46),

$$\begin{aligned}&\langle \pi_f^0(a)^*[\zeta]_f, [\chi]_f\rangle = \langle [\zeta]_f, \pi_f^0(a)[\chi]_f\rangle = \sum_{i=1}^{k}\sum_{j=1}^{l}\langle \xi_i, f(a_i^* a b_j)\eta_j\rangle_{\mathcal{H}} = \\ &\sum_{i=1}^{k}\sum_{j=1}^{l}\langle \xi_i, f((a^* a_i)^* b_j)\eta_j\rangle_{\mathcal{H}} = \langle [\sum_{i=1}^{k}(a^* a_i)\otimes \xi_i]_f, [\sum_{j=1}^{l} b_j \otimes \eta_j]_f\rangle \\ &= \langle \pi_f^0(a^*)[\zeta]_f, [\chi]_f\rangle .\end{aligned}$$

Since this identity is true for all $[\zeta]_f, [\chi]_f \in \mathcal{K}_f^0$ we conclude.

This establishes that π_f^0 is a representation of $\mathcal{A}$ on $\mathcal{K}_f^0$.

By continuity π_f^0 has a unique extension to a representation π_f on the completion $\mathcal{K}_f$ of $\mathcal{K}_f^0$.

Construction of V: Define $V^0 : \mathcal{H} \longrightarrow \mathcal{K}_f^0$ by

$$V^0\xi = [I \otimes \xi]_f \quad \forall\, \xi \in \mathcal{H}. \tag{3.48}$$

An easy calculation shows that V^0 is linear. Now calculate

$$\langle V^0\xi, V^0\xi\rangle = \langle I_n \otimes \xi, I \otimes \xi\rangle_f = \langle \xi, f(I^* I)\xi\rangle_{\mathcal{H}} \le \|f(I)\|\,\langle \xi, \xi\rangle_{\mathcal{H}} \quad \forall \xi \in \mathcal{H}.$$

This shows that V^0 is bounded. Since $f(I^* I) = f(I)$ is positive we know $f(I) = (\sqrt{f(I)})^2$ and thus

$$\left\|V^0\xi\right\| \le \left\|\sqrt{f(I)}\xi\right\|_{\mathcal{H}} \quad \xi \in \mathcal{H}. \tag{3.49}$$

In the case of $f(I) = I$, V^0 is thus an isometry.

Now for $a \in \mathcal{A}$ and $\xi \in \mathcal{H}$ the identity

$$\pi_f^0(a)V^0\xi = [a \otimes \xi]_f$$

follows. We deduce

$$\pi_f^0(\mathcal{A})V^0\mathcal{H} = \mathcal{K}_f^0. \tag{3.50}$$

V^0 is extended by continuity to a bounded linear operator $V : \mathcal{H} \longrightarrow \mathcal{K}_f$ and then the last condition reads

$$[\pi_f(\mathcal{A})V\mathcal{H}] = \mathcal{K}_f, \tag{3.51}$$

where $[\cdots]$ denotes the closure of the linear hull of $\cdots$ in $\mathcal{K}_f$.

Now all preparations for the proof of the Stinespring factorization formula have been done. For $\eta, \xi \in \mathcal{H}$ one finds for all $a \in \mathcal{A}$,

$$\langle V^0\eta, \pi_f^0(a)V^0\xi\rangle = \langle [I \otimes \eta]_f, [a \otimes \xi]_f\rangle = \langle I \otimes \eta, a \otimes \xi\rangle_f = \langle \eta, f(I^*a)\xi\rangle_{\mathcal{H}}$$

and therefore $f(a) = (V^0)^*\pi_f^0(a)V^0$ for all $a \in \mathcal{A}$. By continuous extension the Stinespring factorization formula (3.44) follows. □

Corollary 3.4.9 (Uniqueness under minimality condition) *Let $f : \mathcal{A} \longrightarrow \mathcal{B}$ be a completely positive map as in Theorem 3.4.8 and let*

$$f(a) = U^*\pi(a)U, \quad a \in \mathcal{A} \tag{3.52}$$

be a Stinespring factorization of f with a representation π of $\mathcal{A}$ in a Hilbert space $\mathcal{K}$ and a bounded linear operator $U : \mathcal{H} \longrightarrow \mathcal{K}$. If this factorization satisfies the minimality condition

$$[\pi(\mathcal{A})U\mathcal{H}] = \mathcal{K} \tag{3.53}$$

then, up to a unitary transformation, it is the factorization constructed in Theorem 3.4.8.

Proof. We begin by defining a linear operator $W_0 : \pi_f(\mathcal{A})V\mathcal{H} \longrightarrow \pi(\mathcal{A})U\mathcal{H}$ by the formula

$$W_0\left(\sum_{i=1}^{k}\pi_f(x_i)V\xi_i\right) = \sum_{i=1}^{k}\pi(x_i)U\xi_i \quad x_i \in \mathcal{A},\, \xi_i \in \mathcal{H}. \tag{3.54}$$

Now calculate the inner product of these images in $\mathcal{K}$:

$$\langle W_0\sum_{i=1}^{k}\pi_f(x_i)V\xi_i, W_0\sum_{j=1}^{l}\pi_f(y_i)V\eta_j\rangle_{\mathcal{K}} = \langle\sum_{i=1}^{k}\pi(x_i)U\xi_i, \sum_{j=1}^{l}\pi(y_j)U\eta_j\rangle_{\mathcal{K}} =$$
$$\sum_{i=1}^{k}\sum_{j=1}^{l}\langle U\xi_i, \pi(x_i)^*\pi(y_j)U\eta_j\rangle_{\mathcal{K}} = \sum_{i=1}^{k}\sum_{j=1}^{l}\langle\xi_i, U^*\pi(x_i^*y_j)U\eta_j\rangle_{\mathcal{H}} =$$
$$\sum_{i=1}^{k}\sum_{j=1}^{l}\langle\xi_i, f(x_i^*y_j)\eta_j\rangle_{\mathcal{H}} = \sum_{i=1}^{k}\sum_{j=1}^{l}\langle\xi_i, V^*\pi_f(x_i)^*\pi_f(y_j)V\eta_j\rangle_{\mathcal{H}}$$
$$= \langle\sum_{i=1}^{k}\pi_f(x_i)V\xi_i, \sum_{j=1}^{l}\pi_f(y_i)V\eta_j\rangle.$$

We conclude that $W_0 : \pi_f(\mathcal{A})V\mathcal{H} \longrightarrow \pi(\mathcal{A})U\mathcal{H}$ is isometric and thus extends by continuity to a unitary operator

$$W : [\pi_f(\mathcal{A})V\mathcal{H}] \longrightarrow [\pi(\mathcal{A})U\mathcal{H}],$$

i.e., because of the minimality condition to a unitary operator $W : \mathcal{K}_f \longrightarrow \mathcal{K}$. From the above definition the relations

$$W\pi_f(\cdot)W^* = \pi(\cdot), \quad WV = U$$

follow immediately. □

Corollary 3.4.10 *Let $f : \mathcal{A} \longrightarrow \mathcal{B}$ be a completely positive map as above. Then the inequality*

$$f(a)^*f(a) \le \|f(I)\|\, f(a^*a) \tag{3.55}$$

holds for all $a \in \mathcal{A}$.

Proof. By Theorem 3.4.8 f has a Stinespring factorization $f(a) = V^*\pi(a)V$ and thus

$$\begin{aligned} f(a)^*f(a) &= (V^*\pi(a)V)^*V^*\pi(a)V = V^*\pi(a)^*VV^*\pi(a)V \\ &\le \|V\|^2 V^*\pi(a)^*\pi(a)V = \|V\|^2 f(a^*a). \end{aligned}$$

The estimate $\|V\xi\|_f \le \left\|f(I)^{1/2}\xi\right\|_{\mathcal{H}}$ for $\xi \in \mathcal{H}$ implies $\|V\| \le \left\|f(I)^{1/2}\right\|$ or $\|V\|^2 \le \left\|f(I)^{1/2}\right\|^2 = \|f(I)\|$. □

3.4.5 Completely positive mappings on $\mathcal{B}(\mathcal{H})$

Theorem 3.1.5 determines the structure of representations of the C*-algebra $\mathcal{B}(\mathcal{H})$. If we combine this result with Stinespring's factorization theorem we arrive at a more concrete form of completely positive mappings on $\mathcal{B}(\mathcal{H})$.

Suppose that operators $a_1, \ldots, a_m \in \mathcal{B}(\mathcal{H})$ are given. Define $f_m : \mathcal{B}(\mathcal{H}) \longrightarrow \mathcal{B}(\mathcal{H})$ by

$$f_m(a) = \sum_{j=1}^{m} a_j^* a a_j.$$

Using Corollary 3.4.6 one shows easily that this mapping is completely positive. Next suppose that we are given a sequence $\{a_j\} \subset \mathcal{B}(\mathcal{H})$ of operators for which there is a positive operator B such that for all $m \in \mathbb{N}$

$$S_m = \sum_{j=1}^{m} a_j^* a_j \le B. \tag{3.56}$$

Thus we get a sequence of completely positive mappings f_m on $\mathcal{B}(\mathcal{H})$ which converges.

Lemma 3.4.11 (completely positive maps on $\mathcal{B}(\mathcal{H})$) *For a sequence of operators $a_j \in \mathcal{B}(\mathcal{H})$ which satisfies (3.56) the series*

$$f(a) = \sum_{j=1}^{\infty} a_j^* a a_j, \quad a \in \mathcal{B}(\mathcal{H}) \text{ fixed} \tag{3.57}$$

converges in the ultraweak operator topology on $\mathcal{B}(\mathcal{H})$ and defines a completely positive mapping which satisfies

$$f(I) \leq B. \tag{3.58}$$

Proof. Recall that every $a \in \mathcal{B}(\mathcal{H})$ has a representation as a complex linear combination of four positive elements. Thus it suffices to show this convergence for positive $a \in \mathcal{B}(\mathcal{H})$. In this case we know that $0 \leq a \leq \|a\| I$ and it follows for arbitrary $x \in \mathcal{H}$ and $m \in \mathbb{N}$

$$\sum_{j=1}^{m} \left\| a^{1/2} a_j x \right\|^2 = \sum_{j=1}^{m} \langle x, a_j^* a a_j x \rangle \leq \sum_{j=1}^{m} \|a\| \langle x, a_j^* a_j x \rangle \leq \|a\| \langle x, Bx \rangle,$$

hence this monotone increasing sequence is bounded from above and thus it converges:

$$\sum_{j=1}^{\infty} \langle x, a_j^* a a_j x \rangle \leq \|a\| \langle x, Bx \rangle.$$

The polarization identity (14.5) implies that for arbitrary $x, y \in \mathcal{H}$ the numerical series

$$\sum_{j=1}^{\infty} \langle x, a_j^* a a_j y \rangle$$

converges. This shows that the series (3.57) converges in the weak operator topology and thus defines $f(a) \in \mathcal{B}(\mathcal{H})$. Since the partial sums of the series considered are bounded we also have ultraweak convergence by Lemma 2.3.1 (on bounded sets both topologies coincide).

Finally we show that f is completely positive by showing that it is k-positive for every $k \in \mathbb{N}$. According to Corollary 3.4.6 choose arbitrary $A_1, \ldots, A_k, B_1, \ldots, B_k \in \mathcal{B}(\mathcal{H})$. For every $x \in \mathcal{H}$ we find

$$\langle x, \sum_{i,j=1}^{k} B_i^* f(A_i^* A_j) B_j x \rangle = \lim_{m\to\infty} \langle x, \sum_{i,j=1}^{k} B_i^* f_m(A_i^* A_j) B_j x \rangle \geq 0$$

since f_m is k-positive. We conclude that $\sum_{i,j=1}^{k} B_i^* f(A_i^* A_j) B_j \geq 0$ in $\mathcal{B}(\mathcal{H})$ and hence f is k-positive. □

Combining Stinespring's factorization theorem with Theorem 3.1.5 shows that essentially all completely positive mappings on $\mathcal{B}(\mathcal{H})$ are of the form (3.57).

Theorem 3.4.12 *Every completely positive mapping $f : \mathcal{B}(\mathcal{H}) \longrightarrow \mathcal{B}(\mathcal{H})$ is of the form*

$$f(a) = V_0^* \pi_0(a) V_0 + \sum_{j\in J} a_j^* a a_j, \quad a \in \mathcal{B}(\mathcal{H}) \tag{3.59}$$

with the following specifications:
π_0 is a representation of the quotient algebra$\mathcal{B}(\mathcal{H})/\mathcal{B}_c(\mathcal{H})$ on a Hilbert space $\mathcal{H}_0$ (hence $\pi_0(b) = 0$ for all $b \in \mathcal{B}_c(\mathcal{H})$), V_0 is a bounded linear operator $\mathcal{H} \longrightarrow \mathcal{H}_0$, J is a finite or countable index set, and $a_j \in \mathcal{B}(\mathcal{H})$ satisfy

$$\sum_{j\in J} a_j^* a_j \leq f(I). \tag{3.60}$$

Proof. Theorem 3.4.8 implies that a given completely positive mapping on $\mathcal{B}(\mathcal{H})$ is of the form (3.44) with a representation π of $\mathcal{B}(\mathcal{H})$ on a Hilbert space $\mathcal{K}$ and a bounded linear operator $V : \mathcal{H} \longrightarrow \mathcal{K}$, i. e.,

$$f(a) = V^*\pi(a)V, \quad a \in \mathcal{B}(\mathcal{H}).$$

The general form of representations of $\mathcal{B}(\mathcal{H})$ has been determined in Theorem 3.1.5. According to this result π has the direct sum decomposition (J some finite or countable index set)

$$\pi = \pi_0 \oplus \bigoplus_{j\in J} \pi_j$$

where for $j \in J$ π_j is the identity representation $\pi_j(a) = a$ in the Hilbert space $\mathcal{H}_j = \mathcal{H}$ and where π_0 is a representation of the quotient algebra $\mathcal{B}(\mathcal{H})/\mathcal{B}_c(\mathcal{H})$. This means that there is a unitary operator U from the representation space $\mathcal{K}$ of π onto the direct sum of the Hilbert spaces of these representations

$$\mathcal{H}_0 \oplus \bigoplus_{j\in J} \mathcal{H}_j.$$

Denote the projectors of $U\mathcal{K}$ onto $\mathcal{H}_0$ by P_0, respectively onto $\mathcal{H}_j$ by P_j, $j \in J$. Thus $f(a) = V^*\pi(a)V$ takes the form

$$f(a) = V_0^*\pi_0(a)V_0 + \sum_{j\in J} a_j^* a a_j, \quad a \in \mathcal{B}(\mathcal{H})$$

where $V_0 = P_0UV$ and $a_j = P_jUV$ for $j \in J$. Since $\pi_0(I) \geq 0$ one has the bound (3.60). □

Chapter 4

Positive mappings in quantum physics

Abstract: This chapter offers some results which will help to understand some foundational aspects of quantum mechanics. It relies on some results presented in the last chapter. The first section discusses the general form of σ-additive probability measures on the complete lattice of orthogonal projections on a Hilbert space (Gleason's theorem) and its variations. In quantum mechanics and in quantum information theory quantum channels or quantum operations are defined mathematically as completely positive maps between density operators which do not increase the trace (see for instance 18). Thus in the next section we determine the general form of quantum operations on a separable Hilbert space, i.e., we prove Kraus' first representation theorem

for operations. Usually quantum information theory studies systems of some finite dimension n and then density operators are just positive $n \times n$ matrices with complex coefficients which have trace 1. In this context the relevant C*-algebra is just the space $M_n(\mathbb{C})$ of all $n \times n$ matrices with complex entries, for some $n \in \mathbb{N}$. Therefore in the last section we determine the general form of completely positive maps for these algebras (Choi's results). Of course, this is a special case of Stinespring factorization theorem, but some important aspects are added.

4.1 Gleason's theorem

In Theorem 2.2.2 we learned that the continuous linear functionals on the space of all compact operators on a separable Hilbert space $\mathcal{H}$ are given by trace class operators according to the Trace Formula (2.10). There is a profound related result due to A. Gleason which roughly says that this trace formula holds when we start with a countably additive probability measure on the projections of $\mathcal{H}$ instead of a continuous linear functional on the compact operators on $\mathcal{H}$ (Recall that all finite dimensional projections belong to the space of compact operators).

Gleason's result 11 is very important for the (mathematical) foundation of quantum mechanics. A historical perspective and some key ideas related to this work are presented in 7.

Theorem 4.1.1 (Gleason's theorem) *Let μ be a countable additive probability measure on the projections of a separable Hilbert space $\mathcal{H}$ of dimension greater than 2. Then there is a unique nonnegative trace class operator W of trace 1 such that for every projection P on $\mathcal{H}$ one has*

$$\mu(P) = \mathrm{Tr}(WP). \tag{4.1}$$

The original proof by Gleason relies on methods not related to topics presented in this book. And this proof is quite long. Accordingly we do not present it here. Instead we discuss a weakened version due to P. Busch 6. A proof of Gleason's original result which is more easily accessible is 20.

In 16 the physical meaning of **effects** and their mathematical realization is explained. Denote by $\mathcal{E}(\mathcal{H})$ the set of all effects on the separable Hilbert space $\mathcal{H}$, i.e., the set of all $A \in \mathcal{B}(\mathcal{H})$ which satisfy $0 \leq A \leq I$.

Definition 4.1.2 *A generalized probability measure on all effects on a seperable Hilbert space $\mathcal{H}$ is a function $\mu : \mathcal{E}(\mathcal{H}) \longrightarrow \mathbb{R}$ which satisfies*

1) $0 \leq \mu(E) \leq 1$ *for all* $E \in \mathcal{E}(\mathcal{H})$,

2) $\mu(I) = 1$,

3) *for any sequence* $(E_j) \subset \mathcal{E}(\mathcal{H})$ *such that* $\sum_j E_j \leq I$ *one has*

$$\mu(\sum_j E_j) = \sum_j \mu(E_j).$$

Similar to Gleason's result one would like to know the general form of generalized probability measures on effects. It turns out that the analysis in this case is much simpler, mainly due to the fact that now a more or less standard extension in the ordered vector space of self-adjoint elements in $\mathcal{B}(\mathcal{H})$

$$\mathcal{B}_s(\mathcal{H}) = \mathcal{B}(\mathcal{H})_+ - \mathcal{B}(\mathcal{H})_+$$

which is generated by the positive elements, is possible.

Lemma 4.1.3 *Any generalized probability measure* μ *on* $\mathcal{E}(\mathcal{H})$ *is the restriction of a positive linear functional* $f : \mathcal{B}(\mathcal{H}) \longrightarrow \mathbb{C}$ *to the set of effects* $\mathcal{E}(\mathcal{H})$: $\mu = f|_{\mathcal{E}(\mathcal{H})}$.

Proof. Because of the defining conditions 1) and 3) a generalized probability measure μ on effects is monotone, i.e., if $E, F \in \mathcal{E}(\mathcal{H})$ satisfy $E \leq F$ then $\mu(E) \leq \mu(F)$. If $E \in \mathcal{E}(\mathcal{H})$ and $n \in \mathbb{N}$ are given, condition 3) implies $\mu(E) = n\mu(\frac{1}{n}E)$, since $E = \frac{1}{n}E + \cdots + \frac{1}{n}E$ (n summands). Next suppose $m, n \in \mathbb{N}$ are given and $m \leq n$ so that $m/n \leq 1$, thus $\frac{m}{n}E \in \mathcal{E}(\mathcal{H})$ and the relation $\mu(\frac{1}{n}E) = \frac{1}{n}\mu(E)$ implies ($\cdots$ means m summands)

$$\frac{m}{n}\mu(E) = m\mu(\frac{1}{n}E) = \mu(\frac{1}{n}E + \cdots + \frac{1}{n}E) = \mu(\frac{m}{n}E),$$

and therefore $\mu(qE) = q\mu(E)$ for all rational numbers $q \in [0,1]$.

If $0 < r < 1$ is any real number there are sequences of rational numbers $q_j, p_j \in (0,1)$ such that $p_j \downarrow r$ and $q_j \uparrow r$ and for all $j \in \mathbb{N}$, $0 < q_j \leq r \leq p_j < 1$. Then we know $q_j E \leq rE \leq p_j E$ and therefore by monotonicity of μ

$$q_j \mu(E) = \mu(q_j E) \leq \mu(rE) \leq \mu(p_j E) = p_j \mu(E).$$

In the limit $j \to \infty$ we thus get $r\mu(E) \leq \mu(rE) \leq r\mu(E)$. This implies $\mu(rE) = r\mu(E)$ for all $E \in \mathcal{E}(\mathcal{H})$ and all $r \in [0,1]$.

Next suppose that $A \in \mathcal{B}(\mathcal{H})$ is given satsfying $0 \leq A$ but not $A \leq I$. Then there is $r \geq 1$ such that $E = \frac{1}{r}A \in \mathcal{E}(\mathcal{H})$. But clearly r and E are not unique. If we have $A = r_1 E_1 = r_2 E_2$ we can assume that $r_2 > r_1 \geq 1$ so that $0 < \frac{r_1}{r_2} < 1$. It follows $\mu(E_2) = \mu(\frac{r_1}{r_2}E_1) = \frac{r_1}{r_2}\mu(E_1)$ or $r_1\mu(E_1) = r_2\mu(E_2)$. This allows to define $\mu_1 : \mathcal{B}(\mathcal{H})_+ \longrightarrow \mathbb{R}$ by $\mu_1(A) = r\mu(E)$ whenever $A = rE$ with $E \in \mathcal{E}(\mathcal{H})$ and $r \geq 1$.

Clearly μ_1 is positive homogenous on the convex cone $\mathcal{B}(\mathcal{H})_+$ of nonnegative bounded linear operators on $\mathcal{H}$.

In order to show that μ_1 is additive on $\mathcal{B}(\mathcal{H})_+$ take $A, B \in \mathcal{B}(\mathcal{H})_+$. Then there is $r > 1$ such $(A+B)/r \in \mathcal{E}(\mathcal{H})$. The definition of μ_1 gives $\mu_1(A+B) = r\mu(\frac{1}{r}(A+B)) = r\mu(\frac{1}{r}A) + r\mu(\frac{1}{r}B) = \mu_1(A) + \mu_1(B)$.

Altogether we have shown that μ_1 is an additive positive homogeneous function on the convex cone $\mathcal{B}(\mathcal{H})_+$. Thus, according to a standard procedure in the theory of ordered vector spaces, the functional μ_1 can be extended to a linear functional μ_2 on $\mathcal{B}_s(\mathcal{H})$. If $C = A - B \in \mathcal{B}(\mathcal{H})_+ - \mathcal{B}(\mathcal{H})_+$ define $\mu_2(C) = \mu_1(A) - \mu_1(B)$. It is easy to see that μ_2 is well defined. If C is also represented as $A' - B' \in \mathcal{B}(\mathcal{H})_+ - \mathcal{B}(\mathcal{H})_+$, then it follows $\mu_1(A) - \mu_1(B) = \mu_1(A') - \mu_1(B')$, since $A' + B = A + B'$ implies $\mu_1(A') + \mu_1(B) = \mu_1(A' + B) = \mu_1(A + B') = \mu_1(A) + \mu_1(B')$. In order to show that $\mu_2 : \mathcal{B}_s(\mathcal{H}) \longrightarrow \mathbb{R}$ is additive take $C = A - B$ and $C' = A' - B'$ in $\mathcal{B}(\mathcal{H})_+ - \mathcal{B}(\mathcal{H})_+$ and calculate $\mu_2(C + C') = \mu_2(A - B + A' - B') = \mu_2(A + A' - (B + B')) = \mu_1(A + A') - \mu_1(B + B') = \mu_1(A) + \mu_1(A') - \mu_1(B) - \mu_1(B') = \mu_2(A - B) + \mu_1(A' - B') = \mu_2(C) + \mu_2(C')$.

Clearly, since μ_1 is positive homogeneous, so is μ_2. Next suppose $\lambda < 0$ and $C = A - B \in \mathcal{B}(\mathcal{H})_+ - \mathcal{B}(\mathcal{H})_+$ are given. Then $\lambda C = (-\lambda)B - (-\lambda)A \in \mathcal{B}(\mathcal{H})_+ - \mathcal{B}(\mathcal{H})_+$ and so $\mu_2(\lambda C) = \mu_1((-\lambda)B) - \mu_1((-\lambda)A) = (-\lambda)\mu_1(B) - (-\lambda)\mu_1(A) = \lambda(\mu_1(A) - \mu_1(B)) = \lambda\mu_2(C)$.

It follows that $\mu_2 : \mathcal{B}_s(\mathcal{H}) \longrightarrow \mathbb{R}$ is a positive linear functional which agrees with μ on $\mathcal{E}(\mathcal{H})$. Since $\mathcal{B}(\mathcal{H}) = \mathcal{B}_s(\mathcal{H}) + \mathrm{i}\mathcal{B}_s(\mathcal{H})$ the real linear functional μ_2 is extended to a complex linear functional $f : \mathcal{B}(\mathcal{H}) \longrightarrow \mathbb{C}$ by setting for $A = a + \mathrm{ib} \in \mathcal{B}_s(\mathcal{H}) + \mathrm{i}\mathcal{B}_s(\mathcal{H})$, $f(A) = \mu_2(a) + \mathrm{i}\bar{}_2(\mathrm{b})$. A simple calculation shows that f is indeed complex linear on $\mathcal{B}(\mathcal{H})$ and by construction $\mu = f|_{\mathcal{E}(\mathcal{H})}$. □

Note that in the proof of this lemma Condition 2) has not been used and Condition 3) has been used only for finitely many effects.

Theorem 4.1.4 (Busch) *Any generalized probability measure μ on the set of effects $\mathcal{E}(\mathcal{H})$ of a separable Hilbert space $\mathcal{H}$ is of the form*

$$\mu(E) = \mathrm{Tr}(WE) \quad \textit{for all } E \in \mathcal{E}(\mathcal{H})$$

for some density operator W.

Proof. According to the extension lemma 4.1.3 any generalized probability measure μ on the set of effects is the restriction to this set of a positive linear functional f on $\mathcal{B}(\mathcal{H})$. Such functionals are continuous according to Proposition 3.2.2. Now, since projections are (special) effects, condition 3) says that the functional f is completely additive (see (3.29)). Hence we conclude by Theorem 3.3.3. □

4.2 Kraus form of quantum operations

A quantum mechanical systems undergoes various types of transformations, for instance symmetry transformations, time evolution, and transient interactions with an environment for measurement purposes. These transformations are described by the concept of a quantum operation and the nature of these mappings has been discussed since about 50 years starting with a paper by E. C. G. Sudarshan et al. in 1961.

A mathematically rigorous and comprehensive study of quantum operations has been published by K. Kraus in 1983 in 16. Starting from first (physical) principles it is argued that **quantum operations** are given mathematically by linear mappings

$$\phi : \mathcal{B}_1(\mathcal{H}) \longrightarrow \mathcal{B}_1(\mathcal{H})$$

of the space of trace class operators on a (separable) Hilbert space $\mathcal{H}$ into itself which are completely positive and satisfy

$$\mathrm{Tr}(\phi(W)) \leq 1 \tag{4.2}$$

for all $W \in \mathcal{B}_1(\mathcal{H})$ with $W \geq 0$ and $\mathrm{Tr}(W) = 1$, i.e., for all density matrices on $\mathcal{H}$.

In Definition 3.4.4 we had defined completely positive mappings as mappings between C*-algebras which satisfy certain positivity conditions. Clearly $\mathcal{B}_1(\mathcal{H})$ is not a C*-algebra with unit (if $\mathcal{H}$ is not finite dimensional) but it is a two-sided ideal in the C*-algebra $\mathcal{B}(\mathcal{H})$. Therefore these positivity conditions can be formulated for $\mathcal{B}_1(\mathcal{H})$ in the same way as for the C*-algebra $\mathcal{B}(\mathcal{H})$. And it is in this sense that we understand complete positivity for a map $\phi : \mathcal{B}_1(\mathcal{H}) \longrightarrow \mathcal{B}_1(\mathcal{H})$, i.e., ϕ is completely positive if, and only if, it is k-positive for $k = 1, 2, \ldots$. However in the characterization of positive elements in $\mathcal{B}_1(\mathcal{H})$ there is an important difference to the characterization of positive elements in a C*-algebra. According to the spectral representation of trace class operators (Theorem 2.1.9) $T \in \mathcal{B}_1(\mathcal{H})$ is positive if, and only if, $T = \tau^*\tau$ for some $\tau \in \mathcal{B}_2(\mathcal{H})$ (not in $\mathcal{B}_1(\mathcal{H})$). The characterization of positive elements in $M_k(\mathcal{B}_1(\mathcal{H}))$ for $k \geq 2$ is addressed explicitly later (see Lemma 4.2.4).

4.2.1 Operations and effects

Lemma 4.2.1 *If a positive linear map $\phi : \mathcal{B}_1(\mathcal{H}) \longrightarrow \mathcal{B}_1(\mathcal{H})$ satisfies (4.2), then it is continuous with respect to the trace norm:*

$$\|\phi(T)\|_1 \leq C\,\|T\|_1, \quad C = \sup \mathrm{Tr}(\phi(W)) \leq 1 \tag{4.3}$$

where the sup *is taken over all density matrices W on $\mathcal{H}$.*

Proof. By (4.2) we obviously have that $C = \sup \mathrm{Tr}(\phi(W)) \leq 1$ where the sup is taken over all density matrices. Write $T = T^* \in \mathcal{B}_1(\mathcal{H})$ as $T = T_+ - T_-$ where $T_\pm = (|T| \pm T)/2$. We can assume that $\mathrm{Tr}(T_\pm) > 0$. Then $W_\pm = \frac{1}{\mathrm{Tr}(T_\pm)} T_\pm$ are density matrices and it follows that

$$T = \mathrm{Tr}(T_+)W_+ - \mathrm{Tr}(T_-)W_-$$

and thus $\phi(T) = \mathrm{Tr}(T_+)\phi(W_+) - \mathrm{Tr}(T_-)\phi(W_-)$. According to (2.12) the trace norm of $\phi(T)$ can be calculated as

$$\|\phi(T)\|_1 = \sup_{\|B\|=1} |\mathrm{Tr}(B\phi(T))|.$$

Insert the above expression for $\phi(T)$ and estimate as follows:

$$|\mathrm{Tr}(\mathrm{Tr}(T_+)B\phi(W_+)) - \mathrm{Tr}(T_-)B\phi(W_-))| \leq$$
$$\mathrm{Tr}(T_+)|\mathrm{Tr}(B\phi(W_+))| + \mathrm{Tr}(T_-)|\mathrm{Tr}(B\phi(W_-))|$$

Since ϕ is positive we know

$$|\mathrm{Tr}(B\phi(W_\pm))| \leq \|B\|\,\|\phi(W_\pm)\|_1 = \|B\|\,\mathrm{Tr}(\phi(W_\pm)) \leq \|B\|\,C$$

and thus

$$\|\phi(T)\|_1 \leq \mathrm{Tr}(T_+)C + \mathrm{Tr}(T_-)C = C\,\|T\|_1\,.$$

□

The adjoint ϕ^* of an operation ϕ in the duality between trace class operators and bounded linear operators on $\mathcal{H}$ (see Theorem 2.2.3) is then a linear map

$$\phi^* : \mathcal{B}(\mathcal{H}) \longrightarrow \mathcal{B}(\mathcal{H})$$

which is positive too (see Lemma 4.2.4). From a physical point of view this adjoint is important since it gives the "**effect**" $F = F_\phi$ corresponding to an operation as

$$F = \phi^*(I).$$

Lemma 4.2.2 *Let* $\phi : \mathcal{B}_1(\mathcal{H}) \longrightarrow \mathcal{B}_1(\mathcal{H})$ *be a positive linear mapping such that* $\mathrm{Tr}(\phi(W)) \leq 1$ *for all density matrices W on* $\mathcal{H}$*. Then its dual map* ϕ^* *(in the duality established in Theorem 2.2.3) is a linear map* $\mathcal{B}(\mathcal{H}) \longrightarrow \mathcal{B}(\mathcal{H})$ *which is well defined by*

$$\mathrm{Tr}(\phi^*(B)T) = \mathrm{Tr}(B\phi(T)) \quad \textit{for all } B \in \mathcal{B}(\mathcal{H}), T \in \mathcal{B}_1(\mathcal{H}). \tag{4.4}$$

Proof. Given such a map ϕ it is continuous according to Lemma 4.2.1: $\|\phi(T)\|_1 \leq C\,\|T\|_1$ for all $T \in \mathcal{B}_1(\mathcal{H})$. Fix $B \in \mathcal{B}(\mathcal{H})$; since

$$|\mathrm{Tr}(B\phi(T))| \leq \|B\|\,\|\phi(T)\|_1 \leq \|B\|\,C\,\|T\|_1\,,$$

$T \longrightarrow \mathrm{Tr}(B\phi(T))$ is a continuous linear functional on $\mathcal{B}_1(\mathcal{H})$ and therefore according to Theorem 2.2.3 of the form $\mathrm{Tr}(CT)$ with a unique $C \in \mathcal{B}(\mathcal{H})$. This element C is called $\phi^*(B)$. This applies to every $B \in \mathcal{B}(\mathcal{H})$ and thus defines the adjoint mapping ϕ^*, and by construction Relation (4.4) holds. A straight forward calculation establishes linearity of ϕ^*, using uniqueness in the duality between $\mathcal{B}_1(\mathcal{H})$ and $\mathcal{B}(\mathcal{H})$. □

Corollary 4.2.3 *For a positive linear mapping $\phi : \mathcal{B}_1(\mathcal{H}) \longrightarrow \mathcal{B}_1(\mathcal{H})$ the following statements are equivalent:*

a) $\mathrm{Tr}(\phi(W)) \leq 1$ *for all density matrices W on $\mathcal{H}$;*

b) ϕ is continuous and $\phi^*(I) \leq I$.

Proof. Suppose a) holds. Then, by Lemma 4.2.1 the map ϕ is continuous. Thus according to Lemma 4.2.2 the dual mapping $\phi^* : \mathcal{B}(\mathcal{H}) \longrightarrow \mathcal{B}(\mathcal{H})$ is well defined and Relation (4.4) holds, in particular for all density matrice W and $B = I$,

$$\mathrm{Tr}(\phi^*(I)W) = \mathrm{Tr}(\phi(W)).$$

It follows $\mathrm{Tr}(\phi^*(I)W) \leq 1$ for all W. For $W = [x,x]$, $x \in \mathcal{H}$, $\|x\| = 1$ this says $\mathrm{Tr}(\phi^*(I)[x,x]) = \langle x, \phi^*(I)x\rangle \leq 1$ and hence $\langle x, \phi^*(I)x\rangle \leq \langle x,x\rangle$ for all $x \in \mathcal{H}$ and $\phi^*(I) \leq I$ follows.

Conversely assume b). Since ϕ is continuous the dual map ϕ^* is well defined and (4.4) holds and thus again $\mathrm{Tr}(\phi^*(I)W) = \mathrm{Tr}(\phi(W))$ for all density matrices W. Now $\phi^*(I) \leq I$ implies a)

$$\mathrm{Tr}(\phi(W)) = \mathrm{Tr}(W^{1/2}\phi^*(I)W^{1/2}) \leq \mathrm{Tr}(W^{1/2}W^{1/2}) = \mathrm{Tr}(W) = 1.$$

□

Lemma 4.2.4 *A linear mapping $\phi : \mathcal{B}_1(\mathcal{H}) \longrightarrow \mathcal{B}_1(\mathcal{H})$ is completely positive if, and only if, its adjoint mapping $\phi^* : \mathcal{B}(\mathcal{H}) \longrightarrow \mathcal{B}(\mathcal{H})$ is completely positive.*

Proof. Naturally the proof consists in showing that for all $k \in \mathbb{N}$, the mapping ϕ is k-positive if, and only if, the adjoint mapping ϕ^* is k-positive. We do this explicitly for the case $k = 1$ and indicate the necessary changes for $k \geq 2$.

If $B \in \mathcal{B}(\mathcal{H})$ is given define a linear functional F_B on $\mathcal{B}_1(\mathcal{H})$ by $F_B(T) = \mathrm{Tr}(B\phi(T))$. According to Theorem 2.2.3 the duality is given by the trace formula

$$\mathrm{Tr}(B\phi(T)) = \mathrm{Tr}(\phi^*(B)T) \quad \text{for all } B \in \mathcal{B}(\mathcal{H}), T \in \mathcal{B}_1(\mathcal{H}). \tag{4.5}$$

If ϕ is positive, then we know $\phi(T) \geq 0$ for all $T \in \mathcal{B}_1(\mathcal{H})$, $T \geq 0$, and we have to show that $\phi^*(B) \geq 0$ for all $B \in \mathcal{B}(\mathcal{H})$, $B \geq 0$. According to Theorem 2.1.9 $\phi(T) \in \mathcal{B}_1(\mathcal{H})$ is positive if, and only if, it is of the form $\phi(T) = \tau^*\tau$ for some $\tau = \tau^* \in \mathcal{B}_2(\mathcal{H})$. In this case we have

$$\mathrm{Tr}(B\phi(T)) = \mathrm{Tr}(B\tau^*\tau) = \mathrm{Tr}(\tau B\tau^*) \geq 0 \quad \text{for all } B \geq 0.$$

The duality relation implies

$$\mathrm{Tr}(\phi^*(B)T) \geq 0 \quad \text{for all } T \in \mathcal{B}_1(\mathcal{H}), T \geq 0.$$

Now choose $x \in \mathcal{H}$ and insert the positive finite rank operator $T = [x,x]$ defined by $[x,x]y = x\langle x,y\rangle$, $y \in \mathcal{H}$, into this estimate to get

$$0 \leq \mathrm{Tr}(\phi^*(B)T) = \langle x, \phi^*(B)x\rangle$$

and thus $\phi^*(B) \geq 0$ for $B \geq 0$.

Conversely assume that ϕ^* is a positive mapping so that $\phi^*(B) \geq 0$ for all $B \geq 0$. Then, by Lemma 3.4.1 (or the square root lemma) for some $b \in \mathcal{B}(\mathcal{H})$ we know $\phi^*(B) = b^*b$ and the duality relation yields

$$\mathrm{Tr}(B\phi(T)) = \mathrm{Tr}(\phi^*(B)T) = \mathrm{Tr}(b^*bT) = \mathrm{Tr}(bTb^*) \geq 0$$

for all $T \geq 0$. As above insert $B = [x,x]$ to get $\langle x, \phi(T)x\rangle \geq 0$ whenever $T \geq 0$ and hence the mapping ϕ is positive.

Now assume $k \geq 2$; abbreviate $\mathcal{A} = \mathcal{B}_1(\mathcal{H})$ and $\mathcal{B} = \mathcal{B}(\mathcal{H})$. We have to show that $\phi_k : M_k(\mathcal{A}) \longrightarrow M_k(\mathcal{A})$ is positive if, and only if, $\phi_k^* : M_k(\mathcal{B}) \longrightarrow M_k(\mathcal{B})$ is positive. Recall that $A = [a_{ij}] \in M_k(\mathcal{A})$ respectively $B = [b_{ij}] \in M_k(\mathcal{B})$ act on the Hilbert space $\mathcal{H}^k = \mathcal{H} \times \mathcal{H} \times \cdots \times \mathcal{H}$ (k components). Under standard matrix operations we have $M_k(\mathcal{B}) = \mathcal{B}(\mathcal{H}^k)$ and similarly $M_k(\mathcal{A}) = \mathcal{B}_1(\mathcal{H}^k)$ (see the Exercises). For the relation of traces in $\mathcal{H}$ and in $\mathcal{H}^k$ one finds (see again the Exercises for this chapter)

$$\mathrm{Tr}_{\mathcal{H}^k}([T_{ij}]) = \sum_{j=1}^{k} \mathrm{Tr}(T_{jj})$$

when Tr denotes the trace in $\mathcal{H}$. Thus we get the extended duality formula

$$\mathrm{Tr}_{\mathcal{H}^k}([b_{ij}]\phi_k([T_{ij}])) = \mathrm{Tr}_{\mathcal{H}^k}(\phi_k^*([b_{ij}])[T_{ij}])) \tag{4.6}$$

since

$$\mathrm{Tr}_{\mathcal{H}^k}([b_{ij}]\phi_k([T_{ij}])) = \sum_{i,j=1}^{k} \mathrm{Tr}(b_{ij}\phi(T_{ji})) = \sum_{i,j=1}^{k} \mathrm{Tr}(\phi^*(b_{ij})T_{ji}).$$

Therefore we can argue for $\phi_k : \mathcal{B}_1(\mathcal{H}^k) \longrightarrow \mathcal{B}_1(\mathcal{H}^k)$ and $\phi_k^* : \mathcal{B}(\mathcal{H}^k) \longrightarrow \mathcal{B}(\mathcal{H}^k)$ as above for ϕ and ϕ^*. □

4.2.2 The representation theorem for operations

Naturally the question about the general mathematical form of a quantum operation arises. The answer has been given in 16. In Section 3.4.5 we had studied completely positive maps on $\mathcal{B}(\mathcal{H})$. Here we begin by investigating completely positive maps on trace class operators and find some extensions of the earlier results.

Proposition 4.2.5 *For a sequence of operators $a_j \in \mathcal{B}(\mathcal{H})$ which satisfies (3.56) with bound B the series*

$$\phi(T) = \sum_{j=1}^{\infty} a_j T a_j^*, \quad T \in \mathcal{B}_1(\mathcal{H}) \text{ fixed} \tag{4.7}$$

converges in trace norm and defines a completely positive mapping on $\mathcal{B}_1(\mathcal{H})$. The *related series (3.57), i.e.,*

$$f(a) = \sum_{j=1}^{\infty} a_j^* a a_j, \quad a \in \mathcal{B}(\mathcal{H}) \text{ fixed}$$

converges ultraweakly and defines the adjoint of ϕ, i.e.,

$$\phi^*(a) = \sum_{j=1}^{\infty} a_j^* a a_j, \quad a \in \mathcal{B}(\mathcal{H}) \text{ fixed.} \tag{4.8}$$

Furthermore

$$\phi^*(I) \leq B.$$

Proof. Given $0 \leq T \in \mathcal{B}_1(\mathcal{H})$ define for $m \in \mathbb{N}$,

$$\phi_m(T) = \sum_{j=1}^{m} a_j T a_j^*.$$

Clearly $\phi_m(T)$ is nonnegative and of trace class; thus for $m > n$ we find

$$\begin{aligned}\|\phi_m(T) - \phi_n(T)\|_1 &= \left\| \sum_{j=n+1}^{m} a_j T a_j^* \right\|_1 = \operatorname{Tr}\left(\sum_{j=n+1}^{m} a_j T a_j^* \right) \\ &= \operatorname{Tr}\left(\sum_{j=n+1}^{m} a_j^* a_j T \right) = \operatorname{Tr}(S_m T) - \operatorname{Tr}(S_n T)\end{aligned}$$

where the operators $S_m = \sum_{j=1}^{m} a_j^* a_j$ where introduced in (3.56). Because of the ultraweak converges $S_m \longrightarrow S$ according to Lemma 3.4.11 we know $\operatorname{Tr}(S_m T) \longrightarrow \operatorname{Tr}(ST)$ and we conclude that $(\phi_m(T))$ is a Cauchy sequences with respect to the trace norm and therefore this sequence converges in trace norm to a unique $\phi(T) \in \mathcal{B}_1(\mathcal{H})$.

Since the trace norm dominates the operator norm we also have convergence of (4.7) in operator norm and thus also ultraweakly. Since every trace class operator is the complex linear combination of four positive ones the series (4.7) converges for every $T \in \mathcal{B}_1(\mathcal{H})$ with respect to the topologies as indicated above. The complete positivity of ϕ follows as in the proof of Lemma 3.4.11.

These continuity properties allow to determine the dual ϕ^* of ϕ easily. This dual is determined by

$$\mathrm{Tr}(B\phi(T)) = \mathrm{Tr}(\phi^*(B)T) \quad \text{for all } B \in \mathcal{B}(\mathcal{H}), T \in \mathcal{B}_1(\mathcal{H}).$$

We have

$$\mathrm{Tr}(B\phi(T)) = \lim_{m\to\infty} \mathrm{Tr}(B\phi_m(T))$$

and by property c) of Corollary 2.1.8

$$\mathrm{Tr}(B\sum_{j=1}^{m} a_j T a_j^*) = \mathrm{Tr}(\sum_{j=1}^{m} a_j^* B a_j T).$$

According to Lemma 3.4.11 we know $\lim_{m\to\infty} \sum_{j=1}^{m} a_j^* B a_j = f(b)$ in the ultraweak topology, hence $\mathrm{Tr}(\phi^*(B)T) = \mathrm{Tr}(f(B)T)$ for all $T \in \mathcal{B}_1(\mathcal{H})$. Thus we conclude. □

The following representation theorem is the version Kraus has given in his 1983 Springer Lecture Notes. Its proof is actually much more complicated than I originally thought, even after substantial preparations (Stinespring's factorization theorem, Naimark's characterization of representations of $\mathcal{B}(\mathcal{H})$, characterizations of completely positive maps), and I also think that in some points it is not quite accurate. The elimination of the representation of the Calkin algebra in the representation formula for operations looks quite strange to me (I offer my own proof) and also the bound $\mathrm{Tr}(\phi(W)) \leq I$ for all density matrices W is not taken into account properly. In order to do so I have added the important Corollary 4.2.3 .

Therefore I would be interested to see an updated proof of Kraus' result. Maybe you know some source?

Also, since the presentation and the proof of this result takes much more space than I originally envisaged we might think about to include only a shortened and simplified version. Maybe you have some suggestions.

Theorem 4.2.6 (First representation theorem of Kraus) *Given an operation* $\phi : \mathcal{B}_1(\mathcal{H}) \longrightarrow \mathcal{B}_1(\mathcal{H})$*, there exists a finite or countable family* $\{a_j : j \in J\}$ *of bounded linear operators on* $\mathcal{H}$*, satisfying*

$$\sum_{j \in J_0} a_j^* a_j \leq I \quad \textit{for all finite } J_0 \subset J, \tag{4.9}$$

such that for every $T \in \mathcal{B}_1(\mathcal{H})$ *and every* $B \in \mathcal{B}(\mathcal{H})$ *one has*

$$\phi(T) = \sum_{j \in J} a_j T a_j^* \tag{4.10}$$

respectively

$$\phi^*(B) = \sum_{j \in J} a_j^* B a_j. \tag{4.11}$$

The effect F corresponding to ϕ thus has the representation

$$F = \phi^*(I) = \sum_{j \in J} a_j^* a_j. \tag{4.12}$$

In the case that the index set J is infinite, i.e., $J = \mathbb{N}$, the series in (4.10) converges with respect to the trace norm while the series (4.11) - (4.12) converge in the ultraweak operator topology.

Conversely, if a countable family $\{a_j : j \in J\}$ of bounded linear operators on $\mathcal{H}$ is given which satisfies (4.9) then Equation (4.10) defines an operation ϕ whose adjoint ϕ^* is given by (4.11) and the effect F corresponding to this operation is (4.12).

Proof. Suppose we are given a completely positive map $\phi : \mathcal{B}_1(\mathcal{H}) \longrightarrow \mathcal{B}_1(\mathcal{H})$ satisfying (4.2). Lemma 4.2.4 implies that the adjoint map $\phi^* : \mathcal{B}(\mathcal{H}) \longrightarrow \mathcal{B}(\mathcal{H})$ is completely positive too, thus according to Theorem 3.4.12 ϕ^* is of the form (3.59)

$$\phi^*(B) = V_0^* \pi_0(B) V_0 + \sum_{j \in J} a_j^* B a_j, \quad B \in \mathcal{B}(\mathcal{H})$$

with bounded linear operators $a_j \in \mathcal{B}(\mathcal{H})$ satisfying

$$\sum_{j \in J} a_j^* a_j \leq \phi^*(I)$$

and where the representation π_0 vanishes for all $B \in \mathcal{B}_c(\mathcal{H})$. According to Corollary 4.2.3 the bound $\phi^*(I) \leq I$ is known and hence Condition (4.9) holds.

Proposition 4.2.5 implies that the map $\phi_1^*(B) = \sum_{j \in J} a_j^* B a_j$ on $\mathcal{B}(\mathcal{H})$ is the adjoint of the mapping $\phi_1(T) = \sum_{j \in J} a_j T a_j^*$ on $\mathcal{B}_1(\mathcal{H})$. In order to conclude we need to determine the map ϕ_0 on $\mathcal{B}_1(\mathcal{H})$ whose adjoint is the map $\phi_0^*(B) = V_0^* \pi_0(B) V_0$ on $\mathcal{B}(\mathcal{H})$. This map is defined through the duality relation

$$\mathrm{Tr}(\phi_0^*(B) T) = \mathrm{Tr}(B \phi_0(T)) \quad \text{for all} \quad B \in \mathcal{B}(\mathcal{H}), T \in \mathcal{B}_1(\mathcal{H}).$$

Since the representation π_0 of $\mathcal{B}(\mathcal{H})$ vanishes on the subspace $\mathcal{B}_c(\mathcal{H})$ we know $\mathrm{Tr}(B\phi_0(T)) = 0$ for all $B \in \mathcal{B}_c(\mathcal{H})$, hence in particular for all $x, y \in \mathcal{H}$ setting $B = [xy]$,

$$\langle y, \phi_0(T)x \rangle = \mathrm{Tr}([xy]\phi_0(T)) = 0,$$

and therefore $\phi_0(T) = 0$ for all $T \in \mathcal{B}_1(\mathcal{H})$ and thus $\phi_0^* = 0$ on $\mathcal{B}(\mathcal{H})$. It follows

$$\phi^*(I) = \sum_{j \in J} a_j^* a_j.$$

The converse has already been proven in Proposition 4.2.5 and Lemma 3.4.11 with the bound $B = I$ when we observe Corollary 4.2.3. □

Remark 4.2.7 *Sometimes one requires that an operation ϕ is* **trace preserving**, *i.e.,* $\mathrm{Tr}(\phi(W)) = 1$ *for all density matrices W. This will be the case when in our representation (4.10) the operators a_j satisfy*

$$\sum_{j \in J} a_j^* a_j = I. \tag{4.13}$$

In order to prove this recall that according to (4.12) one has

$$\sum_{j \in J} a_j^* a_j = \phi^*(I)$$

and that we know $\phi^*(I) \leq I$. The duality relation says

$$\mathrm{Tr}(\phi(W)) = \mathrm{Tr}(\phi^*(I)W)$$

for all density matrices W. Thus, if $\phi^*(I) = I$ *then* $\mathrm{Tr}(\phi(W)) = 1$ *for all density matrices and* ϕ *is trace preserving. Conversely suppose that the operation* ϕ *is trace preserving but* $\phi^*(I) \neq I$*. Then, since* $\phi^*(I) \leq I$ *is known, there is* $x \in \mathcal{H}$, $\|x\| = 1$ *such that* $\langle x, \phi^*(I)x \rangle < 1$*. If the density matrix* $W = [x,x]$ *is inserted into the duality relation one gets*

$$\mathrm{Tr}(\phi([x,x])) = \mathrm{Tr}(\phi^*(I)[x,x]) = \langle x, \phi^*(I)x \rangle < 1,$$

hence a contradiction and therefore $\phi^*(I) = I$ *holds.*

4.3 Choi's results for finite dimensional completely positive maps

Naturally in the case of mappings $f : \mathcal{A} \longrightarrow \mathcal{B}$ with $\mathcal{A} = M_n(\mathbb{C})$ and $\mathcal{B} = M_m(\mathbb{C})$ we can use additional structural information to strengthen the statements of Stinespring's factorization theorem (Theorem 3.4.8) and to simplify the proofs. This has been done in 1975 by M. Choi 8 with inspiration from electrical circuit theory (n-port systems) by using the simple fact that these matrix algebras $M_n(\mathbb{C})$ have a basis

$$e^{(n;ij)}, \; i,j = 1,2,\ldots,n \tag{4.14}$$

where $e^{(n;ij)}$ denotes the $n \times n$ matrix with the entry 1 in the ith row and jth column and all other entries are 0.

In terms of this basis we can write $a \in M_n(\mathbb{C})$ as follows:

$$a = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \sum_{i,j=1}^{n} a_{ij} e^{(n;ij)}, \quad a_{ij} \in \mathbb{C}. \tag{4.15}$$

And this allows to determine the general form of a linear map $f : M_n(\mathbb{C}) \longrightarrow M_m(\mathbb{C})$ easily. For $a \in M_n(\mathbb{C})$ as above one finds by linearity

$$f(a) = \sum_{i,j=1}^{n} a_{ij} f(e^{(n;ij)}).$$

Since $f(e^{(n;ij)}) \in M_m(\mathbb{C})$ it has a unique expansion with respect to the basis

$$e^{(m;kl)},\ k,l = 1,2,\ldots,m,$$

i.e.

$$f(e^{(n;ij)}) = \sum_{k,l=1}^{m} f(e^{(n;ij)})_{kl} e^{(m;kl)}, \quad f(e^{(n;ij)})_{kl} \in \mathbb{C}.$$

Thus we can say that there is a one-to-one correspondence between linear maps $f : M_n(\mathbb{C}) \longrightarrow M_m(\mathbb{C})$ and system of complex numbers F^{ij}_{kl}, $i, j = 1, \ldots, n$, $k, l = 1, \ldots, m$ such that

$$f(a) = \sum_{k,l=1}^{m} \sum_{i,j=1}^{n} a_{ij} F^{ij}_{kl} e^{(m;kl)} \tag{4.16}$$

with a as in (4.15).

Theorem 4.3.1 (Choi's characterization of completely positive maps) *For a linear map $f : M_n(\mathbb{C}) \longrightarrow M_m(\mathbb{C})$ the following statements are equivalent:*

(a) f is n-positive, i.e., the map $f_n : M_n(M_n(\mathbb{C})) \longrightarrow M_n(M_m(\mathbb{C}))$ defined in (3.40) is positive;

(b) the matrix $C_f \in M_n(M_m(\mathbb{C}))$ defined by

$$C_f = \begin{pmatrix} f(e^{(n;11)}) & \cdots & f(e^{(n;1n)}) \\ \vdots & \cdots & \vdots \\ f(e^{(n;n1)}) & \cdots & f(e^{(n;nn)}) \end{pmatrix} \tag{4.17}$$

is positive where $e^{(n;ij)}$ is specified in (4.14); it is called the **Choi -matrix** *of f.*

(c) f has the form

$$f(a) = \sum_{\mu=1}^{nm} V_\mu a V_\mu^* \,, a \in M_n(\mathbb{C}) \tag{4.18}$$

with $m \times n$ matrices V_μ, and thus f is completely positive.

Proof. If a linear map f is of the form (4.18) it is a straightforward calculation to show that f is completely positive, just as in the case of the Stinespring factorization. Thus it is clear that (c) implies (a).

(a) $\Rightarrow$ (b): Note that the matrix $E \in M_n(M_n(\mathbb{C}))$ given by

$$E = \begin{pmatrix} e^{(n;11)} & \dots & e^{(n;1n)} \\ \vdots & \dots & \vdots \\ e^{(n;n1)} & \dots & e^{(n;nn)} \end{pmatrix} \tag{4.19}$$

satisfies $E^* = E$ (since $(e^{(n;ij)})^* = e^{(n;ji)}$ and $E^2 = E$, thus $E = E^*E$ is positive in $M_n(M_n(\mathbb{C}))$ by Lemma 3.4.1. Since f is assumed to be n-positive $f_n(E) = C_f$ is positive in $M_n(M_m(\mathbb{C}))$.

(b) $\Rightarrow$ (c): By definition, the matrix C_f acts on $\mathcal{H}_m^n \cong \mathbb{C}^{nm}$. If (b) is assumed this matrix is positive and thus its spectrum is contained in $[0, \left\|C_f\right\|]$. Its spectral representation is of the form

$$C_f = \sum_{\nu=1}^{nm} \lambda_k Q_k, \quad 0 \le \lambda_k \le \left\|C_f\right\|$$

where Q_k is the projector onto the eigen-space corresponding to the eigenvalue λ_k.

Denote by P_i the projection from $\mathcal{H}_m^n = \mathcal{H}_m \times \mathcal{H}_m \times \cdots \times \mathcal{H}_m$ (n times) to the ith component $\mathcal{H}_m$, i.e., $P_i(z_1,\dots,z_i,\dots,z_n) = z_i$, for all $z_j \in \mathcal{H}_m$, $j = 1,\dots,n$. Then (4.17) shows

$$f(e^{(n;ij)}) = P_i C_f P_j, \quad i,j = 1,\dots,n\,,$$

and the spectral representation thus implies

$$f(e^{(n;ij)}) = \sum_{k=1}^{nm} \lambda_k P_i Q_k P_j\,.$$

The normalized eigenvector $v^{(k)}$ for the eigenvalue λ_k belongs to the space $\mathcal{H}_m^n$ and thus has a decomposition $v^{(k)} = (v_1^{(k)}, \ldots, v_n^{(k)})$ with $v_i^{(k)} \in \mathcal{H}_m$ for $i = 1, \ldots, n$. With the standard convention for the tensor product the projector Q_k can be realized as $Q_k = v^{(k)} \otimes (v^{(k)})^*$. This allows to rewrite the above formula for $f(e^{(n;ij)})$ as

$$f(e^{(n;ij)}) = \sum_{k=1}^{nm} \lambda_k P_i v^{(k)} \otimes (v^{(k)})^* P_j = \sum_{k=1}^{nm} \lambda_k v_i^{(k)} \otimes (v_j^{(k)})^* .$$

Denote by $e^{(n;i)}$, $i = 1, \ldots, n$ the standard basis of $\mathcal{H}_n$. For $k = 1, \ldots, mn$ define linear operators $V^{(k)} : \mathcal{H}_n \longrightarrow \mathcal{H}_m$ by their action on this basis

$$V^{(k)} e^{(n;i)} = \sqrt{\lambda_k} v_i^{(k)}, \quad i = 1, \ldots, n .$$

Hence we can continue our chain of identities for $f(e^{(n;ij)})$ by

$$f(e^{(n;ij)}) = \sum_{k=1}^{nm} (V^{(k)} e^{(n;i)}) \otimes (V^{(k)} e^{(n;j)})^* = \sum_{k=1}^{nm} V^{(k)} (e^{(n;i)}) \otimes (e^{(n;j)})^*)(V^{(k)})^*$$

or, since $e^{(n;i)} \otimes (e^{(n;j)})^* = e^{(n;ij)}$,

$$f(e^{(n;ij)}) = \sum_{k=1}^{nm} V^{(k)} e^{(n;ij)} (V^{(k)})^* \tag{4.20}$$

and thus by (4.15) f has the form (4.18). This proves (c). □

Remark 4.3.2 *(a) This result of M. D. Choi is quite remarkable. It shows in particular that a linear map on the matrix algebra $M_n(\mathbb{C})$ with values in $M_m(\mathbb{C})$ is already completely positive when it is n-positive.*

(b) In addition it is shown that such a linear map is n-positive whenever it is n-positive on the elements of the (standard) basis.

(c) *It determines the explicit form of completely positive maps which is considerable more specific than the Stinespring factorization.*

(d) *In the case of matrix algebras the proof of the Stinespring factorization indicates that a linear map is completely positive if is is $n^2 \times m$-positive (the dimension of the space $M_n(\mathbb{C}) \otimes \mathbb{C}^m$ is n^2m).*

(e) *The map $f \longrightarrow C_f$ defined in Equation (4.17) is often called* **Jamiolkowski isomorphism** *or* **Choi-Jamiolkowski isomorphism**. *It appeared first in 13.*

Corollary 4.3.3 (Finite-dimensional representations of $M_n(\mathbb{C})$) *Let $\pi : M_n(\mathbb{C}) \longrightarrow M_m(\mathbb{C})$ be a finite-dimensional representation of the matrix algebra $M_n(\mathbb{C})$. Then there are $m \times n$ matrices V^μ, $\mu = 1, \ldots, mn$ satisfying*

$$V^{\mu *} V^\nu = \delta_{\mu,\nu} I_n$$

such that

$$\pi(a) = \sum_{k=1}^{nm} V^{(k)} a (V^{(k)})^* \quad \forall\, a \in M_n(\mathbb{C}).$$

4.4 Exercises

1. For $k = 2, 3, \ldots$ and a separable Hilbert space $\mathcal{H}$ denote $\mathcal{H}^k = \mathcal{H} \times \mathcal{H} \cdots \mathcal{H}$ (k components). With the standard operations and the natural scalar product $\mathcal{H}^k$ is a Hilbert space in which the given Hilbert space is embedded by isometric mappings $J_1 : \mathcal{H} \longrightarrow \mathcal{H} \times \{0\} \times \cdots \times \{0\}$, $J_2 : \mathcal{H} \longrightarrow \{0\} \times \mathcal{H} \times \{0\} \times \cdots \times \{0\}$, …, $J_k : \mathcal{H} \longrightarrow \{0\} \times \cdots \times \{0\} \times \mathcal{H}$. Show: If $B_\kappa = \left\{ e_j^\kappa : j \in \mathbb{N} \right\}$ is an orthonormal basis of $\mathcal{H}$, then $J_1(B_1) \times J_2(B_2) \times \cdots \times J_k(B_k)$ is an orthonormal basis of $\mathcal{H}^k$.

2. Using the notation introduced in the text show $M_k(\mathcal{B}(\mathcal{H})) = \mathcal{B}(\mathcal{H}^k)$ and $M_k(\mathcal{B}_1(\mathcal{H})) = \mathcal{B}_1(\mathcal{H}^k)$.
   Hints: In order to show $M_k(\mathcal{B}_1(\mathcal{H})) \subseteq \mathcal{B}_1(\mathcal{H}^k)$ use a suitable characterization of trace class operators as given in Proposition 2.1.5 and observe Exercise 1.

3. Observe Exercise 1 to prove the 'trace formula'

$$\mathrm{Tr}_{\mathcal{H}^k}([T_{ij}]) = \sum_{j=1}^{k} \mathrm{Tr}(T_{jj})$$

for $[T_{ij}] \in M_k(\mathcal{B}_1(\mathcal{H}))$.

Bibliography

[1] J.F. Adams, R. A.and Fournier. *Sobolev Spaces*. Elsevier, 2003.

[2] R. A. Adams. *Sobolev Spaces*. Academic Press, , Boston, MA, 1975. ISBN 978-0-12-044150-1.

[3] P. Blanchard and E. Brüning. Reply to “comment on ‘remarks on the structure of states of composite quantum systems and envariance’[phys. lett. a 355 (2006)180]"[phys. lett. a 375 (2011) 1160]. *Phys. Lett. A*, 375:1163–1165, 2011.

[4] Ph. Blanchard and E. Brüning. *Mathematical Methods in Physics – Distributions, Hilbert space operators and variational methods*, volume 26 of *Progress in Mathematical Physics*. Birkhäuser, Boston Basel Berlin, 2003.

[5] Ph. Blanchard and E. Brüning. *Mathematical Methods in Physics – Distributions, Hilbert space operators and variational methods*, volume 69 of *Progress in Mathematical Physics*. Birkhäuser, Boston Basel Berlin, 2015.

[6] P. Busch. Quantum states and generalized observables: a simple proof of gleason's theorem. *Phys. Rev. Lett.*, 91(12/120403):1–4, 2003.

[7] P. R. Chernoff. Andy gleason and quantum mechanics. *Notices AMS*, 56(10):1253–1259, 2009.

[8] M. D. Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10:285 – 290, 1975.

[9] E. B. Davies. *Linear Operator and their Sprectra*, volume 106 of *Cambridge studies in advanced mathematics*. Cambridge University Press, 2007.

[10] B. K. Driver. *Analysis Tools with Applications*. Springer, Berlin Heidelberg New York London Hong Kong Milan Paris Tokyo, 2003.

[11] A. M. Gleason. Measures on the closed subspaces of a hilbert space. *J. Math. Mech.*, 6:885–893, 1957.

[12] C. J. Isham. *Lectures on quantum theory : mathematical and structural foundations*. Imperial College Press, London, 1995.

[13] A Jamiolkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Rep. Math. Phys.*, 3:275, 1972.

[14] J. M. Jauch. *Foundations of quantum mechanics*. Addison-Wesley, Reading, Massachusetts, 1973.

[15] A. N. Kolmogorov. über Kompaktheit der Funktionenmengen bei der Konvergenz im Mittel. *Nachr. Ges. Wiss. Göttingen*, 9:60–63, 1931.

[16] K. Kraus. *States, Effects, and Operations*, volume 190 of *Lecture Notes in Physics*. Springer-Verlag, Berlin Heidelberg New York Tokyo, 1983.

[17] N. Meyers and J. Serrin. H=w. *Proc. Nat. Acad. Sci.*, 51:1055–1056, 1964.

[18] N. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000.

[19] J.R. Ringrose R. V. Kadison. *Fundamentals of the Theory of Operator Algebras Volume I: Elementary Theory*. Academic Press, 1983.

[20] F. Richman and D. Bridges. Aconstruspace proof of gleason' theorem. *J. Funct. Anal.*, 162:287–312, 1999.

[21] M. Riesz. Sur les ensembles compacts de fonctions sommables. *Acta Szeged Sect. Math.*, 6:136–142, 1933.

[22] W. Rudin. *Functional analysis*. McGraw Hill, New York, 1973.

[23] W. F. Stinespring. Positive functions on C^*-algebras. *Proceedings of the American Mathematical Society*, 6 No. 2:211 – 216, 1955.

[24] M. Takesaki. *Theory of Operator algebras I*, volume 124 of *Encyclopedia of Mathematical Sciences - Operator Algebras and Non-Commutative Geometry*. Springer, Berlin Heidelber New York, 2002.