Fumio Hiai

Quantum f-Divergences
in von Neumann Algebras

Reversibility of Quantum Operations

March 10, 2020

Preface

After I wrote in 1981 a joint paper [55] with M. Tsukada and M. Ohya on the sufficiency of von Neumann subalgebras and the relative entropy in a specialized situation, I liked to extend the sufficiency notion to general von Neumann subalgebras. But I felt some difficulty to do that because of the lack of the conditional expectation onto a general von Neumann subalgebra. In the meantime, I was shocked by D. Petz' remarkable paper [92] in 1986, where he settled the problem mentioned above by use of the generalized conditional expectation due to L. Accardi and C. Cecchini [1] in 1982. In his 1988 paper [93] Petz followed the same idea to introduce the notion of sufficient (or reversible) channels (more precisely, unital normal 2-positive maps) between von Neumann algebras, and he further collaborated with A. Jenčová on the subject in [62, 63].

After joint work of the 2011 paper [53], I thought that it would be nice to extend the materials for [53] to the von Neumann algebra setting, including a renewal of Petz' theory on the sufficiency and reversibility in von Neumann algebras. But that was not undertaken due to my laziness before I looked at the papers [17, 59, 60], where the sandwiched Rényi divergence, a new Rényi divergence extensively developed in these years, was extended to the von Neumann algebra setting. Those papers strongly motivated me to develop a general theory of quantum f-divergences in von Neumann algebra with applications to the sufficiency and reversibility problem. I planned to first write three papers on three different kinds of f-divergences (the standard f-divergences, the maximal f-divergences, and the measured f-divergences) in von Neumann algebras. The first two papers were published in [48, 49], which are surveyed in Chaps. 2 and 4 of this monograph. However, I write the third one in Chap. 5 instead of publishing a separate paper, so Chap. 5 of the monograph is new. In the main part of Chaps. 6–8, I revisit the work [92, 93, 95, 62] with slight extensions and some modifications.

I dedicate this monograph to the memory of Dénes Petz (1953–2018).

case. I thank Mosonyi also for his inviting me to an intensive course [50] in 2019, April at Department of Mathematical Analysis, Budapest University of Technology and Economics. Appendix A of this monograph is a short extract from those lecture notes. I learned from Jenčová (as [61]) important points in Secs. 3.2 and 3.3, in particular, Lemma 3.5 is indebted to her. Discussions with Kosaki in recent joint work [51] on operator connections of unbounded positive operators and positive forms are helpful to write Appendix D.

I am grateful to Masayuki Nakamura (Springer Japan) for his constant supports, without which it is not possible for me to write this monograph.

Chiba, Japan,
March 2020   *Fumio Hiai*

Contents

Chapter 1
Introduction

The notion of quantum divergences has played a significant role in quantum information, which defines important quantum quantities to discriminate between states of a quantum system. A quantum system is mathematically described, in most cases, by an operator algebra $\mathcal{A}$ on a Hilbert space, either finite-dimensional or infinite-dimensional, and a quantum divergence is generally given as a function $S(\psi\|\varphi)$ of two states or more generally, two positive linear functionals ψ and φ on $\mathcal{A}$. This monograph is aimed at presenting a comprehensive survey on quantum f-divergences and showing their significant role in the reversibility problem of quantum operations in the general von Neumann algebra setting.

A general and rigorous framework of quantum information should be that of von Neumann algebras, while, in the present circumstances, quantum information literature is mostly presented in the finite-dimensional setting or the matrix setting. Indeed, the reader can consult the textbooks [44, 96, 87, 115] for broad and intensive developments of quantum information in the last quarter century. As widely believed, the von Neumann algebra framework is the most suitable in studying non-commutative (= quantum) probability and integration from the pure mathematical side [102, 29, 86] and also in developing quantum statistical mechanics from mathematical physics side [20, 21]. This idea provides us a good motivation to study quantum information in the von Neumann algebra setting. C^*-algebras, a more abstract class of operator algebras, are useful as well in studying quantum physics, but discussions in C^*-algebras can be, in many cases, reduced to those in von Neumann algebras by taking the GNS (Gelfand-Naimark-Segal) construction associated with a relevant state on a C^*-algebra. Furthermore, several well-developed machines are at our disposal in theory of von Neumann algebras such as, starting from the structure theory based on the Tomita-Takesaki theory, the standard forms, the relative modular operators, the non-commutative L^p-spaces, and so on. Those basics of von Neumann algebras are briefly surveyed, for the convenience of the reader, in Appendix A of the monograph as preliminaries for quantum information in the von Neumann algebra setting.

In the first half of the monograph we present a comprehensive survey of quantum f-divergences, which are introduced for normal positive linear functionals ψ, φ on a

von Neumann algebra M with a parametrization of convex functions (often operator convex functions) f on $(0, \infty)$. In Chaps. 2–5 we deal with three different kinds of quantum f-divergences. First, the standard f-divergences $S_f(\psi\|\varphi)$ discussed in Chap. 2 are defined by use of the relative modular operators due to Araki [9], which are a special case of quasi-entropies introduced by Petz [90, 91]. The most notable example (when $f(t) = t\log t$) is the relative entropy $D(\psi\|\varphi)$ having a long history as the quantum version of the Kullback-Leibler divergence in classical theory, first introduced by Umegaki [114] in semifinite von Neumann algebras and extended by Araki [8, 9] to the general von Neumann algebra case. The content of Chap. 2 is a survey from [48], and a main point there is a variational expression (see Theorem 2.5) of $S_f(\psi\|\varphi)$ extending Kosaki's expression [70] of the relative entropy. See [48] for the proofs of the results given in Chap. 2, and more detailed discussions in the finite-dimensional case are found in [52, 53].

The maximal f-divergences $\widehat{S}_f(\psi\|\varphi)$ are discussed in Chap. 4. When specialized to the finite-dimensional case, those f-divergences are defined as the trace value of the operator perspective $D_\varphi^{1/2} f(D_\varphi^{-1/2} D_\psi D_\varphi^{1/2}) D_\varphi^{1/2}$ of the density operators D_ψ, D_φ, whose detailed discussions in the finite-dimensional case are found in [52, 79]. A special example of the maximal f-divergence (when $f(t) = t\log t$) is a variant of the relative entropy introduced by Belavkin and Staszewski [14], denoted by $D_{\mathrm{BS}}(\psi\|\varphi)$. The content of Chap. 4 is a survey from [49], where, apart from the definition of $\widehat{S}_f(\psi\|\varphi)$, the main results are the integral expression (see Theorem 4.8) and the variational expression (see Theorem 4.17) in terms of reverse tests introduced following Matsumoto's idea in [79].

Assume that f is an operator convex function on $(0, \infty)$. Among many properties of the quantum f-divergence S_f (resp,, $\widehat{S}_f$), the most important is the monotonicity property (also called the DPI, the data-processing inequality)

$$S_f(\psi \circ \gamma \| \varphi \circ \gamma) \le S_f(\psi\|\varphi), \qquad \widehat{S}_f(\psi \circ \gamma \| \varphi \circ \gamma) \le \widehat{S}_f(\psi\|\varphi) \qquad (1.1)$$

for any unital normal Schwarz (resp., simply positive) map $\gamma : N \to M$ between von Neumann algebras N, M (in particular, normal completely positive maps or quantum channels). Here, a linear map $\gamma : N \to M$ is a Schwarz map if $\gamma(x^*x) \ge \gamma(x)^*\gamma(x)$ for all $x \in N$. In Chap. 5 we introduce the measured f-divergences $S_f^{\mathrm{meas}}(\psi\|\varphi)$ by the supremum of $S_f(\mathcal{M}(\psi)\|\mathcal{M}(\varphi)) := \sum_{j=1}^n \varphi(A_j) f(\psi(A_j)/\varphi(A_j))$, the classical f-divergence of $\mathcal{M}(\psi) := (\psi(A_j))$ and $\mathcal{M}(\varphi) = (\varphi(A_j))$ over all measurements (or quantum-classical channels) $\mathcal{M} = (A_j)_{1 \le j \le n}$ in M. The projective measured f-divergence $S_f^{\mathrm{pr}}(\psi\|\varphi)$ is also defined to be the supremum taken with restriction to projective measurements in M. The measured f-divergences S_f^{meas} and S_f^{pr} are also defined by using $\widehat{S}_f$ in place of S_f because $\widehat{S}_f(\psi\|\varphi) = S_f(\psi\|\varphi)$ for $\psi, \varphi \in M_*^+$ if M is an abelian von Neumann algebra. By construction we have the relation

$$S_f^{\mathrm{pr}} \le S_f^{\mathrm{meas}} \le S_f \le \widehat{S}_f.$$

A more intrinsic explanation of the above relation is given at the end of Sec. 5.1. For the measured f-divergences S_f^{meas} and S_f^{pr}, in Chap. 5 we discuss their basic

properties, variational expressions, and the existence of optimal measurements. The part of the measured f-divergences in the von Neumann algebra setting is new while discussions in the finite-dimensional case are found in [52, 80]. So the presentation of Chap. 5 is mostly self-contained including proofs.

Together with the relative entropy, among the most used quantum divergences is the α-Rényi divergence (or the α-Rényi relative entropy) $D_\alpha(\psi\|\varphi)$ with parameter $\alpha \in (0, \infty) \setminus \{1\}$, which is essentially the standard f-divergence for $f(t) = t^\alpha$ (for $\alpha > 1$) or $-t^\alpha$ (for $0 < \alpha < 1$). In these years it has also been widely known that another variant of the α-Rényi divergence $\widetilde{D}_\alpha(\psi\|\varphi)$ with parameter $\alpha \in [1/2, \infty) \setminus \{1\}$, called the sandwiched α-Rényi divergence, is equally useful in quantum information, in particular, in quantum state discrimination, see [83, 84] for example. In the finite-dimensional case with the density operators D_ψ, D_φ, the divergences D_α and $\widetilde{D}_\alpha$ are given by

$$D_\alpha(\psi\|\varphi) := \frac{1}{\alpha - 1} \log \frac{\mathrm{Tr}\, D_\psi^\alpha D_\varphi^{1-\alpha}}{\psi(1)},$$

$$\widetilde{D}_\alpha(\psi\|\varphi) := \frac{1}{\alpha - 1} \log \frac{\mathrm{Tr}\, (D_\varphi^{\frac{1-\alpha}{2\alpha}} D_\psi D_\varphi^{\frac{1-\alpha}{2\alpha}})^\alpha}{\psi(1)}.$$

The monograph [111] is a good source for Rényi type quantum divergences and their mathematical backgrounds in the finite-dimensional setting. The sandwiched Rényi divergence $\widetilde{D}_\alpha$ has recently been extended to the von Neumann algebra setting by Berta, Scholz and Tomamichel [17] and Jenčová [59, 60], while the description of D_α in von Neumann algebras is found in [48]. Chap. 3 is a survey from those articles. In particular, it was shown in [60] that the two definitions of $\widetilde{D}_\alpha$ in [17, 59] are equivalent. Its proof different from that in [60] is given in Sec. 3.3.

In the second half in Chap. 6–8 we develop the reversibility theory for quantum operations between von Neumann algebras via quantum f-divergences. Let $\gamma : N \to M$ be a unital normal 2-positive map, and let $\psi, \varphi \in M_*^+$. We say that γ is reversible for $\{\psi, \varphi\}$ if there exists a unital normal 2-positive map $\beta : M \to N$ such that

$$\psi \circ \gamma \circ \beta = \psi \quad \text{and} \quad \varphi \circ \gamma \circ \beta = \varphi.$$

When M_0 is a von Neumann subalgebra of M, we say that M_0 is sufficient for $\{\psi, \varphi\}$ if there exists a unital normal 2-positive map $\beta : M \to M_0$ such that $\psi \circ \beta = \psi$ and $\varphi \circ \beta = \varphi$. If γ is reversible for $\{\psi, \varphi\}$, then the double use of the monotonicity inequality(1.1) gives

$$S_f(\psi \circ \gamma \| \varphi \circ \gamma) = S_f(\psi \| \varphi) \tag{1.2}$$

for any standard f-divergence S_f with an operator convex function f. Similarly, if a subalgebra M_0 is sufficient for $\{\psi, \varphi\}$, then

$$S_f(\psi|_{M_0} \| \varphi|_{M_0}) = S_f(\psi \| \varphi). \tag{1.3}$$

The reverse direction of the above is our concern of Chap. 6, that is, we are interested in the problem whether the equality in (1.2) (resp., in (1.3)) with a finite value implies the reversibility of γ (resp., the sufficiency of M_0) for $\{\psi, \varphi\}$. The problem was formerly studied by Petz [92, 93] (also in [88, Chap. 9]) and by Jenčová and Petz [62, 63], where the treated quantum divergences were the relative entropy and the transition probability (or equivalently, the standard f-divergence for $f(t) = -t^{1/2}$). It was assumed in [92, 93] that $\psi, \varphi \in M_*^+$ are faithful, while the assumption of ψ being faithful was removed in [62]. Moreover, a detailed discussion is found in [53] (also [58]) on the reversibility in the finite-dimensional case via standard f-divergences under a mild assumption on f. Other discussions on the sufficiency in von Neumann algebras are found in [77, 78] for example.

In Secs. 6.1–6.4 we revisit the sufficiency and reversibility results in [92, 93, 62] with slight improvements and furthermore obtain the reversibility (see Theorem 6.19) via a general standard f-divergences under the assumption that the support of μ_f has a limit point in $(0, \infty)$, where μ_f is the representing measure of the integral expression of an operator convex function f. (Note that μ_f is supported on the whole $(0, \infty)$ when $f(t) = t \log t$ for the relative entropy and when $f(t) = -t^{1/2}$ for the transition probability.) In recent papers [59, 60] Jenčová proved the reversibility theorems via the sandwiched Rényi divergences in von Neumann algebras, which are briefly reported in Sec. 6.5.

A significant point in the reversibility theorems in Chap. 6 is that when a unital normal 2-positive map $\gamma : N \to M$ is reversible for $\{\psi, \varphi\}$, we have a reverse map $\beta = \gamma_\varphi^* : M \to N$ that is canonically defined as a certain dual map of γ with respect to φ. The map γ_φ^* is often called Petz' recovery map and its construction described in Sec. 6.1 is originally due to Accardi and Cecchini [1].

In a quantum mechanical system we have final data by measurements (or quantum-classical channels). From this point it is reasonable for us to consider the reversibility for $\{\psi, \varphi\}$ in M_*^+ via a measurement channel, that is, there exist unital normal positive maps $\alpha : \mathcal{A} \to M$ and $\beta : M \to \mathcal{A}$ with a commutative von Neumann algebra $\mathcal{A}$ such that $\psi = \psi \circ \alpha \circ \beta$ and $\varphi = \varphi \circ \alpha \circ \beta$. The approximate reversibility via measurement channels is also considered as follows: there exist unital normal positive maps $\alpha_k : \mathcal{A}_k \to M$ and $\beta_k : M \to \mathcal{A}_k$ $(k \in \mathbb{N})$ with commutative $\mathcal{A}_k$ such that

$$\psi \circ \alpha_k \circ \beta_k \longrightarrow \psi \quad \text{and} \quad \varphi \circ \alpha_k \circ \beta_k \longrightarrow \varphi \tag{1.4}$$

in the $\sigma(M_*, M)$-topology. Then a natural problem is to characterize the above (approximate) reversibility by the approximate attainability condition that there exist $\alpha_k : \mathcal{A}_k \to M$ $(k \in \mathbb{N})$ as above with commutative $\mathcal{A}_k$ such that

$$\lim_k S_f(\psi \circ \alpha_k \| \varphi \circ \alpha_k) = S_f(\psi \| \varphi) < +\infty$$

for a certain f. This condition is equivalently stated as

$$S_f^{\mathrm{meas}}(\psi \| \varphi) = S_f(\psi \| \varphi) < +\infty. \tag{1.5}$$

The problem was studied by Petz [95]. It was shown in [95] that if the approximate reversibility via measurements as in (1.4) is satisfied for faithful ψ, φ, then ψ, φ commute. This means that the approximate reversibility above holds only when ψ, φ are in the classical position. In Chap. 7 we extend this result to ψ, φ with $s(\psi) \leq s(\varphi)$ (see Theorems 7.3 and 7.6). Furthermore, it was shown in [95] that when $\delta\varphi \leq \psi \leq \delta^{-1}\varphi$ for some $\delta > 0$, the condition (1.5) for $S_f = D$, the relative entropy, implies the approximate reversibility in (1.4). We extend this result to S_f with a general f satisfying the same support condition as in Theorem 6.19 and ψ, φ with $s(\psi) = s(\varphi)$ (see Theorem 7.3).

In the proofs of the reversibility and the approximate reversibility theorems in Chaps. 6 and 7, a main ingredient is to obtain the (approximate) preservation of Connes' cocycle derivatives, as done in Secs. 6.3 and 7.1. Both Araki's relative modular operator $\Delta_{\psi,\varphi}$ ([9]) and Connes' cocycle derivative $[D\psi : D\varphi]_t$ ([23]) are certain kinds of the Radon-Nikodym derivative. In the finite-dimensional $M = B(\mathcal{H})$ case with the density operators D_ψ, D_φ, those are given as $\Delta_{\psi,\varphi}(X) = D_\psi X D_\varphi^{-1}$ for $X \in B(\mathcal{H})$, where $M = B(\mathcal{H})$ is represented by the left multiplication on $B(\mathcal{H})$ with the Hilbert-Schmidt inner product, and $[D\psi : D\varphi]_t = D_\psi^{it} D_\varphi^{-it}$ for $t \in \mathbb{R}$. So $\Delta_{\psi,\varphi}$ and $[D\psi : D\varphi]_t$ are essentially in the same vein. A concise account on Connes' cocycle derivatives is included in Appendix A.7 for the convenience of the reader.

The reversibility problem may be considered for the maximal f-divergence $\widehat{S}_f$ in place of S_f in Sec. 6.4. But it is known [52, Example 4.8] that this is not true even in the finite-dimensional case, that is, the equality case (with a finite value) in the monotonicity inequality for $\widehat{S}_f$ in (1.1) does not necessarily imply the reversibility of γ for $\{\psi, \varphi\}$. Even though the situation is negative like that, it is meaningful to characterize the preservation

$$\widehat{S}_f(\psi \circ \gamma \| \varphi \circ \gamma) = \widehat{S}_f(\psi \| \varphi) \tag{1.6}$$

under a unital normal positive map $\gamma : N \to M$, which was discussed in [52] in the finite-dimensional case. Our concern of Chap. 8 is to extend the discussions in [52] to the von Neumann algebra case. Among many equivalent conditions, we show that if (1.6) holds with a finite value for some nonlinear operator convex function f and any $\psi, \varphi \in M_*^+$, then the same equality holds for all operator convex functions f (see Theorem 8.4). In the course of discussions, we are led to consider operator connections (developed by Kubo and Ando [73]) for positive functionals in M_*^+. The idea is interested in its own, so a somewhat detailed account on operator connections of elements in M_*^+ is given in Appendix D separately.

The remaining part is four appendices, which are used in the main body of the monograph. In the first long Appendix A we give concise accounts on selected topics of von Neumann algebras, including the Tomita-Takesaki theory, the standard forms, the relative modular operators, Haagerup's L^p-spaces, Connes' cocycle derivatives, and so on. These might be beneficial for the reader who is not familiar with von Neumann algebras. Appendix B contains a few preliminaries on positive self-adjoint operators. Appendix C is concerned with operator convex functions on $(0, 1)$, which arise in the integral expression of Theorem 4.8 and may be of independent interest.

The last Appendix D is concerned with operator connections of normal positive functionals, which is a variant of Kubo-Ando's theory in [73].

Chapter 2
Standard f-Divergences

2.1 Definition

Let M be a general von Neumann algebra, and M_*^+ be the positive cone of the predual M_* consisting of normal positive linear functionals on M. See the first paragraph of Sec. A.2 for basics of con Neumann algebras. Throughout the text, we consider M in its *standard form*

$$(M, \mathcal{H}, J, \mathcal{P}),$$

that is, M is faithfully represented on a Hilbert space $\mathcal{H}$ with a conjugate-linear involution J and a self-dual cone $\mathcal{P}$ called the *natural cone*, satisfying properties (a)–(d) given in Appendix A.3. Any von Neumann algebra has a standard form, which is unique up to unitary conjugation, see Theorem A.18.

Let $\psi, \varphi \in M_*^+$, whose vector representatives are denoted by $\Psi, \Phi \in \mathcal{P}$, respectively, so that

$$\psi(x) = \langle \Psi, x\Psi \rangle, \qquad \varphi(x) = \langle \Phi, x\Phi \rangle, \qquad x \in M.$$

The M-support projection $s(\varphi) = s_M(\varphi) \in M$ and the M'-support projection $s_{M'}(\varphi) \in M'$ are the orthogonal projections onto $\overline{M'\Phi}$ and onto $\overline{M\Phi}$, respectively, and similarly for $s(\psi) = s_M(\psi)$, $s_{M'}(\psi)$. We have the *relative modular operator* $\Delta_{\psi,\varphi}$ introduced in [9], that is described in Appendix A.4. The support projection of $\Delta_{\psi,\varphi}$ is $s_M(\psi)s_{M'}(\varphi)$. We write the spectral decomposition of $\Delta_{\psi,\varphi}$ as

$$\Delta_{\psi,\varphi} = \int_{[0,\infty)} t \, dE_{\psi,\varphi}(t). \tag{2.1}$$

Assume that $f : (0, \infty) \to \mathbb{R}$ is a convex function. The limits

$$f(0^+) := \lim_{t \searrow 0} f(t), \qquad f'(\infty) := \lim_{t \to \infty} \frac{f(t)}{t}$$

exist in $(-\infty, +\infty]$. Below we will understand the expression $bf(a/b)$ for $a = 0$ or $b = 0$ in the following way:

$$bf(a/b) := \begin{cases} f(0^+)b & \text{for } a = 0, b \geq 0, \\ f'(\infty)a & \text{for } a > 0, b = 0, \end{cases} \tag{2.2}$$

where we use the convention that $(+\infty)0 := 0$ and $(+\infty)a := +\infty$ for $a > 0$. In particular, we set $0f(0/0) = 0$.

The next definition is a specialization of the *quasi-entropy* introduced in [68, 90] with modifications.

Definition 2.1 For each $\varphi, \psi \in M_*^+$ and f as above, we define the *standard f-divergence* $S_f(\psi\|\varphi)$ of ψ, φ by

$$S_f(\psi\|\varphi)$$
$$:= \int_{(0,\infty)} f(t)\, d\|E_{\psi,\varphi}(t)\Phi\|^2 + f(0^+)\varphi(1 - s_M(\psi)) + f'(\infty)\psi(1 - s_M(\varphi)), \tag{2.3}$$

where we note that the integral above is on $(0, \infty)$ instead of $[0, \infty)$.

The integral in (2.3) may be written as $\langle \Phi, f(\Delta_{\psi,\varphi})\Phi \rangle$ in terms of the self-adjoint operator $f(\Delta_{\psi,\varphi}) := \int_{(0,\infty)} f(t)\, dE_{\psi,\varphi}(t)$ on $s_M(\psi)s_{M'}(\varphi)\mathcal{H}$, though, to be more precise, it should be understood in the sense of a lower-bounded form (see [99]). From the convexity of f, there are $a, b \in \mathbb{R}$ such that $f(t) \geq a + bt$, $t > 0$. Then expression (2.3) is well defined with values in $(-\infty, +\infty]$ since

$$\int_{(0,\infty)} f(t)\, d\|E_{\psi,\varphi}(t)\Phi\|^2 \geq \int_{(0,\infty)} (a + bt)\, d\|E_{\psi,\varphi}(t)\Phi\|^2$$
$$= a\|s_M(\psi)s_{M'}(\varphi)\Phi\|^2 + b\|\Delta_{\psi,\varphi}^{1/2}\Phi\|^2$$
$$= a\|s_M(\psi)\Phi\|^2 + b\|s_M(\varphi)\Psi\|^2$$
$$= a\varphi(s_M(\psi)) + b\psi(s_M(\varphi)) > -\infty$$

thanks to $J\Delta_{\psi,\varphi}^{1/2}\Phi = s_M(\varphi)\Psi$. By the above computation and (2.3) we also see that

$$S_{a+bt}(\psi\|\varphi) = a\varphi(1) + b\psi(1). \tag{2.4}$$

Example 2.2 The following are the special cases of M abelian (classical case) and $M = B(\mathcal{H})$.

(1) Let $M = L^\infty(X, \mu)$ be an abelian von Neumann algebra, where $(X, \mathcal{X}, \mu)$ is a σ-finite measure space. For every $\psi, \varphi \in L^1(X, \mu)_+ \cong M_*^+$, since $\Delta_{\psi,\varphi}$ is the multiplication of $1_{\{\varphi>0\}}(\psi/\varphi)$ (see Examples A.14 (1) and A.24 (1)), we have

$$S_f(\psi\|\varphi) = \int_{\{\psi>0\}\cap\{\varphi>0\}} \varphi f(\psi/\varphi)\, d\mu + f(0^+) \int_{\{\psi=0\}} \varphi\, d\mu + f'(\infty) \int_{\{\varphi=0\}} \psi\, d\mu,$$

which equals the classical f-divergence $S_f(\psi|\varphi) = \int_X \psi f(\psi/\varphi)\,d\mu$ under the convention (2.2).

(2) Let $M = B(\mathcal{H})$ be a factor of type I, where $\mathcal{H}$ is an arbitrary Hilbert space. The standard form of $B(\mathcal{H})$ is described in Example A.14 (2). For every $\psi, \varphi \in B(\mathcal{H})_*^+$ the relative modular operators $\Delta_{\psi,\varphi}$ on $C_2(\mathcal{H})$ is given in Example A.24 (2) in terms of the spectral decompositions $D_\psi = \sum_{a>0} a P_a$ and $D_\varphi = \sum_{b>0} b Q_b$ of the density operators of ψ, φ. From (A.15) the definition of $S_f(\psi\|\varphi)$ in (2.3) is rewritten as

$$S_f(\psi\|\varphi) = \sum_{a,b>0} b f(ab^{-1}) \operatorname{Tr} P_a Q_b + f(0^+) \operatorname{Tr}(I - D_\psi^0) D_\varphi + f'(\infty) \operatorname{Tr} D_\psi (I - D_\varphi^0),$$

where D_φ^0 is the support projection of D_φ. The above coincides with an expression in [52, Proposition 3.2] when $\dim \mathcal{H} < \infty$.

The next proposition summarizes simple properties of $S_f(\psi\|\varphi)$.

Proposition 2.3 ([48]) *Let* $\psi, \varphi \in M_*^+$.

(1) Invariance: *If* $\alpha : N \to M$ *is a $*$-isomorphism between von Neumann algebras, then*

$$S_f(\psi \circ \alpha \| \varphi \circ \alpha) = S_f(\psi\|\varphi).$$

(2) *In the case* $\psi = 0$ *or* $\varphi = 0$ *or* $\psi = \varphi$,

$$S_f(0\|\varphi) = f(0^+)\varphi(1), \quad S_f(\psi\|0) = f'(\infty)\psi(1), \quad S_f(\varphi\|\varphi) = f(1)\varphi(1).$$

(3) Homogeneity: *For every* $\lambda \in [0, \infty)$,

$$S_f(\lambda\psi\|\lambda\varphi) = \lambda S_f(\psi\|\varphi).$$

(4) Additivity: *Let* $M = M_1 \oplus M_2$ *be the direct sum of von Neumann algebras* M_1 *and* M_2. *If* $\psi_i, \varphi_i \in (M_i)_*^+$ *for* $i = 1, 2$, *then*

$$S_f(\psi_1 \oplus \psi_2 \| \varphi_1 \oplus \varphi_2) = S_f(\psi_1\|\varphi_1) + S_f(\psi_2\|\varphi_2).$$

(5) Transpose: *Let* $\widetilde{f}$ *be the* transpose *of* f *defined by*

$$\widetilde{f}(t) := t f(t^{-1}), \qquad t \in (0, \infty).$$

Then $S_f(\psi\|\varphi) = S_{\widetilde{f}}(\varphi\|\psi)$. *(Note that* $\widetilde{f}(0^+) = f'(\infty)$ *and* $\widetilde{f}'(\infty) = f(0^+)$.)

Example 2.4 The relative entropy is the most utilized quantum divergence, which is the standard f-divergence for $f(t) = t \log t$ on $(0, \infty)$ with the transpose $\widetilde{f}(t) = -\log t$. Since $f(0^+) = 0$ and $f'(\infty) = +\infty$, the *relative entropy* $D(\psi\|\varphi)$ of $\psi, \varphi \in M_*^+$ [8, 9] is

$$D(\psi\|\varphi) = S_{t\log t}(\psi\|\varphi) = \begin{cases} \int_{(0,\infty)} t\log t \, d\|E_{\psi,\varphi}(t)\Phi\|^2 & \text{if } s(\psi) \leq s(\varphi), \\ +\infty & \text{otherwise,} \end{cases}$$

or $D(\psi\|\varphi) = S_{-\log t}(\varphi\|\psi)$. When $M = B(\mathcal{H})$, $D(\psi\|\varphi)$ is *Umegaki's relative entropy* [114], whose alternative familiar expression is

$$D(\psi\|\varphi) = \begin{cases} \text{Tr}\, D_\psi(\log D_\psi - \log D_\varphi) & \text{if } D_\psi^0 \leq D_\varphi^0, \\ +\infty & \text{otherwise,} \end{cases}$$

where D_ψ, D_φ is as in Example 2.2 (2).

2.2 Variational Expression of Standard f-Divergences

We extend the variational expression of the relative entropy given in [70] to standard f-divergences. The extended expression will be quite useful to verify various properties of standard f-divergences. Throughout this section we assume that a function $f : (0,\infty) \to \mathbb{R}$ is *operator convex*, i.e., the operator inequality

$$f(\lambda A + (1-\lambda)B) \leq \lambda f(A) + (1-\lambda)f(B), \qquad 0 \leq \lambda \leq 1$$

holds for every invertible $A, B \in B(\mathcal{H})_+$ of any $\mathcal{H}$. Also, a function $h : (0,\infty) \to \mathbb{R}$ is said to be *operator monotone* if $A \leq B \implies h(A) \leq h(B)$ for every invertible $A, B \in B(\mathcal{H})_+$ of any $\mathcal{H}$. It is well-known that an operator monotone function h on $(0,\infty)$ is automatically *operator concave* (i.e., $-h$ is operator convex). For general theory on operator monotone and operator convex functions, see, e.g., [18, 45]. Note here [52, Proposition A.1] that if f is operator convex on $(0,\infty)$, then so is the transpose $\widetilde{f}(t) = tf(t^{-1})$.

Recall [75] (also [35, Theorem 5.1] for a more general form) that any operator convex function f on $(0,\infty)$ has the integral expression

$$f(t) = a + b(t-1) + c(t-1)^2 + \int_{[0,\infty)} \frac{(t-1)^2}{t+s} \, d\mu(s), \quad t \in (0,\infty), \qquad (2.5)$$

where $a, b \in \mathbb{R}$, $c \geq 0$ and μ is a positive measure on $[0,\infty)$ satisfying $\int_{[0,\infty)} (1+s)^{-1} \, d\mu(s) < +\infty$, and moreover a, b, c and μ are uniquely determined. Letting $d := \mu(\{0\}) \geq 0$ we also write

$$f(t) = a + b(t-1) + c(t-1)^2 + d\frac{(t-1)^2}{t} + \int_{(0,\infty)} \frac{(t-1)^2}{t+s} \, d\mu(s), \quad t \in (0,\infty).$$

One can easily verify that

$$f(0^+) = a - b + c + (+\infty)d + \int_{(0,\infty)} s^{-1}\, d\mu(s),$$

$$f'(\infty) = b + (+\infty)c + d + \int_{(0,\infty)} d\mu(s).$$

For each $n \in \mathbb{N}$ we define

$$f_n(t) := a + b(t-1) + c\,\frac{n(t-1)^2}{t+n} + d\,\frac{(t-1)^2}{t+(1/n)}$$
$$+ \int_{[1/n,n]} \frac{(t-1)^2}{t+s}\, d\mu(s), \qquad t \in (0,\infty). \tag{2.6}$$

Then f_n is operator convex on $(0,\infty)$, and we have

$$f_n(0^+) = a - b + c + nd + \int_{[1/n,n]} s^{-1}\, d\mu(s) < +\infty, \tag{2.7}$$

$$f_n'(\infty) = b + nc + d + \int_{[1/n,n]} d\mu(s) < +\infty, \tag{2.8}$$

$$f_n(0^+) \nearrow f(0^+), \quad f_n'(\infty) \nearrow f'(\infty), \quad f_n(t) \nearrow f(t) \tag{2.9}$$

as $n \to \infty$ for all $t \in (0,\infty)$. Moreover, for every $\psi, \varphi \in M_*^+$ we have

$$S_{f_n}(\psi\|\varphi) \nearrow S_f(\psi\|\varphi) \qquad \text{as } n \nearrow \infty.$$

Next, we define

$$h_n(t) := \int_{(0,\infty)} \frac{t(1+s)}{t+s}\, d\nu_n(s), \qquad t \in [0,\infty),$$

where ν_n is a finite positive measure supported on $[1/n, n]$ given by

$$d\nu_n(s) := c(1+n)\delta_n + d(1+n)\delta_{1/n} + 1_{[1/n,n]}(s)\frac{1+s}{s}\, d\mu(s) \tag{2.10}$$

with the point masses δ_n at n and $\delta_{1/n}$ at $1/n$. Then h_n is an operator monotone function on $[0,\infty)$, and f_n is written as

$$f_n(t) = f_n(0^+) + f_n'(\infty)t - h_n(t), \qquad t \in (0,\infty). \tag{2.11}$$

Now, let L be a subspace of M containing 1, and assume that L is dense in M with respect to the strong* operator topology. Our variational expression of $S_f(\psi\|\varphi)$ is

Theorem 2.5 ([48]) *Let f be an operator convex function on $(0,\infty)$. For each $n \in \mathbb{N}$ let $f_n(0^+)$, $f_n'(\infty)$ and ν_n be given in (2.7), (2.8) and (2.10), respectively. Then for every $\psi, \varphi \in M_*^+$,*

$$S_f(\psi\|\varphi) = \sup_{n\in\mathbb{N}} \sup_{x(\cdot)} \left[f_n(0^+)\varphi(1) + f_n'(\infty)\psi(1) \right.$$

$$\left. - \int_{[1/n,n]} \{\varphi((1-x(s))^*(1-x(s))) + s^{-1}\psi(x(s)x(s)^*)\}(1+s)\,d\nu_n(s) \right], \quad (2.12)$$

where the second supremum is taken over all L-valued (finitely many values) step functions $x(\cdot)$ on $(0,\infty)$.

To prove this, note from (2.11) and (2.4) that

$$S_{f_n}(\psi\|\varphi) = f_n(0^+)\varphi(1) + f_n'(\infty)\psi(1) + S_{-h_n}(\psi\|\varphi)$$

$$= f_n(0^+)\varphi(1) + f_n'(\infty)\psi(1) - \int_{(0,\infty)} h_n(t)\,d\|E_{\psi,\varphi}(t)\Phi\|^2.$$

Since $h_n(0) = h_n'(\infty) = 0$, we can apply [70, Theorem 2.2] to the above integral term, so we have

$$S_{f_n}(\psi\|\varphi) = \sup_{x(\cdot)} \left[f_n(0^+)\varphi(1) + f_n'(\infty)\psi(1) \right.$$

$$\left. - \int_{[1/n,n]} \{\varphi((1-x(s))^*(1-x(s))) + s^{-1}\psi(x(s)x(s)^*)\}(1+s)\,d\nu_n(s) \right].$$

Taking $\sup_n$ of both sides of the above yields (2.12).

Example 2.6 Consider $f(t) = -\log t$, whose integral expression in (2.5) is

$$-\log t = -(t-1) + \int_{(0,\infty)} \frac{(t-1)^2}{(t+s)(1+s)^2}\,ds.$$

Hence, in this case,

$$a = c = d = 0, \quad b = -1, \quad d\mu(s) = \frac{1}{(1+s)^2}\,ds, \quad f(0^+) = +\infty, \quad f'(\infty) = 0.$$

Moreover, compute

$$f_n(0^+) = 1 + \int_{1/n}^n \frac{1}{s(1+s)^2}\,ds = \frac{2}{n+1} + \log n,$$

$$f_n'(\infty) = -1 + \int_{1/n}^n \frac{1}{(1+s)^2}\,ds = -\frac{2}{n+1},$$

$$d\nu_n(s) = 1_{[1/n,n]}(s)\frac{1}{s(1+s)}\,ds.$$

For every $\psi, \varphi \in M_*^+$ the *relative entropy* $D(\varphi\|\psi)$ is $S_{t\log t}(\varphi\|\psi) = S_{-\log t}(\psi\|\varphi)$ (see Example 2.4), for which one can write expression (2.12) as

$$D(\varphi\|\psi) = \sup_{n\in\mathbb{N}} \sup_{x(\cdot)} \left[\varphi(1)\log n + (\varphi(1) - \psi(1))\frac{2}{n+1} \right.$$
$$\left. - \int_{[1/n,n]} \{\varphi((1 - x(s))^*(1 - x(s))) + s^{-1}\psi(x(s)x(s)^*)\}s^{-1}\,ds \right].$$

This expression is similar to but a bit different from the variational expression

$$D(\varphi\|\psi) = \sup_{n\in\mathbb{N}} \sup_{x(\cdot)} \left[\varphi(1)\log n \right.$$
$$\left. - \int_{[1/n,\infty)} \{\varphi((1 - x(s))^*(1 - x(s))) + s^{-1}\psi(x(s)x(s)^*)\}s^{-1}\,ds \right]$$

in [70, Theorem 3.2].

2.3 Properties of Standard f-Divergences

Most of the important properties of standard f-divergences can immediately be verified from the variational expression in Theorem 2.5, similarly to [70] where the relative entropy was treated. In the rest of this section assume that f is an operator convex function on $(0,\infty)$.

Theorem 2.7 ([48]) *Let $\psi, \varphi, \psi_i, \varphi_i \in M_*^+$ for $i = 1, 2$.*

(i) Joint lower semicontinuity: *The map $(\psi, \varphi) \in M_*^+ \times M_*^+ \mapsto S_f(\psi\|\varphi) \in (-\infty, \infty]$ is jointly lower semicontinuous in the $\sigma(M_*, M)$-topology.*

(ii) Joint convexity: *The map in (i) is jointly convex and jointly subadditive, i.e., for every $\psi_i, \varphi_i \in M_*^+$, $1 \leq i \leq k$,*

$$S_f\left(\sum_{i=1}^k \psi_i \,\Big\|\, \sum_{i=1}^k \varphi_i\right) \leq \sum_{i=1}^k S_f(\psi_i\|\varphi_i).$$

(iii) *If $f(0^+) \leq 0$ and $\varphi_1 \leq \varphi_2$, then $S_f(\psi\|\varphi_1) \geq S_f(\psi\|\varphi_2)$. Also, if $f'(\infty) \leq 0$ and $\psi_1 \leq \psi_2$, then $S_f(\psi_1\|\varphi) \geq S_f(\psi_2\|\varphi)$.*

(iv) Monotonicity: *Let N be another von Neumann algebra and $\gamma : N \to M$ be a unital positive linear map that is normal (i.e., if $\{x_\alpha\}$ is an increasing net in M_+ with $x_\alpha \nearrow x \in M_+$, then $\gamma(x_\alpha) \nearrow \gamma(x)$) and is a Schwarz map (i.e., $\gamma(x^*x) \geq \gamma(x)^*\gamma(x)$ for all $x \in N$). Then*

$$S_f(\psi \circ \gamma\|\varphi \circ \gamma) \leq S_f(\psi\|\varphi). \tag{2.13}$$

In particular, if N is a unital von Neumann subalgebra of M, then

$$S_f(\psi|_N\|\varphi|_N) \leq S_f(\psi\|\varphi). \tag{2.14}$$

(v) Peierls-Bogolieubov inequality:

$$S_f(\psi\|\varphi) \geq \varphi(1)f(\psi(1)/\varphi(1)). \tag{2.15}$$

Assume that f is non-linear and $\psi, \varphi \neq 0$. Then equality holds in (2.15) if and only if $\psi = (\psi(1)/\varphi(1))\varphi$.

(vi) Strict positivity: *Assume that f is non-linear with $f(1) = 0$ and $\psi(1) = \varphi(1) > 0$. Then $S_f(\psi\|\varphi) \geq 0$, and $S_f(\psi\|\varphi) = 0$ if and only if $\psi = \varphi$.*

(vii) Martingale convergence: *If $\{M_\alpha\}$ is an increasing net of unital von Neumann subalgebras of M such that $(\bigcup_\alpha M_\alpha)'' = M$, then*

$$S_f(\psi|_{M_\alpha}\|\varphi|_{M_\alpha}) \nearrow S_f(\psi\|\varphi).$$

For $\varphi \in M_*^+$ and a projection $e \in M$, we write $e\varphi e$ for the restriction of φ to the reduced von Neumann algebra eMe. The next corollary is easily seen from properties in Theorem 2.7.

Corollary 2.8 (1) *If $e \in M$ is a projection such that $s_M(\psi), s_M(\varphi) \leq e$, then*

$$S_f(\psi\|\varphi) = S_f(e\psi e\|e\varphi e).$$

(2) *If $\psi_i, \varphi_i \in M_*^+$, $i = 1, 2$, and $s_M(\psi_1) \vee s_M(\varphi_1) \perp s_M(\psi_2) \vee s_M(\varphi_2)$, then*

$$S_f(\psi_1 + \psi_2\|\varphi_1 + \varphi_2) = S_f(\psi_1\|\varphi_1) + S_f(\psi_2\|\varphi_2).$$

(3) *If $\omega_1, \omega_2 \in M_*^+$ and $S_f(\omega_1\|\omega_2) < +\infty$, then for every $\psi, \varphi \in M_*^+$,*

$$S_f(\psi\|\varphi) = \lim_{\varepsilon \searrow 0} S_f(\psi + \varepsilon\omega_1\|\varphi + \varepsilon\omega_2).$$

In particular, for every $\psi, \varphi, \omega \in M_*^+$,

$$S_f(\psi\|\varphi) = \lim_{\varepsilon \searrow 0} S_f(\psi + \varepsilon\omega\|\varphi + \varepsilon\omega). \tag{2.16}$$

The relative entropy $D(\psi\|\varphi)$ has the scaling property

$$D(\lambda\psi\|\mu\varphi) = \lambda D(\psi\|\varphi) + \lambda\psi(1)\log\frac{\lambda}{\mu}, \qquad \lambda, \mu > 0,$$

from which it is clear that if $D(\psi\|\varphi) < +\infty$ then $D(\lambda\psi\|\mu\varphi) < +\infty$ for all $\lambda, \mu > 0$. The next proposition says that this holds for any standard f-divergence, which will be used in the proof of Theorem 7.6 of Sec. 7.2.

Proposition 2.9 *Let $\psi, \varphi \in M_*^+$. If $S_f(\psi\|\varphi) < +\infty$, then $S_f(\lambda\psi\|\mu\varphi) < +\infty$ and $S_f(\lambda\psi + \mu\varphi\|\varphi) < +\infty$ for all $\lambda, \mu > 0$.*

Proof Since $S_f(\lambda\psi\|\lambda\mu) = \lambda S_f(\psi\|\varphi)$, we may assume $\mu = 1$ for the first assertion. Since $\Delta_{\lambda\psi,\varphi} = \lambda\Delta_{\psi,\varphi}$, we have

$$\Delta_{\lambda\psi,\varphi} = \int_{[0,\infty)} \lambda t \, dE_{\psi,\varphi}(t) = \int_{[0,\infty)} t \, dE_{\psi,\varphi}(\lambda^{-1}t)$$

Note that

$$S_f(\psi\|\varphi) = \int_{(0,\infty)} f(t)\, d\|E_{\psi,\varphi}(t)\Phi\|^2$$
$$+ f(0^+)\varphi(1 - s(\psi)) + f'(+\infty)\psi(1 - s(\varphi)) < +\infty,$$

$$S_f(\lambda\psi\|\varphi) = \int_{(0,\infty)} f(t)\, d\|E_{\psi,\varphi}(\lambda^{-1}t)\Phi\|^2$$
$$+ f(0^+)\varphi(1 - s(\psi)) + f'(+\infty)\lambda\psi(1 - s(\varphi))$$
$$= \int_{(0,\infty)} f(\lambda t)\, d\|E_{\psi,\varphi}(t)\Phi\|^2$$
$$+ f(0^+)\varphi(1 - s(\psi)) + f'(+\infty)\lambda\psi(1 - s(\varphi)),$$

so it suffices to prove that

$$\int_{(0,\infty)} f(\lambda t)\, d\|E_{\psi,\varphi}(t)\Phi\|^2 < +\infty. \tag{2.17}$$

One can write $f(t) = f_0(t) + at + b$ for $t > 0$ with an operator convex function f_0 on $(0,\infty)$ such that $f_0(1) = 0$ and $f_0(t) \geq 0$ for all $t > 0$. So we may assume that $f(1) = 0$ and $f(t) \geq 0$ for all $t > 0$.

Assume that $\lambda > 1$. Note that $g(t) := \frac{f(t)-f(1)}{t-1}$ (where $g(1) = f'(1) = 0$) is operator monotone and so operator concave on $(0,\infty)$. Hence for every $t > 1$ one has $\frac{g(t)}{t-1} \geq \frac{g(\lambda t)}{\lambda t-1}$ so that

$$f(\lambda t) \leq \left(\frac{\lambda t - 1}{t - 1}\right)^2 f(t), \qquad t > 1.$$

Since $\left(\frac{\lambda t-1}{t-1}\right)^2 \leq 2\lambda^2$ for all sufficiently large t, one can choose a $c > 0$ such that

$$f(\lambda t) \leq 2\lambda^2 f(t) + c, \qquad t \geq 1,$$

and so

$$\int_{[1,\infty)} f(\lambda t)\, d\|E_{\psi,\varphi}(t)\Phi\|^2 \leq \int \{2\lambda^2 f(t) + c\}\, d\|E_{\psi,\varphi}(t)\Phi\|^2 < +\infty.$$

Moreover, for every $t \in (0, \lambda^{-1})$ one has $\frac{f(t)}{t-1} \leq \frac{f(\lambda t)}{\lambda t-1}$ so that

$$f(\lambda t) \leq \frac{1 - \lambda t}{1 - t} f(t) \leq f(t), \qquad 0 < t < \lambda^{-1}.$$

Hence one can choose a $d > 0$ such that

$$f(\lambda t) \leq f(t) + d, \qquad 0 < t \leq 1,$$

and so

$$\int_{(0,1]} f(\lambda t)\, d\|E_{\psi,\varphi}(t)\Phi\|^2 \le \int_{(0,1]} \{f(t) + d\}\, d\|E_{\psi,\varphi}(t)\Phi\|^2 < +\infty.$$

Therefore, (2.17) follows.

Next, assume that $0 < \lambda < 1$. One can apply the above argument to the transpose $\widetilde{f}(t) := t f(t^{-1})$ and $\lambda^{-1} > 1$. Hence one can choose $c', d' > 0$ such that

$$\widetilde{f}(\lambda^{-1}t) \le 2\lambda^{-2}\tilde{f}(t) + c', \qquad t \ge 1,$$
$$\widetilde{f}(\lambda^{-1}t) \le \tilde{f}(t) + d', \qquad 0 < t \le 1.$$

By replacing t with t^{-1} these are rewritten as

$$f(\lambda t) \le 2\lambda^{-1}\tilde{f}(t) + c'\lambda t, \qquad 0 < t \le 1,$$
$$f(\lambda t) \le \lambda f(t) + d'\lambda t, \qquad t \ge 1,$$

from which (2.17) follows as in the above case where $\lambda > 1$.

For the second assertion we may assume by the first that $\lambda + \mu = 1$. Then Theorem 2.7 (ii) gives $S_f(\lambda\psi + \mu\varphi\|\varphi) \le \lambda S_f(\psi\|\varphi) + \mu S_f(\varphi\|\varphi) < +\infty.$ $\qquad\square$

The next continuity property is not included in the martingale convergence in Theorem 2.7, since eMe is not a unital von Neumann subalgebra of M.

Proposition 2.10 ([48]) *Let $\{e_\alpha\}$ be an increasing net of projections in M such that $e_\alpha \nearrow 1$. Then for every $\psi, \varphi \in M_*^+$,*

$$\lim_\alpha S_f(e_\alpha\psi e_\alpha\|e_\alpha\varphi e_\alpha) = S_f(\psi\|\varphi).$$

When $f \ge 0$ in Proposition 2.10, from the monotonicity of S_f we see that $S_f(e_\alpha\psi e_\alpha\|e_\alpha\varphi e_\alpha)$ is increasing as $e_\alpha \nearrow 1$. But this is not the case unless $f \ge 0$.

Chapter 3
Rényi divergences and sandwiched Rényi divergences

3.1 Rényi Divergences

The notion of Rényi divergences $D_\alpha(\psi\|\varphi)$ for $\alpha \geq 0$ in the general von Neumann algebra setting is defined as follows:

Definition 3.1 Let $\psi, \varphi \in M_*^+$. Since the vector representative Φ of φ is in $\mathcal{D}(\Delta_{\psi,\varphi}^{1/2})$, the domain of $\Delta_{\psi,\varphi}^{1/2}$, note that $\Phi \in \mathcal{D}(\Delta_{\psi,\varphi}^{\alpha/2})$ for any $\alpha \in [0, 1]$. Define the quantities $Q_\alpha(\psi\|\varphi)$ for $\alpha \geq 0$ as follows: When $0 \leq \alpha < 1$,

$$Q_\alpha(\psi\|\varphi) := \|\Delta_{\psi,\varphi}^{\alpha/2}\Phi\|^2 \in [0, \infty), \tag{3.1}$$

and when $\alpha > 1$,

$$Q_\alpha(\psi\|\varphi) := \begin{cases} \|\Delta_{\psi,\varphi}^{\alpha/2}\Phi\|^2 & \text{if } s_M(\psi) \leq s_M(\varphi) \text{ and } \Phi \in \mathcal{D}(\Delta_{\psi,\varphi}^{\alpha/2}), \\ +\infty & \text{otherwise.} \end{cases} \tag{3.2}$$

Moreover, when $\alpha = 1$, define $Q_1(\psi\|\varphi) := \psi(1)$.

Assume that $\psi \neq 0$. For each $\alpha \in [0, \infty) \setminus \{1\}$ the α-*Rényi divergence* $D_\alpha(\psi\|\varphi)$ is defined by

$$D_\alpha(\psi\|\varphi) := \frac{1}{\alpha - 1} \log \frac{Q_\alpha(\psi\|\varphi)}{\psi(1)}. \tag{3.3}$$

In particular, $Q_0(\alpha\|\varphi) = \varphi(s_M(\psi))$ and $D_0(\psi\|\varphi) = -\log[\varphi(s_M(\psi))/\psi(1)]$.

Define convex functions f_α on $[0, \infty)$ by

$$f_\alpha(t) := \begin{cases} t^\alpha & \text{if } \alpha \geq 1, \\ -t^\alpha & \text{if } 0 < \alpha < 1. \end{cases}$$

Then for every $\psi, \varphi \in M_*^+$, $Q_\alpha(\psi\|\varphi)$ is given as

$$Q_\alpha(\psi\|\varphi) = \begin{cases} S_{f_\alpha}(\psi\|\varphi) & \text{if } \alpha \geq 1, \\ -S_{f_\alpha}(\psi\|\varphi) & \text{if } 0 < \alpha < 1. \end{cases}$$

When $0 < \alpha < 1$, since $f_\alpha(0) = f'_\alpha(\infty) = 0$, by (3.1)

$$S_{f_\alpha}(\psi\|\varphi) = \int_{(0,\infty)} (-t^\alpha)\, d\|E_{\psi,\varphi}(t)\Phi\|^2 = -Q_\alpha(\psi\|\varphi).$$

When $\alpha = 1$, by (2.4), $S_{f_1}(\psi\|\varphi) = \psi(1) = Q_1(\psi\|\varphi)$. When $\alpha > 1$, since $f_\alpha(0) = 0$ and $f'_\alpha(\infty) = +\infty$, by (3.2),

$$S_{f_\alpha}(\psi\|\varphi) = \int_{(0,\infty)} t^\alpha\, d\|E_{\psi,\varphi}(t)\Phi\|^2 + (+\infty)\psi(1 - s_M(\varphi)) = Q_\alpha(\psi\|\varphi).$$

Therefore, $Q_\alpha(\psi\|\varphi)$ is essentially a standard f-divergence and so $D_\alpha(\psi\|\varphi)$ is a variant of standard f-divergences.

Properties of Q_α and D_α are summarized in the next theorem, some of which are also found in, e.g., [76, 82, 46, 81, 17] though mostly in the finite-dimensional situation. Many of them are shown based on Theorem 2.7 applied to S_{f_α}.

Theorem 3.2 ([48]) *Let $\psi, \varphi \in M_*^+$ with $\psi \neq 0$.*

(1) *Unless $s_M(\psi) \perp s_M(\varphi)$, $Q_\alpha(\psi\|\varphi) > 0$ for all $\alpha \geq 0$ and the function $\alpha \in [0, \infty) \mapsto \log Q_\alpha(\psi\|\varphi)$ is convex.*
(2) Limit values: *The limit $D_1(\psi\|\varphi) := \lim_{\alpha \nearrow 1} D_\alpha(\psi\|\varphi)$ exists and*

$$D_1(\psi\|\varphi) = \frac{D(\psi\|\varphi)}{\psi(1)},$$

where $D(\psi\|\varphi)$ is the relative entropy. Moreover, if $D_\alpha(\psi\|\varphi) < +\infty$ for some $\alpha > 1$, then $\lim_{\alpha \searrow 1} D_\alpha(\psi\|\varphi) = D_1(\psi\|\varphi)$. Moreover, if $\psi(1) = \varphi(1)$, then $\lim_{\alpha \searrow 0} \frac{1}{\alpha} D_\alpha(\psi\|\varphi) = D_1(\varphi\|\psi)$.
(3) *The function $\alpha \in [0, \infty) \mapsto D_\alpha(\psi\|\varphi)$ is monotone increasing.*
(4) Joint lower semicontinuity: *For every $\alpha \in [0, 2]$, the map*

$$(\psi, \varphi) \in (M_*^+ \setminus \{0\}) \times M_*^+ \longmapsto D_\alpha(\psi\|\varphi) \in (-\infty, +\infty]$$

is jointly lower semicontinuous in the $\sigma(M_*, M)$-topology.
(5) *The map $(\psi, \varphi) \in M_*^+ \times M_*^+ \mapsto Q_\alpha(\psi\|\varphi)$ is jointly concave and jointly superadditive for $0 \leq \alpha \leq 1$, and jointly convex and jointly subadditive for $1 \leq \alpha \leq 2$. Hence, when $0 \leq \alpha \leq 1$, $D_\alpha(\psi\|\varphi)$ is jointly convex on $\{(\psi, \varphi) \in M_*^+ \times M_*^+ : \psi(1) = c\}$ for any fixed $c > 0$.*
(6) *Let $\psi_i, \varphi_i \in M_*^+$ for $i = 1, 2$. If $0 \leq \alpha < 1$, $\psi_1 \leq \psi_2$ and $\varphi_1 \leq \varphi_2$, then $Q_\alpha(\psi_1\|\varphi_1) \leq Q_\alpha(\psi_2\|\varphi_2)$. If $1 \leq \alpha \leq 2$ and $\varphi_1 \leq \varphi_2$, then $Q_\alpha(\psi\|\varphi_1) \geq Q_\alpha(\psi\|\varphi_2)$. If $\varphi_1 \leq \varphi_2$, then $D_\alpha(\psi\|\varphi_1) \geq D_\alpha(\psi\|\varphi_2)$ for all $\alpha \in [0, 2]$.*
(7) Monotonicity: *For each $\alpha \in [0, 2]$, $D_\alpha(\psi\|\varphi)$ is monotone under unital normal Schwarz maps, i.e.,*

$$D_\alpha(\psi \circ \gamma \| \varphi \circ \gamma) \leq D_\alpha(\psi\|\varphi)$$

for any unital normal Schwarz map $\gamma : N \to M$ as in Theorem 2.7 (iv).
(8) Strict positivity: *Let $\alpha \in (0, \infty)$ and $\psi, \varphi \neq 0$. The inequality*

$$D_\alpha(\psi \| \varphi) \geq \log \frac{\psi(1)}{\varphi(1)}$$

holds, and equality holds in (3.10) if and only if $\psi = (\psi(1)/\varphi(1))\varphi$. If $\psi(1) = \varphi(1)$, then $D_\alpha(\psi \| \varphi) \geq 0$, and $D_\alpha(\psi \| \varphi) = 0$ if and only if $\psi = \varphi$.

Remark 3.3 (1) In Theorem 3.2 (2) the assumption that $D_\alpha(\psi \| \varphi) < +\infty$ for some $\alpha > 1$ is essential to have $\lim_{\alpha \searrow 1} D_\alpha(\psi \| \varphi) = D_1(\psi \| \varphi)$ in the infinite-dimensional case.

(2) The convexity of Q_α for $1 \leq \alpha \leq 2$ in Theorem 3.2 (5) cannot extend to $\alpha > 2$ even in the finite-dimensional case and in separate arguments. This implies that the monotonicity property of D_α in (7) fails to hold for $\alpha > 2$, because the monotonicity of Q_α under unital CP maps yields its joint convexity (see Sec. 6.1 for the definition of CP (completely positive) maps). Also, the monotone decreasing of $\varphi \mapsto Q_\alpha(\psi \| \varphi)$ for $1 \leq \alpha \leq 2$ in (6) cannot extend to $\alpha > 2$.

(3) In connection with Theorem 3.2 (5), in the finite-dimensional case it was shown [82, Theorem II.1] that $\varphi \mapsto D_\alpha(\psi \| \varphi)$ with $\psi \in M_*^+ \setminus \{0\}$ fixed is convex on M_*^+ for any $\alpha \in [0, 2]$. It is natural to expect that this extends to the general von Neumann algebra case.

3.2 Description of Rényi Divergences

When $M = B(\mathcal{H})$ on a finite-dimensional Hilbert space $\mathcal{H}$, the main term Q_α of the Rényi divergence D_α in (3.3) has a simple expression

$$Q_\alpha(\psi \| \varphi) = \operatorname{Tr}\left(D_\psi^\alpha D_\varphi^{1-\alpha}\right) \tag{3.4}$$

in terms of the density operators D_ψ, D_φ, where $D_\varphi^{1-\alpha}$ is defined with restriction to the support of φ. In this section we give a similar expression of Q_α in the general von Neumann algebra case.

It is convenient for us to work in the framework of *Haagerup's L^p-spaces $L^p(M)$*, which are constructed inside the space of τ-measurable operators affiliated with the crossed product $M \rtimes_\sigma \mathbb{R}$ of M by the modular automorphism group σ_t ($t \in \mathbb{R}$). See Appendix A.6 for Haagerup's L^p-spaces and the beginning of Appendix A.5 for operators affiliated with M. Here recall that $M_* \cong L^1(M)$ by an order-isomorphic linear bijection $\varphi \in M_* \mapsto h_\varphi \in L^1(M)$, the tr functional on $L^1(M)$ is defined by $\operatorname{tr}(h_\varphi) = \varphi(1)$ for $\varphi \in M_*$, and the standard form of M is given as

$$(M, L^2(M), J = {}^*, L^2(M)_+),$$

where $x \in M$ (resp., $JxJ \in M'$) acts as the left (resp., the right) multiplication by x on $L^2(M)$. In particular, for any $\varphi \in M_*^+$ the support projection $s(\varphi) = s_M(\varphi)$ is

equal to $s(h_\varphi)$ and it acts as the left multiplication on $L^2(M)$ while the M'-support projection $s_{M'}(\varphi) = Js(\varphi)J$ does as the right multiplication by $s(\varphi)$. See Appendix A.6 for more detailed description of Haagerup's $L^p(M)$.

Let $\psi, \varphi \in M_*^+$ with corresponding $h_\psi, h_\varphi \in L^1(M)_+$. For each $p \geq 0$ define

$$\mathcal{D}_p(\psi, \varphi) := \{\xi \in L^2(M) : h_\psi^p \xi s(\varphi) = \eta h_\varphi^p \text{ for some } \eta \in L^2(M)s(\varphi)\}. \quad (3.5)$$

We prove that the domain $\mathcal{D}(\Delta_{\psi,\varphi}^p)$ coincides with $\mathcal{D}_p$ for any $p \geq 0$. The next lemma was shown in [49, Lemma A.3].

Lemma 3.4 ([49]) *For every $\psi, \varphi \in M_*^+$ and $0 \leq p \leq 1/2$,*

$$\mathcal{D}(\Delta_{\psi,\varphi}^p) = \mathcal{D}_p(\psi, \varphi).$$

Moreover, if $\xi, \eta \in L^2(M)$ are given as in (3.5)*, then*

$$\Delta_{\psi,\varphi}^p \xi = \Delta_{\psi,\varphi}^p(\xi s(\varphi)) = \eta. \quad (3.6)$$

The next extension to the whole $p \geq 0$ is due to Jenčová [61].

Lemma 3.5 ([61]) *The results of Lemma 3.4 hold for all $p \geq 0$.*

Proof Define

$$\mathcal{D}_0 := \{\xi \in L^2(M) : t \in \mathbb{R} \mapsto \Delta_{\psi,\varphi}^{it}\xi \in L^2(M) \text{ extends to an entire function}\}.$$

By a familiar regularization technique with Gaussian kernels (used also in the last part of the proof) it is seen that $\mathcal{D}_0$ is dense in $L^2(M)$ and is a core of $\Delta_{\psi,\varphi}^p$ for any $p \geq 0$. Let $\widetilde{N}$ be the space of τ-measurable operators affiliated with $N := M \rtimes_\sigma \mathbb{R}$, the crossed product algebra (see the first paragraph of Appendix A.6). Let $\xi \in \mathcal{D}_0$. By [110, Chap. II, Lemma 18 and Proposition 26] we see that the $\widetilde{N}$-valued functions $z \mapsto h_\psi^z \xi s(\varphi)$ and $z \mapsto (\Delta_{\psi,\varphi}^z \xi)h_\varphi^z$ are analytic in Re $z > 0$. By Lemma 3.4 these functions coincide for $z = p \in (0, 1/2]$, so they must be equal for all z with Re $z > 0$. This implies that $\mathcal{D}_0 \subset \mathcal{D}_p(\psi, \varphi)$ for all $p > 0$ and (3.6) holds for all $\xi \in \mathcal{D}_0$ and $p > 0$.

Now let $p > 1/2$ and let T_p denote the operator with domain $\mathcal{D}_p(\psi, \varphi)$ defined by $T_p\xi := \eta$, which is clearly a densely-defined linear operator on $L^2(M)$. By the previous paragraph, T_p coincides with $\Delta_{\psi,\varphi}^p$ on $\mathcal{D}_0$. We will show that T_p is a closed operator with a core $\mathcal{D}_0$; then the lemma follows.

So let $\{\xi_n\}$ be a sequence in $\mathcal{D}_p(\psi, \varphi)$ such that $\xi_n \to \xi$ and $T_p\xi_n \to \eta$ in $L^2(M)$. Then by [110, Chap. II, Proposition 26], $\xi_n \to \xi$ and $T_p\xi_n \to \eta$ in $\widetilde{N}$. Since $\widetilde{N}$ is a complete Hausdorff topological $*$-algebra (see [110, Chap. I, Theorem 28]), one has

$$h_\psi^p \xi = \lim_n h_\psi^p \xi_n = \lim_n T_p\xi_n h_\varphi^p = \eta h_\varphi^p,$$

so that $\xi \in \mathcal{D}_p(\psi, \varphi)$ and $T_p\xi = \eta$. Hence T_p is closed. To show that $\mathcal{D}_0$ is a core of T_p, let $\xi \in \mathcal{D}_p(\psi, \varphi)$ and $\eta = T_p\xi$. Set

$$\xi_n := \left(\frac{n}{\pi}\right)^{1/2} \int_{-\infty}^{\infty} e^{-nt^2} \Delta_{\psi,\varphi}^{it} \xi \, dt + \xi(1 - s(\varphi)),$$

$$\eta_n := \left(\frac{n}{\pi}\right)^{1/2} \int_{-\infty}^{\infty} e^{-nt^2} \Delta_{\psi,\varphi}^{it} \eta \, dt,$$

then $\xi_n \in \mathcal{D}_0$ and $\xi_n \to \xi$, $\eta_n \to \eta$ in $L^2(M)$ (see [69, Sec. 10]). Using the formula (A.18) of Appendix A.6 one can easily verify that

$$h_\psi^p \Delta_{\psi,\varphi}^{it} \xi = h_\psi^{it} h_\psi^p \xi h_\varphi^{-it} = h_\psi^{it} \eta h_\varphi^p h_\varphi^{-it} = (\Delta_{\psi,\varphi}^{it} \eta) h_\varphi^p, \qquad t \in \mathbb{R}.$$

From [110, Chap. I, Theorem 28] again we have

$$h_\psi^p \xi_n s(\varphi) = \left(\frac{n}{\pi}\right)^{1/2} \int_{-\infty}^{\infty} e^{-nt^2} h_\psi^p \Delta_{\psi,\varphi}^{it} \xi \, dt$$

$$= \left(\frac{n}{\pi}\right)^{1/2} \int_{-\infty}^{\infty} e^{-nt^2} (\Delta_{\psi,\varphi}^{it} \eta) h_\varphi^p \, dt = \eta_n h_\varphi^p$$

so that $T_p \xi_n = \eta_n \to \eta$, as desired. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

The following theorem provides an explicit description of Q_α (hence the Rényi divergence $D_\alpha(\psi\|\varphi)$) in terms of $h_\psi, h_\varphi \in L^1(M)_+$.

Theorem 3.6 *Let $\psi, \varphi \in M_*^+$.*

(1) When $0 \le \alpha < 1$,

$$Q_\alpha(\psi\|\varphi) = \mathrm{tr}\,(h_\psi^\alpha h_\varphi^{1-\alpha}). \tag{3.7}$$

(2) When $s(\psi) \le s(\varphi)$ and $\alpha > 1$, the following conditions are equivalent:

(i) $h_\varphi^{1/2} \in \mathcal{D}(\Delta_{\psi,\varphi}^{\alpha/2})$;

(ii) $h_\psi^{1/2} \in \mathcal{D}(\Delta_{\psi,\varphi}^{(\alpha-1)/2})$;

(iii) *there exists an $\eta \in L^2(M)s(\varphi)$ such that $h_\psi^{\alpha/2} = \eta h_\varphi^{(\alpha-1)/2}$.*

If the above conditions hold, then η in (iii) *is unique and $Q_\alpha(\psi\|\varphi) = \|\eta\|_2^2$.*

Proof (1) For $0 < \alpha < 1$, since $h_\psi^{\alpha/2} h_\varphi^{1/2} = (h_\psi^{\alpha/2} h_\varphi^{(1-\alpha)/2}) h_\varphi^{\alpha/2}$, we have by (3.6)

$$Q_\alpha(\psi\|\varphi) = \|\Delta_{\psi,\varphi}^{\alpha/2} h_\varphi^{1/2}\|^2 = \|h_\psi^{\alpha/2} h_\varphi^{(1-\alpha)/2}\|^2$$

$$= \mathrm{tr}\,(h_\varphi^{(1-\alpha)/2} h_\psi^\alpha h_\varphi^{(1-\alpha)/2}) = \mathrm{tr}\,(h_\psi^\alpha h_\varphi^{1-\alpha}).$$

(2) Assume that $s(\psi) \le s(\varphi)$ and $\alpha > 1$. Since $h_\varphi^{1/2} \in \mathcal{D}(\Delta_{\psi,\varphi}^{1/2})$ and $\Delta_{\psi,\varphi}^{1/2} h_\varphi^{1/2} = h_\psi^{1/2} s(\varphi) = h_\psi^{1/2}$, it follows that (i) $\Longleftrightarrow$ (ii) and in this case $\Delta_{\psi,\varphi}^{\alpha/2} h_\varphi^{1/2} = \Delta_{\psi,\varphi}^{(\alpha-1)/2} h_\psi^{1/2}$ (see [104, Theorem 9.20] for details). Hence Lemma 3.5 with $p = (\alpha-1)/2$ implies that (ii) $\Longleftrightarrow$ (iii) and in this case $Q_\alpha(\psi\|\varphi) = \|\eta\|_2^2$. The uniqueness of η in (iii) is obvious. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 3.7 In the above (2), if $h_\psi^{\alpha/2} = \eta h_\varphi^{(\alpha-1)/2}$ with $\eta \in L^2(M)s(\varphi)$, then we may write $\eta = h_\psi^{\alpha/2} h_\varphi^{(1-\alpha)/2}$ in a formal sense, so that

$$Q_\alpha(\psi\|\varphi) = \|h_\psi^{\alpha/2} h_\varphi^{(1-\alpha)/2}\|^2 = \mathrm{tr}\,(h_\psi^{\alpha/2} h_\varphi^{1-\alpha} h_\psi^{\alpha/2}) = \mathrm{tr}\,(h_\psi^\alpha h_\varphi^{1-\alpha}),$$

which is the same expression as in (1). The expression is in the same form as (3.4) in the finite-dimensional case when the usual trace is replaced with tr and the density operators are with h_ψ, h_φ.

Corollary 3.8 *When* $0 \le \alpha < 1$*, the map* $(\psi, \varphi) \in (M_*^+\backslash\{0\})\times M_*^+ \mapsto D_\alpha(\psi\|\varphi) \in \mathbb{R}$ *is jointly continuous in the norm topology.*

Proof By (3.7) it suffices to show that $(\psi, \varphi) \in M_*^+ \times M_*^+ \mapsto h_\psi^\alpha h_\varphi^{1-\alpha} \in L^1(M)$ is continuous. Let $\psi_n, \psi, \varphi_n, \varphi \in M_*^+$ be such that $\|\psi_n - \psi\|_1 \to 0$ and $\|\varphi_n - \varphi\|_1 \to 0$. Then

$$
\begin{aligned}
\|h_{\psi_n}^\alpha h_{\varphi_n}^{1-\alpha} - h_\psi^\alpha d_\varphi^{1-\alpha}\|_1 &\le \|(h_{\psi_n}^\alpha - h_\psi^\alpha)h_{\varphi_n}^{1-\alpha}\|_1 + \|h_\psi^\alpha(h_{\varphi_n}^{1-\alpha} - h_\varphi^{1-\alpha})\|_1 \\
&\le \|h_{\psi_n}^\alpha - h_\psi^\alpha\|_{1/\alpha} \|h_{\varphi_n}^{1-\alpha}\|_{1/(1-\alpha)} \\
&\qquad + \|h_\psi^\alpha\|_{1/\alpha} \|h_{\varphi_n}^{1-\alpha} - h_\varphi^{1-\alpha}\|_{1/(1-\alpha)} \\
&\le \|h_{\psi_n} - h_\varphi\|_1^\alpha \|h_{\varphi_n}\|_1^{1-\alpha} + \|h_\psi\|_1^\alpha \|h_{\varphi_n} - h_\varphi\|_1^{1-\alpha} \\
&\longrightarrow 0,
\end{aligned}
$$

where the second inequality is Hölder's inequality for Haagerup's L^p-norms (see Theorem A.38) and the third follows from Kosaki's generalized Powers-Størmer inequality given in the appendix in [54]. $\qquad\square$

3.3 Sandwiched Rényi Divergences

When $\mathcal{H}$ is finite-dimensional, for every $\psi, \varphi \in B(\mathcal{H})_+^*$ ($\psi \ne 0$) with the density operators D_ψ, D_φ and every $\alpha \in (0, \infty) \setminus \{1\}$, the *$\alpha$-sandwiched Rényi divergence* introduced in [85, 116] is

$$\widetilde{D}_\alpha(\psi\|\varphi) := \frac{1}{\alpha - 1} \log \frac{\widetilde{Q}_\alpha(\psi\|\varphi)}{\psi(1)},$$

where

$$\widetilde{Q}_\alpha(\psi\|\varphi) := \begin{cases} \mathrm{Tr}\left(D_\varphi^{\frac{1-\alpha}{2\alpha}} D_\psi D_\varphi^{\frac{1-\alpha}{2\alpha}}\right)^\alpha & \text{if } 0 < \alpha < 1 \text{ or } s(\psi) \le s(\varphi), \\ +\infty & \text{otherwise.} \end{cases} \tag{3.8}$$

These sandwiched variants of Rényi divergences have been in active consideration with applications, e.g., [13, 34, 83, 84]. A more general notion of the *α-z-Rényi relative entropies* was first introduced in [57, Sec. 4.3.3] and further discussed in

[12]. In this section we will survey the α-sandwiched Rényi divergences in von Neumann algebras, recently developed in [17, 59, 60].

Definition 3.9 ([59]) Let $\psi, \varphi \in M_*^+$ with $\psi \neq 0$, and $1 < \alpha < \infty$. The α-*sandwiched Rényi divergence* due to Jenčová is

$$\widetilde{D}_\alpha^{(\mathrm{J})}(\psi\|\varphi) := \frac{1}{\alpha - 1} \log \frac{\widetilde{Q}_\alpha^{(\mathrm{J})}(\psi\|\varphi)}{\psi(1)}.$$

where

$$\widetilde{Q}_\alpha^{(\mathrm{J})}(\psi\|\varphi) := \begin{cases} \|h_\psi\|_{\alpha,\varphi}^\alpha & \text{if } s(\psi) \leq s(\varphi) \text{ and } h_\psi \in L^\alpha(M; \varphi), \\ +\infty & \text{otherwise.} \end{cases}$$

Here, $L^\alpha(M; \varphi)$ is *Kosaki's L^α-space* in the symmetric case and $\|\cdot\|_{\alpha,\varphi}$ is the norm on $L^\alpha(M; \varphi)$ (more precisely, $\|\cdot\|_{\alpha,\varphi,\theta}$ with $\theta = 1/2$), see Appendix A.8.

An alternative definition of the sandwiched divergence in von Neumann algebra given in [17] is based on *Araki-Masuda's L^p-norm* [10]. So here we briefly recall the definition in [10]. Let $\psi, \varphi \in M_*^+$. For $2 \leq p \leq \infty$ Araki-Masuda's L^p-norm of the vector representative $h_\varphi^{1/2}$ with respect to φ is

$$\|h_\psi^{1/2}\|_{p,\varphi} := \begin{cases} \sup\left\{\|\Delta_{\omega,\varphi}^{\frac{1}{2} - \frac{1}{p}} h_\psi^{1/2}\| : \omega \in M_*^+, \omega(1) = 1\right\} & \text{if } s(\psi) \leq s(\varphi), \\ +\infty & \text{otherwise.} \end{cases} \tag{3.9}$$

For $1 \leq p < 2$, Araki-Masuda's L^p-norm of $h_\psi^{1/2}$ is

$$\|h_\psi^{1/2}\|_{p,\varphi} := \inf\left\{\|\Delta_{\omega,\varphi}^{\frac{1}{2} - \frac{1}{p}} h_\psi^{1/2}\| : \omega \in M_*^+, \omega(1) = 1, s(\omega) \geq s(\psi)\right\}, \tag{3.10}$$

where $\|\Delta_{\omega,\psi}^{\frac{1}{2} - \frac{1}{p}} h_\psi^{1/2}\|$ means $+\infty$ if $h_\psi^{1/2}$ is not in $\mathcal{D}(\Delta_{\omega,\psi}^{\frac{1}{2} - \frac{1}{p}})$. (Note that the above p-norm of vector representatives can defined under any $*$-representation of M on any Hilbert space $\mathcal{H}$, while we here use the standard representation on Haagerup's $L^2(M)$.)

Definition 3.10 ([17]) For every $\psi, \varphi \in M_*^+$ with $\psi \neq 0$ and $\alpha \in [1/2, \infty) \setminus \{1\}$, the α-*sandwiched Rényi divergence* due to Berta-Scholz-Tomamichel (called the *Araki-Masuda divergence* in [17]) is

$$\widetilde{D}_\alpha^{(\mathrm{BST})}(\psi\|\varphi) := \frac{1}{\alpha - 1} \log \widetilde{Q}_\alpha^{(\mathrm{BST})}(\psi\|\varphi), \quad \text{where} \quad \widetilde{Q}_\alpha^{(\mathrm{BST})}(\psi\|\varphi) := \|h_\psi^{1/2}\|_{2\alpha,\varphi}^{2\alpha}.$$

The next theorem is a main result of this section, which was proved by Jenčová. The proof below is a bit different from that in [60].

Theorem 3.11 ([60]) *Let $\psi, \varphi \in M_*^+$ with $\psi \neq 0$.*

(1) *For every $\alpha \in (1, \infty)$ we have*

$$\widetilde{Q}_\alpha^{(\mathrm{BST})}(\psi \| \varphi) = \widetilde{Q}_\alpha^{(\mathrm{J})}(\psi \| \varphi). \tag{3.11}$$

Hence

$$\widetilde{D}_\alpha^{(\mathrm{BST})}(\psi \| \varphi) = \widetilde{D}_\alpha^{(\mathrm{J})}(\psi \| \varphi).$$

(2) *For every $\alpha \in [1/2, 1)$ we have*

$$\widetilde{Q}_\alpha^{(\mathrm{BST})}(\psi \| \varphi) = \mathrm{tr} \left(h_\varphi^{\frac{1-\alpha}{2\alpha}} h_\psi h_\varphi^{\frac{1-\alpha}{2\alpha}} \right)^\alpha. \tag{3.12}$$

Hence

$$\widetilde{D}_\alpha^{(\mathrm{BST})}(\psi \| \varphi) = \frac{1}{\alpha - 1} \log \frac{\mathrm{tr}\, (h_\varphi^{\frac{1-\alpha}{2\alpha}} h_\psi h_\varphi^{\frac{1-\alpha}{2\alpha}})^\alpha}{\psi(1)}.$$

To prove the theorem, we first give lemmas.

Lemma 3.12 *Assume that $1 < \alpha < \infty$.*

(1) $\widetilde{D}_\alpha^{(\mathrm{J})}(\psi \| \varphi)$ *is jointly lower semicontinuous in $\psi, \varphi \in M_*^+$.*
(2) $\widetilde{Q}_\alpha^{(\mathrm{J})}(\psi \| \varphi)$ *is jointly convex in $\psi, \varphi \in M_*^+$.*
(3) $\widetilde{Q}_\alpha^{(\mathrm{BST})}(\psi \| \varphi)$ *is jointly convex in $\psi, \varphi \in M_*^+$.*

Proof See [59] for the proofs of (1) and (2).

(3) Consider a unital $*$-homomorphism (hence CP) $\gamma : M \to M \oplus M$ given by $\gamma(x) := x \oplus x$. For $\psi_i, \varphi_i \in M_*^+$ ($i = 1, 2$) and $\lambda \in (0, 1)$ set $\psi := \lambda\varphi_1 \oplus (1-\lambda)\psi_2$ and $\varphi := \lambda\varphi_1 \oplus (1-\lambda)\varphi_2$. Since $\psi \circ \gamma = \lambda\psi_1 + (1-\lambda)\psi_2$ and $\varphi \circ \gamma = \lambda\varphi_1 + (1-\lambda)\varphi_2$, from the monotonicity property of $\widetilde{D}_\alpha^{(\mathrm{BST})}$ (equivalent to that of $\widetilde{Q}_\alpha^{(\mathrm{BST})}$) given in [17, Theorem 14] it follows that

$$\widetilde{Q}_\alpha^{(\mathrm{BST})}(\lambda\psi_1 + (1-\lambda)\psi_2 \| \lambda\varphi_1 + (1-\lambda)\varphi_2) \le \widetilde{Q}_\alpha^{(\mathrm{BST})}(\psi \| \varphi).$$

So it suffices to prove that

$$\widetilde{Q}_\alpha^{(\mathrm{BST})}(\psi \| \varphi) = \lambda \widetilde{Q}_\alpha^{(\mathrm{BST})}(\psi_1 \| \varphi_1) + (1-\lambda)\widetilde{Q}_\alpha^{(\mathrm{BST})}(\psi_2 \| \varphi_2). \tag{3.13}$$

Note that $L^2(M \oplus M) = L^2(M) \oplus L^2(M)$ and

$$h_\psi^{1/2} = \lambda^{1/2} h_{\psi_1}^{1/2} \oplus (1-\lambda)^{1/2} h_{\psi_2}^{1/2}, \quad h_\varphi^{1/2} = \lambda^{1/2} h_{\varphi_1}^{1/2} \oplus (1-\lambda)^{1/2} h_{\varphi_2}^{1/2}.$$

Let $p := 2\alpha \in (2, \infty)$. For every $\omega \in (M \oplus M)_*^+$ with $\omega(1) = 1$, we can write $\omega = t\omega_1 \oplus (1-t)\omega_2$ where $\omega_i \in M_*^+$, $\omega_i(1) = 1$ and $0 \le t \le 1$. Since

$$\Delta_{\omega,\varphi} = \Delta_{t\omega_1,\lambda\varphi_1} \oplus \Delta_{(1-t)\omega_2,(1-\lambda)\varphi_2} = \frac{t}{\lambda}\Delta_{\omega_1,\varphi_1} \oplus \frac{1-t}{1-\lambda}\Delta_{\omega_2,\varphi_2},$$

we have

$$\|\Delta_{\omega,\varphi}^{\frac{1}{2}-\frac{1}{p}} h_\psi^{1/2}\| = \left\{ t^{1-\frac{2}{p}}\lambda^{\frac{2}{p}}\|\Delta_{\omega_1,\varphi_1}^{\frac{1}{2}-\frac{1}{p}} h_{\psi_1}^{1/2}\|^2 + (1-t)^{1-\frac{2}{p}}(1-\lambda)^{\frac{2}{p}}\|\Delta_{\omega_2,\varphi_2}^{\frac{1}{2}-\frac{1}{p}} h_{\psi_2}^{1/2}\|^2 \right\}^{1/2}.$$

Therefore,

$$\|h_\psi^{1/2}\|_{p,\varphi} = \sup_{t,\omega_1,\varphi_2} \left\{ t^{1-\frac{2}{p}} \lambda^{\frac{2}{p}} \|\Delta_{\omega_1,\varphi_1}^{\frac{1}{2}-\frac{1}{p}} h_{\psi_1}^{1/2}\|^2 + (1-t)^{1-\frac{2}{p}}(1-\lambda)^{\frac{2}{p}} \|\Delta_{\omega_2,\varphi_2}^{\frac{1}{2}-\frac{1}{p}} h_{\psi_2}^{1/2}\|^2 \right\}^{1/2}$$

$$\text{(the sup is over } 0 \le t \le 1, \omega_1, \omega_2 \in M_*^+, \omega_1(1) = \omega_2(1) = 1)$$

$$= \sup_{0 \le t \le 1} \left\{ t^{1-\frac{2}{p}} \lambda^{\frac{2}{p}} \|h_{\psi_1}^{1/2}\|_{p,\varphi_1}^2 + (1-t)^{1-\frac{2}{p}}(1-\lambda)^{\frac{2}{p}} \|h_{\psi_2}^{1/2}\|_{p,\varphi_2}^2 \right\}^{1/2}$$

$$= \left\{ \lambda \|h_{\psi_1}^{1\cdot2}\|_{p,\varphi_1}^p + (1-\lambda)\|h_{\psi_2}^{1/2}\|_{p,\varphi_2}^p \right\}^{1/p}$$

so that $\|h_\psi^{1/2}\|_{p,\varphi}^p = \lambda\|h_{\psi_1}^{1/2}\|_{p,\varphi_1}^p + (1-\lambda)\|h_{\psi_2}^{1/2}\|_{p,\varphi_2}^p$, that is (3.18). $\qquad\square$

Lemma 3.13 *Let $\alpha \in (1,\infty)$ and $\psi,\varphi \in M_*^+$. Then*

(1) $\widetilde{D}_\alpha^{(J)}(\psi\|\varphi) = \lim_{\varepsilon\searrow0} \widetilde{D}_\alpha^{(J)}(\psi\|\varphi + \varepsilon\psi)$, and
(2) $\widetilde{D}_\alpha^{(BST)}(\psi\|\varphi) = \lim_{\varepsilon\searrow0} \widetilde{D}_\alpha^{(BST)}(\psi\|\varphi + \varepsilon\psi)$ if $s(\psi) \le s(\varphi)$.

Proof (1) It is clear that $\widetilde{Q}_\alpha^{(J)}(\psi\|\psi) = 1$. Hence by Lemma 3.12 (2),

$$\widetilde{Q}_\alpha^{(J)}(\psi\|(\varphi + \varepsilon\psi)/(1+\varepsilon)) \le \frac{\widetilde{Q}_\alpha^{(J)}(\psi\|\varphi) + \varepsilon}{1+\varepsilon}.$$

Moreover, as easily seen and noted in [59, (14)],

$$\widetilde{D}_\alpha^{(J)}(\psi\|(\varphi + \varepsilon\psi)/(1+\varepsilon)) = \widetilde{D}_\alpha^{(J)}(\psi\|\varphi + \varepsilon\psi) + \log(1+\varepsilon).$$

Therefore,

$$\limsup_{\varepsilon\searrow0} \widetilde{D}_\alpha^{(J)}(\psi\|\varphi + \varepsilon\psi) = \limsup_{\varepsilon\searrow0} \widetilde{D}_\alpha^{(J)}(\psi\|(\varphi + \varepsilon\psi)/(1+\varepsilon)) \le \widetilde{D}_\alpha^{(J)}(\psi\|\varphi).$$

On the other hand, by Lemma 3.12 (1), $\widetilde{D}_\alpha^{(J)}(\psi\|\varphi) \le \liminf_{\varepsilon\searrow0} \widetilde{D}_\alpha^{(J)}(\psi\|\varphi + \varepsilon\psi)$.

(2) The scaling property of $\widetilde{D}_\alpha^{(BST)}$ is the same as [59, (14)] for $\widetilde{D}_\alpha^{(J)}$. Hence as in the above proof of (1), by Lemma 3.12 (3) we have

$$\limsup_{\varepsilon\searrow0} \widetilde{D}_\alpha^{(BST)}(\psi\|\varphi + \varepsilon\psi) \le \widetilde{D}_\alpha^{(BST)}(\psi\|\varphi).$$

So it suffices to prove that

$$\widetilde{D}_\alpha^{(BST)}(\psi\|\varphi) \le \liminf_{\varepsilon\searrow0} \widetilde{D}_\alpha^{(BST)}(\psi\|\varphi + \varepsilon\psi),$$

that is, for $p := 2\alpha \in (2,\infty)$,

$$\|h_\psi^{1/2}\|_{p,\varphi} \le \liminf_{\varepsilon\searrow0} \|h_\psi^{1/2}\|_{p,\varphi+\varepsilon\psi}. \tag{3.14}$$

Let $\varepsilon_n \searrow 0$ be arbitrary and let $\omega \in M_*^+$. By [9, Lemma 4.1] we have

$$(I + \Delta_{\omega,\varphi+\varepsilon_n\psi}^{1/2})^{-1}Js(\varphi)J \longrightarrow (I + \Delta_{\omega,\varphi}^{1/2})^{-1}Js(\varphi)J \quad \text{strongly.}$$

Since $s(\psi) \leq s(\varphi)$, both of $\Delta_{\omega,\varphi+\varepsilon_n\psi}$ and $\Delta_{\omega,\varphi}$ have the support projection $s(\omega)Js(\varphi)J \ (\leq Js(\varphi)J)$. Hence the above strong convergence means that

$$(1 + \Delta_{\omega,\varphi+\varepsilon_n\psi}^{1/2})^{-1} \longrightarrow (1 + \Delta_{\omega,\varphi}^{1/2})^{-1} \quad \text{strongly.}$$

Moreover, since $s(\varphi + \varepsilon_1\psi) = s(\varphi)$, we find by Proposition B.4 of Appendix B that

$$(1 + \Delta_{\omega,\varphi+\varepsilon_1\psi})^{-1} \geq (1 + \Delta_{\omega,\varphi+\varepsilon_2\psi})^{-1} \geq \cdots \geq (1 + \Delta_{\omega,\varphi})^{-1}.$$

Therefore, by Lemma B.7 we have

$$\|\Delta_{\omega,\varphi}^{\frac{1}{2}-\frac{1}{p}}h_\psi^{1/2}\|^2 = \lim_{n\to\infty}\|\Delta_{\omega,\varphi+\varepsilon_n\psi}^{\frac{1}{2}-\frac{1}{p}}h_\psi^{1/2}\|^2,$$

which implies (3.14) in view of (3.9). $\square$

Proof (Theorem 3.11) (1) If $s(\psi) \not\leq s(\varphi)$, then both sides of (3.11) are $+\infty$ by Definitions 3.9 and 3.10. Hence by Lemma 3.13 we may prove equality (3.11) in the case where $\psi \leq \lambda\varphi$ for some $\lambda > 0$. Then by Lemma A.21, $h_\psi^{1/2} = Ah_\varphi^{1/2}$ for some $A \in s(\varphi)Ms(\varphi)$. Let $p := 2\alpha \in (2, \infty)$ and $\omega \in M_*^+$ with $\omega(1) = 1$. From (A.19) of Appendix A.6 one has

$$\Delta_{\omega,\varphi}^{\frac{1}{2}-\frac{1}{p}}h_\psi^{1/2} = \Delta_{\omega,\varphi}^{\frac{1}{2}-\frac{1}{p}}Ah_\varphi^{1/2} = h_\omega^{\frac{1}{2}-\frac{1}{p}}Ah_\varphi^{\frac{1}{p}}$$

so that

$$\|\Delta_{\omega,\varphi}^{\frac{1}{2}-\frac{1}{p}}h_\psi^{1/2}\|^2 = \text{tr}\left(h_\omega^{\frac{1}{2}-\frac{1}{p}}Ah_\varphi^{\frac{2}{p}}A^*h_\omega^{\frac{1}{2}-\frac{1}{p}}\right) = \text{tr}\,(h_\omega^{(p-2)/p}Ah_\varphi^{2/p}A^*). \tag{3.15}$$

Apply Hölder's inequality for Haagerup's L^p-norms to $h_\omega^{(p-2)/p} \in L^{p/(p-2)}(M)$ and $Ah_\varphi^{2/p}A^* \in L^{p/2}(M)$ to obtain

$$|\text{tr}\,(h_\omega^{(p-2)/p}Ah_\varphi^{2/p}A^*)| \leq \|Ah_\varphi^{2/p}A^*\|_{p/2} \tag{3.16}$$

thanks to $\|h_\omega^{(p-2)/p}\|_{p/(p-2)} = (\text{tr}\,h_\omega)^{(p-2)/p} = 1$. On the other hand, set $h := (Ah_\varphi^{2/p}A^*)^{p/2}/\|Ah_\varphi^{2/p}A^*\|_{p/2}^{p/2}$. Then $h \in L^1(M)_+$ and $\text{tr}\,(h) = 1$, so that $h = h_{\omega_0}$ for some $\omega_0 \in M_*^+$ with $\omega_0(1) = 1$. We then have

$$\text{tr}\,(h_{\omega_0}^{(p-2)/p}Ah_\varphi^{2/p}A^*) = \frac{\text{tr}\,((Ah_\varphi^{2/p}A^*)^{p/2})}{\text{tr}\,((Ah_\varphi^{2/p}A^*)^{p/2})^{(p-2)/p}} = \|Ah_\varphi^{2/p}A^*\|_{p/2}. \tag{3.17}$$

Combining (3.15)–(3.17) yields $\|h_\psi^{1/2}\|_{p,\varphi}^2 = \|Ah_\varphi^{2/p}A^*\|_{p/2}$ so that

$$\|h_\psi^{1/2}\|_{2\alpha,\varphi}^{2\alpha} = \|Ah_\varphi^{1/\alpha}A^*\|_\alpha^\alpha = \|h_\varphi^{\frac{1}{2\alpha}}A^*Ah_\varphi^{\frac{1}{2\alpha}}\|_\alpha^\alpha. \tag{3.18}$$

Now define

$$h_0 := (h_\varphi^{\frac{1}{2\alpha}} A^* A h_\varphi^{\frac{1}{2\alpha}})^\alpha \in L^1(M),$$

for which we have, with $1/\alpha + 1/\beta = 1$,

$$h_\varphi^{\frac{1}{2\beta}} h_0^{1/\alpha} h_\varphi^{\frac{1}{2\beta}} = h_\varphi^{\frac{1}{2\alpha}+\frac{1}{2\beta}} A^* A h_\varphi^{\frac{1}{2\alpha}+\frac{1}{2\beta}} = h_\varphi^{1/2} A^* A h_\varphi^{1/2} = h_\psi.$$

This implies that $h_\psi \in L^\alpha(M; \varphi)$ by Theorem A.61 (in the $\theta = 1/2$ case) and

$$\|h_\psi\|_{\alpha,\varphi} = (\mathrm{tr}\, h_0)^{1/\alpha} = \|h_\varphi^{\frac{1}{2\alpha}} A^* A h_\varphi^{\frac{1}{2\alpha}}\|_\alpha. \tag{3.19}$$

By (3.18) and (3.19) we thus obtain $\|h_\psi^{1/2}\|_{2\alpha,\varphi}^{2\alpha} = \|h_\psi\|_{\alpha,\varphi}^\alpha$, which yields (3.11) by Definitions 3.9 and 3.10.

(2) Assume that $1/2 \leq \alpha < 1$. First, note that the RHS of (3.12) is well defined since $(h_\varphi^{\frac{1-\alpha}{2\alpha}} h_\psi h_\varphi^{\frac{1-\alpha}{2\alpha}})^\alpha$ is in $L^1(M)$. Since $\Delta_{\omega,\varphi}^{-1} = J\Delta_{\varphi,\omega} J$ for every $\omega, \varphi \in M_*^+$ by Proposition A.23 (3), the definition of $\|h_\psi^{1/2}\|_{2\alpha,\varphi}$ in (3.10) is rewritten as

$$\|h_\psi^{1/2}\|_{2\alpha,\varphi} = \inf\{\|\Delta_{\varphi,\omega}^{\frac{1}{2\alpha}-\frac{1}{2}} h_\psi^{1/2}\| : \omega \in M_*^+, \omega(1) = 1, s(\omega) \geq s(\psi)\}.$$

Using Lemma 3.4 we find that

$$\|h_\psi^{1/2}\|_{2\alpha,\varphi}^2 = \inf\{\|\eta\|_2^2 : \omega \in M_*^+, \omega(1) = 1, s(\omega) \geq s(\psi),$$
$$\eta \in L^2(M), h_\varphi^{\frac{1-\alpha}{2\alpha}} h_\psi^{1/2} = \eta h_\omega^{\frac{1-\alpha}{2\alpha}}\}, \tag{3.20}$$

where we note that $h_\psi^{1/2} s(\omega) = h_\psi^{1/2}$ (thanks to $s(\omega) \geq s(\psi)$), $\eta h_\omega^{\frac{1-\alpha}{2\alpha}} = \eta s(\omega) h_\omega^{\frac{1-\alpha}{2\alpha}}$ and $\|\eta s(\omega)\|_2^2 \leq \|\eta\|_2^2$. Now let $p := \alpha$ and $q := \frac{\alpha}{1-\alpha}$ so that $1/p - 1/q = 1$. By letting $b := h_\omega^{\frac{1-\alpha}{\alpha}} \in L^q(M)_+$ expression (3.20) can be rewritten as

$$\|h_\psi^{1/2}\|_{2\alpha,\varphi}^2 = \inf\{\|\eta\|_2^2 : b \in L^q(M)_+, \|b\|_q = 1, s(b) \geq s(\psi),$$
$$\eta \in L^2(M), h_\varphi^{\frac{1-\alpha}{2\alpha}} h_\psi^{1/2} = \eta b^{1/2}\}.$$

Furthermore, let $a := |h_\varphi^{\frac{1-\alpha}{2\alpha}} h_\psi^{1/2}|^2$ and $h_\varphi^{\frac{1-\alpha}{2\alpha}} h_\psi^{1/2} = v a^{1/2}$ be the polar decomposition where $v \in M$ is a partial isometry with $v^* v = s(a) \leq s(\psi)$. Since $v a^{1/2} = \eta b^{1/2}$ $\iff$ $a^{1/2} = v^* \eta b^{1/2}$ $\iff$ $v a^{1/2} = v v^* \eta b^{1/2}$ and $\|\eta\|_2 \geq \|v^* \eta\|_2 = \|v v^* \eta\|_2$, it follows that

$$\|h_\psi^{1/2}\|_{2\alpha,\varphi}^2 = \inf\{\|\eta\|_2^2 : b \in L^q(M)_+, \|b\|_q = 1, s(b) \geq s(\psi),$$
$$\eta \in L^2(M), a^{1/2} = \eta b^{1/2}\}. \tag{3.21}$$

Now we show that the RHS of (3.21) is equal to $\|a\|_p$. For this we may assume that $a \neq 0$ (the result is immediate when $a = 0$). If b and η are as in the above RHS, then

Hölder's inequality (Theorem A.38) implies that

$$\|a\|_p^{1/2} = \|a^{1/2}\|_{2p} = \|\eta b^{1/2}\|_{2p} \leq \|\eta\|_2 \|b^{1/2}\|_{2q} = \|\eta\|_2 \|b\|_q^{1/2} = \|\eta\|_2.$$

Furthermore, set $\eta := \|a\|_p^{p/2q} a^{p/2}$ and $b = \|a\|_p^{-p/q} a^{p/q}$; then it is immediate to check that $\eta \in L^2(M)$, $b \in L^q(M)_+$, $\|b\|_q = 1$, $a^{1/2} = \eta b^{1/2}$ and $\|\eta\|_2^2 = \|a\|_p$. Note that $s(\eta) = s(b) = s(a) \leq s(\psi)$. Choose a $b_0 \in L^q(M)_+$ with $s(b_0) = s(\psi) - s(a)$ and for $\varepsilon > 0$ set $\eta_\varepsilon := \|b + \varepsilon b_0\|_q^{1/2}\eta$ and $b_\varepsilon := \|b + \varepsilon b_0\|_q^{-1}(b + \varepsilon b_0)$. Then $s(b_\varepsilon) = s(\psi)$, $a^{1/2} = \eta_\varepsilon b_\varepsilon^{1/2}$ and $\|\eta_\varepsilon\|_2^2 \to \|\eta\|_2^2 = \|a\|_p$ as $\varepsilon \searrow 0$. Therefore, the RHS of (3.21) is $\|a\|_p$ so that

$$\|h_\psi^{1/2}\|_{2\alpha,\varphi}^2 = \|a\|_p = \left[\mathrm{tr} \left(h_\varphi^{\frac{1-\alpha}{2\alpha}} h_\psi h_\varphi^{\frac{1-\alpha}{2\alpha}} \right)^\alpha \right]^{1/\alpha},$$

which is (3.12) by Definition 3.10 when $1/2 \leq \alpha < 1$.   □

In view of Theorem 3.11 we simply denote the α-sandwiched Rényi divergence by

$$\widetilde{D}_\alpha(\psi\|\varphi) = \frac{1}{\alpha - 1} \log \frac{\widetilde{Q}_\alpha(\psi\|\varphi)}{\psi(1)} \tag{3.22}$$

for $\psi, \varphi \in M_*^+$ and $\alpha \in [1/2, \infty) \setminus \{1\}$, where

$$\widetilde{Q}_\alpha(\psi\|\varphi) = \begin{cases} \widetilde{Q}_\alpha^{(\mathrm{BST})}(\psi\|\varphi) = \widetilde{Q}_\alpha^{(\mathrm{J})}(\psi\|\varphi) & \text{if } 1 < \alpha < \infty, \\ \widetilde{Q}_\alpha^{(\mathrm{BST})}(\psi\|\varphi) = \mathrm{tr}\left(h_\varphi^{\frac{1-\alpha}{2\alpha}} h_\psi h_\varphi^{\frac{1-\alpha}{2\alpha}} \right)^\alpha & \text{if } 1/2 \leq \alpha < 1. \end{cases} \tag{3.23}$$

Remark 3.14 The expression for $1/2 \leq \alpha < 1$ in (3.23) has a complete resemblance to the matrix case in (3.8). For $1 < \alpha < \infty$, if $h_\psi \in L^\alpha(M;\varphi)$ (Kosaki's L^α-space), then there exists an $x \in (s(\varphi)L^\alpha(M)s(\varphi))_+$ such that $h_\psi = h_\varphi^{\frac{\alpha-1}{2\alpha}} x h_\varphi^{\frac{\alpha-1}{2\alpha}}$ and $\|h_\psi\|_{\alpha,\varphi} = \|x\|_\alpha$. Hence we may write

$$\widetilde{Q}_\alpha(\psi\|\varphi) = \mathrm{tr}\, x^\alpha = \mathrm{tr}\left(h_\varphi^{\frac{1-\alpha}{2\alpha}} h_\psi h_\varphi^{\frac{1-\alpha}{2\alpha}} \right)^\alpha$$

in a formal sense, which is the same expression as (3.8) again.

In particular, when $\alpha = 1/2$, note that

$$\widetilde{Q}_{1/2}(\psi\|\varphi) = F(\psi,\varphi), \qquad \widetilde{D}_{1/2}(\psi\|\varphi) = -2\log \frac{F(\psi,\varphi)}{\psi(1)}, \tag{3.24}$$

where $F(\psi,\varphi) := \mathrm{tr}\,(h_\varphi^{1/2} h_\psi h_\varphi^{1/2})^{1/2} = \|h_\psi^{1/2} h_\varphi^{1/2}\|_1$ is the so-called *fidelity* and will be discussed in Sec. 5.3. On the other hand, the Rényi divergence for $\alpha = 1/2$ is

$$D_{1/2}(\psi\|\varphi) = -2\log \frac{P(\psi,\varphi)}{\psi(1)}, \tag{3.25}$$

where $P(\psi, \varphi) := \mathrm{tr}\,(h_\varphi^{1/2} h_\varphi^{1/2}) = \langle h_\psi^{1/2}, h_\varphi^{1/2}\rangle$ is the so-called *transition probability*, introduced by Raggio [98] in the von Neumann algebra setting.

Properties of $\widetilde{D}_\alpha$ and $\widetilde{Q}_\alpha$ are summarized in the next theorem, which are more or less similar to those of D_α and Q_α in Theorem 3.2.

Theorem 3.15 ([17, 59]) *Let $\psi, \varphi \in M_*^+$ with $\psi \neq 0$.*

(1) *The function $\alpha \mapsto \widetilde{D}_\alpha(\psi\|\varphi)$ is monotone increasing on $[1/2, 1) \cup (1, \infty)$.*
(2) Limit values:

$$\lim_{\alpha \nearrow 1} \widetilde{D}_\alpha(\psi\|\varphi) = \frac{D(\psi\|\varphi)}{\psi(1)},$$

where $D(\psi\|\varphi)$ is the relative entropy. *Moreover, if $\widetilde{D}_\alpha(\psi\|\varphi) < +\infty$ for some $\alpha > 1$, then $\lim_{\alpha \searrow 1} \widetilde{D}_\alpha(\psi\|\varphi) = D(\psi\|\varphi)/\psi(1)$. On the other hand,*

$$\lim_{\alpha \to \infty} \widetilde{D}_\alpha(\psi\|\varphi) = D_{\max}(\psi\|\varphi),$$

where

$$D_{\max}(\psi\|\varphi) := \log \inf\{t > 0 : \psi \leq t\varphi\}$$

is the max-relative entropy *introduced in [28] (here $\inf \emptyset = +\infty$ as usual).*
(3) Joint lower semicontinuity: *The map $(\psi, \varphi) \in (M_*^+ \setminus \{0\}) \times M_*^+ \mapsto \widetilde{D}_\alpha(\psi\|\varphi)$ is jointly lower semicontinuous in the norm topology for every $\alpha \in (1, \infty)$ and jointly continuous in the norm topology for every $\alpha \in [1/2, 1)$.*
(4) Monotonicity: *When $\alpha \in (1, \infty)$, for any unital normal positive map $\gamma : N \to M$ between von Neumann algebras and for every $\psi, \varphi \in M_*^+$,*

$$\widetilde{D}_\alpha(\psi \circ \gamma \| \varphi \circ \gamma) \leq \widetilde{D}_\alpha(\psi\|\varphi).$$

When $\alpha \in [1/2, 1)$, the same holds if γ is a unital normal CP map.
(5) *The map $(\psi, \varphi) \in M_*^+ \times M_*^+ \mapsto \widetilde{Q}_\alpha(\psi\|\varphi)$ is jointly convex for $1 < \alpha < \infty$ and jointly concave for $1/2 \leq \alpha < 1$. Hence, when $1/2 \leq \alpha < 1$, $\widetilde{D}_\alpha(\psi\|\varphi)$ is jointly convex on $\{(\psi, \varphi) \in M_*^+ \times M_*^+ : \psi(1) = c\}$ for any fixed $c > 0$.*
(6) Relation between $\widetilde{D}_\alpha$ and D_α: *For every $\alpha \in [1/2, \infty) \setminus \{1\}$,*

$$\widetilde{D}_\alpha(\psi\|\varphi) \leq D_\alpha(\psi\|\varphi). \tag{3.26}$$

In particular, $F(\psi, \varphi) \geq P(\psi, \varphi)$, see (3.24) and (3.25). Moreover, $D_2(\psi\|\varphi) \leq D_{\max}(\psi\|\varphi) \leq D_\infty(\psi\|\varphi) := \lim_{\alpha \to \infty} D_\alpha(\psi\|\varphi)$.
(7) *Let $\psi_i, \varphi_i \in M_*^+$ for $i = 1, 2$. If $1/2 \leq \alpha < 1$, $\psi_1 \leq \psi_2$ and $\varphi_1 \leq \varphi_2$, then $\widetilde{Q}_\alpha(\psi_1\|\varphi_1) \leq \widetilde{Q}_\alpha(\psi_2\|\varphi_2)$. If $1 < \alpha < \infty$, $\psi_1 \geq \psi_2$ and $\varphi_1 \leq \varphi_2$, then $\widetilde{Q}_\alpha(\psi_1\|\varphi_1) \geq \widetilde{Q}_\alpha(\psi_2\|\varphi_2)$. Hence, if $\varphi_1 \leq \varphi_2$, then $\widetilde{Q}_\alpha(\psi\|\varphi_1) \geq \widetilde{Q}_\alpha(\psi\|\varphi_2)$ for all $\alpha \in [1/2, \infty) \setminus \{1\}$.*
(8) Strict positivity: *Let $\alpha \in [1/2, \infty) \setminus \{1\}$ and $\psi, \varphi \neq 0$. The inequality*

$$\widetilde{D}_\alpha(\psi\|\varphi) \geq \log \frac{\psi(1)}{\varphi(1)}$$

holds, and equality holds if and only if $\psi = (\psi(1)/\varphi(1))\varphi$. If $\psi(1) = \varphi(1)$, then $\widetilde{D}_\alpha(\psi\|\varphi) \geq 0$, and $\widetilde{D}_\alpha(\psi\|\varphi) = 0$ if and only if $\psi = \varphi$.

Proof (1) See [17, Lemma 8] and [59, Proposition 3.7].

(2) See [17, Theorem 13, Lemma 9] and [59, Proposition 3.8].

(3) The joint lower semicontinuity of $\widetilde{D}_\alpha$ for $1 < \alpha < \infty$ was given in [59, Proposition 3.10]. For the joint continuity of $\widetilde{D}_\alpha$ (or $\widetilde{Q}_\alpha$) for $1/2 \leq \alpha < 1$, by (3.12) it suffices to show that $(\psi, \varphi) \in M_*^+ \times M_*^+ \mapsto h_\varphi^{\frac{1-\alpha}{2\alpha}} h_\psi h_\varphi^{\frac{1-\alpha}{2\alpha}} \in L^\alpha(M)$ is continuous, whose proof is similar to that of Corollary 3.8 by using $\|a + b\|_p^p \leq \|a\|_p^p + \|b\|_p^p$ for $a, b \in L^p(M)$ $(0 < p \leq 1)$ (see [33, Theorem 4.9 (iii)]), Hölder's inequality and Kosaki's generalized Powers-Størmer inequality.

(4) See [59, Theorem 3.14] and [17, Theorem 14] (in fact, we have used the latter theorem in the proof of Lemma 3.12 (3)).

(5) The joint convexity of $\widetilde{Q}_\alpha$ for $1 < \alpha < \infty$ was given in [59, Corollary 3.16]. For $1/2 \leq \alpha < 1$ the monotonicity of $\widetilde{D}_\alpha$ in (4) means the "reverse" monotonicity of $\widetilde{Q}_\alpha$ under a unital normal CP map. Then the joint concavity of $\widetilde{Q}_\alpha$ for $1/2 \leq \alpha < 1$ can be shown similarly to the proof of [59, Corollary 3.16] (or Lemma 3.12 (3)) by considering $\gamma : M \to M \oplus M$, $\gamma(x) := x \oplus x$, and the expression of $\widetilde{Q}_\alpha$ in Theorem 3.11 (2).

(6) See [17, Theorem 12] and [59, Corollary 3.6], and see [17, (95)] for $D_2 \leq D_{\max}$. Furthermore, in view of (3.22) note that inequality (3.26) for $1/2 \leq \alpha < 1$ is equivalently written as

$$\| |h_\psi^{1/2} h_\varphi^{\frac{1-\alpha}{2\alpha}}|^\alpha \|_2 \leq \|h_\psi^{\alpha/2} h_\varphi^{\frac{1-\alpha}{2}}\|_2,$$

which is a special case of Kosaki's extension of the Araki-Lieb-Thirring inequality to Haagerup's L^p-norms [71, Theorem 4].

(7) The case $1 < \alpha < \infty$ was given in [59, Proposition 3.9]. When $1/2 \leq \alpha < 1$, Lemma B.5 of Appendix B implies that $\left(h_{\varphi_1}^{\frac{1-\alpha}{2\alpha}} h_{\psi_1} h_{\varphi_1}^{\frac{1-\alpha}{2\alpha}}\right)^\alpha \leq \left(h_{\varphi_1}^{\frac{1-\alpha}{2\alpha}} h_{\psi_2} h_{\varphi_1}^{\frac{1-\alpha}{2\alpha}}\right)^\alpha$. Since $\frac{1-\alpha}{\alpha} \in (0, 1]$, by Lemma B.5 again one has $h_{\varphi_1}^{\frac{1-\alpha}{\alpha}} \leq h_{\varphi_2}^{\frac{1-\alpha}{\alpha}}$ and hence $\left(h_{\psi_2}^{1/2} h_{\varphi_1}^{\frac{1-\alpha}{\alpha}} h_{\psi_2}^{1/2}\right)^\alpha \leq \left(h_{\psi_2}^{1/2} h_{\varphi_2}^{\frac{1-\alpha}{\alpha}} h_{\psi_2}^{1/2}\right)^\alpha$. Therefore,

$$\begin{aligned}
\operatorname{tr}\left(h_{\varphi_1}^{\frac{1-\alpha}{2\alpha}} h_{\psi_1} h_{\varphi_1}^{\frac{1-\alpha}{2\alpha}}\right)^\alpha &\leq \operatorname{tr}\left(h_{\varphi_1}^{\frac{1-\alpha}{2\alpha}} h_{\psi_2} h_{\varphi_1}^{\frac{1-\alpha}{2\alpha}}\right)^\alpha = \operatorname{tr}\left(h_{\psi_2}^{1/2} h_{\varphi_1}^{\frac{1-\alpha}{\alpha}} h_{\psi_2}^{1/2}\right)^\alpha \\
&\leq \operatorname{tr}\left(h_{\psi_2}^{1/2} h_{\varphi_2}^{\frac{1-\alpha}{\alpha}} h_{\psi_2}^{1/2}\right)^\alpha = \operatorname{tr}\left(h_{\psi_2}^{\frac{1-\alpha}{2\alpha}} h_{\psi_2} h_{\varphi_2}^{\frac{1-\alpha}{2\alpha}}\right)^\alpha,
\end{aligned}$$

that is, $\widetilde{Q}(\psi_1\|\varphi_1) \leq \widetilde{Q}_\alpha(\psi_2\|\varphi_2)$.

(8) By (1) and (4) for $\gamma : \mathbb{C}1 \hookrightarrow M$ one has

$$\widetilde{D}_\alpha(\psi\|\varphi) \geq \widetilde{D}_{1/2}(\psi\|\varphi) \geq -2\log \frac{\psi(1)^{1/2}\varphi(1)^{1/2}}{\psi(1)} = \log \frac{\psi(1)}{\varphi(1)}.$$

If $\psi = (\psi(1)/\varphi(1))\varphi$, then $\widetilde{D}_\alpha(\psi\|\varphi) \leq D_\alpha(\psi\|\varphi) = \log \psi(1)/\varphi(1)$ by (6) and Theorem 3.2 (8), so that $\widetilde{D}_\alpha(\psi\|\varphi) = \log \psi(1)/\varphi(1)$. Conversely, if this equality

holds, then $\widetilde{D}_{1/2}(\psi\|\varphi) = \log \psi(1)/\varphi(1)$, and $\psi = (\psi(1)/\varphi(1))\varphi$ follows as in the proof of [48, Corollary 4.2 (1)]. $\square$

Remark 3.16 The expression of $\widetilde{D}_\alpha$ in (3.22) with $\widetilde{Q}_\alpha(\psi\|\varphi) = \mathrm{tr}\left(h_\varphi^{\frac{1-\alpha}{2\alpha}} h_\psi h_\varphi^{\frac{1-\alpha}{2\alpha}}\right)^\alpha$ makes sense for all $\alpha \in (0, 1)$. In the matrix case, as seen from discussions in [34] that the monotonicity of $\widetilde{D}_\alpha$ ($0 < \alpha < 1$) under CP trace-preserving maps is equivalent to the joint concavity of $\widetilde{Q}_\alpha$. But we note by [46, Proposition 5.1] that $\widetilde{Q}_\alpha$ is not jointly concave (even not separately concave in the second variable) for $0 < \alpha < 1/2$. Thus, the monotonicity of $\widetilde{D}_\alpha$ under unital normal CP maps in (4) above is restricted to $\alpha \in [1/2, \infty)$, while that of D_α in Theorem 3.2 (7) is restricted to $\alpha \in [0, 2]$. See also [16, Theorem 7] for violation of the monotonicity property of $\widetilde{D}_\alpha$ when $\alpha < 1/2$.

Remark 3.17 Let $\psi, \varphi \in M_*^+$ with $\psi \neq 0$. If ψ, φ commute in the sense that h_ψ, h_φ commute (see Lemma 4.19 and Proposition A.55 of Appendix A.7), then

$$\widetilde{D}_\alpha(\psi\|\varphi) = D_\alpha(\psi\|\varphi) \quad \text{for all } \alpha \in [1/2, \infty) \setminus \{1\}.$$

This is obvious for $\alpha \in [1/2, 1)$, since $\widetilde{Q}_\alpha(\psi\|\varphi) = \mathrm{tr}\, h_\psi^\alpha h_\varphi^{1-\alpha} = Q_\alpha(\psi\|\varphi)$ by (3.23) and (3.7). When $\alpha > 1$, since $\widetilde{D}_\alpha(\psi\|\varphi) \leq D_\alpha(\psi\|\varphi)$ by (3.26), we may assume that $\widetilde{D}_\alpha(\psi\|\varphi) < +\infty$. Then there exists an $x \in (s(\varphi)L^\alpha(M)s(\varphi))_+$ such that $h_\psi = h_\varphi^{\frac{\alpha-1}{2\alpha}} x h_\varphi^{\frac{\alpha-1}{2\alpha}}$. Since h_ψ, h_φ commute, by taking the spectral decomposition of h_ψ, h_φ with the common spectral resolution, we have $h_\psi^{\alpha/2} = \eta h_\varphi^{(\alpha-1)/2}$ with $\eta := x^{\alpha/2} \in s(\varphi)L^2(M)s(\varphi)$. Theorem 3.6 gives $Q_\alpha(\psi\|\varphi) = \|\eta\|_2^2 = \|x\|_\alpha^\alpha = \|h_\psi\|_{\alpha,\varphi}^\alpha = \widetilde{Q}(\psi\|\varphi)$.

We end the section with the next variational expression of $\widetilde{Q}_\alpha$ for all $\alpha \in (0, 1)$ as touched in Remark 3.16, which will be used in Example 5.18 of Sec. 5.3.

Lemma 3.18 *Let* $0 < \alpha < 1$ *and* $\psi, \varphi \in M_*^+$. *Define* $\widetilde{Q}(\psi\|\varphi) := \mathrm{tr}\left(h_\varphi^{\frac{1-\alpha}{2\alpha}} h_\psi h_\varphi^{\frac{1-\alpha}{2\alpha}}\right)^\alpha$. *Then*

$$\widetilde{Q}_\alpha(\psi\|\varphi) \leq \inf_{x \in M_{++}} \left[\alpha\,\mathrm{tr}\,(h_\psi x) + (1-\alpha)\mathrm{tr}\,(h_\varphi^{\frac{1-\alpha}{2\alpha}} x^{-1} h_\varphi^{\frac{1-\alpha}{2\alpha}})^{\frac{\alpha}{1-\alpha}}\right]. \tag{3.27}$$

Furthermore, if $\psi \leq \lambda\varphi$ *for some* $\lambda > 0$, *then the above inequality becomes equality:*

$$\widetilde{Q}_\alpha(\psi\|\varphi) = \inf_{x \in M_{++}} \left[\alpha\,\mathrm{tr}\,(h_\psi x) + (1-\alpha)\mathrm{tr}\,(h_\varphi^{\frac{1-\alpha}{2\alpha}} x^{-1} h_\varphi^{\frac{1-\alpha}{2\alpha}})^{\frac{\alpha}{1-\alpha}}\right]. \tag{3.28}$$

Proof Write $R_\alpha(\psi\|\varphi)$ for the RHS of (3.27). For every $x \in M_{++}$ we have

$$\widetilde{Q}_\alpha(\psi\|\varphi) = \|h_\psi^{1/2} h_\varphi^{\frac{1-\alpha}{2\alpha}}\|_{2\alpha}^2 = \|h_\psi^{1/2} x^{1/2} x^{-1/2} h_\varphi^{\frac{1-\alpha}{2\alpha}}\|_{2\alpha}^2$$

$$\leq \left(\|h_\psi^{1/2} x^{1/2}\|_2 \|x^{-1/2} h_\varphi^{\frac{1-\alpha}{2\alpha}}\|_{\frac{2\alpha}{1-\alpha}}\right)^{2\alpha} \quad \text{(by Hölder's inequality)}$$

$$= [\operatorname{tr}(h_\psi^{1/2} x h_\psi^{1/2})]^\alpha \left[\operatorname{tr}(h_\varphi^{\frac{1-\alpha}{2\alpha}} x^{-1} h_\sigma^{\frac{1-\alpha}{2\alpha}})^{\frac{\alpha}{1-\alpha}}\right]^{1-\alpha}$$

$$\leq \alpha \operatorname{tr}(h_\psi x) + (1-\alpha)\operatorname{tr}(h_\varphi^{\frac{1-\alpha}{2\alpha}} x^{-1} h_\varphi^{\frac{1-\alpha}{2\alpha}})^{\frac{\alpha}{1-\alpha}},$$

which implies that $\widetilde{Q}_\alpha(\psi\|\varphi) \leq R_\alpha(\psi\|\varphi)$.

To prove the latter assertion, first assume that $\lambda^{-1}\varphi \leq \psi \leq \lambda\varphi$ for some $\lambda > 1$. We may assume that ψ, φ are faithful. Write $\beta := \frac{1-\alpha}{\alpha} \in (0, \infty)$, so $(1+\beta)(1-\alpha) = \beta$. Since $\lambda^{-1} h_\varphi^{1+\beta} \leq h_\varphi^{\beta/2} h_\psi h_\varphi^{\beta/2} \leq \lambda h_\varphi^{1+\beta}$, it follows from (A.24) of Lemma A.58 that there are $a_0, b_0 \in M$ such that

$$h_\varphi^{\beta/2} = a_0 (h_\varphi^{\beta/2} h_\psi h_\varphi^{\beta/2})^{\frac{1-\alpha}{2}}, \qquad (h_\varphi^{\beta/2} h_\psi h_\varphi^{\beta/2})^{\frac{1-\alpha}{2}} = b_0 h_\varphi^{\beta/2}.$$

It is immediate to verify that $a_0 b_0 = b_0 a_0 = 1$ so that $b_0 = a_0^{-1}$. Set $x_0 := a_0 a_0^*$ so that $x_0^{-1} = (a_0^{-1})^* a_0^{-1} = b_0^* b_0$. Since

$$h_\varphi^{\beta/2} h_\psi h_\varphi^{\beta/2} = (h_\varphi^{\beta/2} h_\psi h_\varphi^{\beta/2})^{\frac{1-\alpha}{2}} a_0^* h_\psi a_0 (h_\varphi^{\beta/2} h_\psi h_\varphi^{\beta/2})^{\frac{1-\alpha}{2}},$$

one has $a_0^* h_\psi a_0 = (h_\varphi^{\beta/2} h_\psi h_\varphi^{\beta/2})^\alpha$ so that

$$\operatorname{tr}(h_\psi x_0) = \operatorname{tr}(a_0^* h_\psi a_0) = \operatorname{tr}(h_\varphi^{\beta/2} h_\psi h_\varphi^{\beta/2})^\alpha. \tag{3.29}$$

Moreover, since

$$h_\varphi^{\beta/2} x_0^{-1} h_\varphi^{\beta/2} = h_\varphi^{\beta/2} b_0^* b_0 h_\varphi^{\beta/2} = (h_\varphi^{\beta/2} h_\psi h_\varphi^{\beta/2})^{1-\alpha},$$

one has

$$\operatorname{tr}(h_\varphi^{\beta/2} x_0^{-1} h_\varphi^{\beta/2})^{1/\beta} = \operatorname{tr}(h_\varphi^{\beta/2} h_\psi h_\varphi^{\beta/2})^\alpha. \tag{3.30}$$

From (3.29) and (3.30), $R_\alpha(\psi\|\varphi) \leq \operatorname{tr}(h_\varphi^{\beta/2} h_\psi h_\varphi^{\beta/2})^\alpha = \widetilde{Q}_\alpha(\psi\|\varphi)$. Therefore, $Q_\alpha(\psi\|\varphi) = R_\alpha(\psi\|\varphi)$ holds.

Next, assume that $\psi \leq \lambda\varphi$ for some $\lambda > 0$. From the proofs of (3) and (7) of Theorem 3.15 we note that the continuity property of Theorem 3.15 (3) and the monotonicity of $\psi \mapsto \widetilde{Q}_\alpha(\psi\|\varphi)$ hold for all $\alpha \in (0, 1)$. Hence we have

$$\widetilde{Q}_\alpha(\psi\|\varphi) = \inf_{\varepsilon>0} \widetilde{Q}_\alpha(\psi + \varepsilon\varphi\|\varphi)$$

$$= \inf_{\varepsilon>0} \inf_{x\in M_{++}} \left[\alpha\mathrm{tr}\,((h_\psi + \varepsilon h_\varphi)x) + (1-\alpha)\mathrm{tr}\,(h_\varphi^{\beta/2}x^{-1}h_\varphi^{\beta/2})^{1/\beta}\right]$$

$$= \inf_{x\in M_{++}} \inf_{\varepsilon>0} \left[\alpha\mathrm{tr}\,((h_\psi + \varepsilon h_\varphi)x) + (1-\alpha)\mathrm{tr}\,(h_\varphi^{\beta/2}x^{-1}h_\varphi^{\beta/2})^{1/\beta}\right]$$

$$= \inf_{x\in M_{++}} \left[\alpha\mathrm{tr}\,(h_\psi x) + (1-\alpha)\mathrm{tr}\,(h_\varphi^{\beta/2}x^{-1}h_\varphi^{\beta/2})^{1/\beta}\right],$$

where the second equality holds by the above case.   □

Problem 3.19 (1) It is interesting to find whether (3.28) holds for general ψ, φ. But it does not seem true that $\varphi \in M_*^+ \mapsto \mathrm{tr}\,(h_\varphi^{\beta/2}x^{-1}h_\varphi^{\beta/2})^{1/\beta}$ is continuous in the $\sigma(M_*, M)$-topology, so probably we cannot have the $\sigma(M_*, M)$-lower semicontinuity of $\widetilde{D}_\alpha$ even when (5.5) holds for general ψ, φ.

(2) Let $e := s(\varphi)$. Since $\widetilde{Q}(\psi\|\varphi) = \widetilde{Q}_\alpha(e\psi e\|\varphi)$, one can see that the variational expression of $\widetilde{Q}_\alpha(\psi\|\varphi)$ in (3.28) holds true if $e\psi e \le \lambda\varphi$ for some $\lambda > 0$. In the finite-dimensional case, this is automatic so that the above variational expression of $\widetilde{Q}_\alpha$ for $0 < \alpha < 1$ holds for all ψ, φ, as given in [34, Lemma 4]. A similar variational expression of $\widetilde{Q}_\alpha$ for $1 < \alpha < \infty$ is also in [34, Lemma 4], so it is interesting to prove the same expression of $\widetilde{Q}_\alpha$ for $\alpha > 1$ in the von Neumann algebra case.

Chapter 4
Maximal f-Divergences

4.1 Definition and Basic Properties

Let M be a von Neumann algebra with its standard form $(M.\mathcal{H}, J, \mathcal{P})$ as before. Throughout this chapter we assume that f is an operator convex function on $(0, \infty)$. For $\psi, \varphi \in M_*^+$ we write $\psi \sim \varphi$ if $\delta\varphi \leq \psi \leq \delta^{-1}\varphi$ for some $\delta > 0$, and for convenience we set

$$(M_*^+ \times M_*^+)_\sim := \{(\psi, \varphi) \in M_*^+ \times M_*^+ : \psi \sim \varphi\},$$
$$(M_*^+ \times M_*^+)_\leq := \{(\psi, \varphi) \in M_*^+ \times M_*^+ : \psi \leq \alpha\varphi \text{ for some } \alpha > 0\},$$

which are convex sets. We first define the maximal f-divergence for $(\psi, \varphi) \in (M_*^+ \times M_*^+)_\sim$ and then extend it to general $\psi, \varphi \in M_*^+$.

Definition 4.1 Let $(\psi, \varphi) \in (M_*^+ \times M_*^+)_\sim$ with respective vector representatives $\Psi, \Phi \in \mathcal{P}$. By Lemma A.21 of Appendix A.3 we have a unique $A \in s(\varphi)Ms(\varphi)$ such that $\Psi = A\Phi$. Write $T_{\psi/\varphi} := A^*A$, which is a positive invertible operator in $s(\varphi)Ms(\varphi)$, and define a self-adjoint operator $f(T_{\psi/\varphi})$ in $s(\varphi)Ms(\varphi)$ via functional calculus. Then we define the *maximal f-divergence* of ψ with respect to φ by

$$\widehat{S}_f(\psi\|\varphi) := \varphi(f(T_{\psi/\varphi})) \in \mathbb{R}. \tag{4.1}$$

When $\psi, \varphi \in B(\mathcal{H})_*^+$, $\psi \sim \varphi$, on a finite-dimensional Hilbert space $\mathcal{H}$ with the density operators D_ψ, D_φ, the above (4.1) is given as

$$\widehat{S}_f(\psi\|\varphi) = \operatorname{Tr} D_\varphi^{1/2} f(D_\varphi^{-1/2} D_\psi D_\varphi^{-1/2}) D_\varphi^{1/2},$$

that is, $\widehat{S}_f(\psi\|\varphi)$ is the trace value of the *operator perspective* [32, 31] of D_ψ, D_φ by the function f. This notion of quantum divergence was first introduced by Petz and Ruskai [97] and later developed in more detail in [52, 79].

To extend $\widehat{S}_f(\psi\|\varphi)$ to general $\psi, \varphi \in M_*^+$, we need the following:

Lemma 4.2 ([49]) *Let* $\psi, \varphi \in M_*^+$.

(1) *For every $\omega \in M_*^+$ with $\omega \sim \psi + \varphi$, the limit*

$$\lim_{\varepsilon \searrow 0} \widehat{S}_f(\psi + \varepsilon\omega \| \varphi + \varepsilon\omega) \in (-\infty, +\infty]$$

exists, and moreover the limit is independent of the choice of ω as above.

(2) *If $(\psi, \varphi) \in (M_*^+ \times M_*^+)_\sim$, then*

$$\widehat{S}_f(\psi \| \varphi) = \lim_{\varepsilon \searrow 0} \widehat{S}_f(\psi + \varepsilon\varphi \| (1 + \varepsilon)\varphi).$$

The lemma says that the next definition is well defined independently of the choice of ω and extend Definition 4.1 for the case $\psi \sim \varphi$.

Definition 4.3 For every $\psi, \varphi \in M_*^+$ define the *maximal f-divergence* $\widehat{S}_f(\psi \| \varphi)$ by

$$\widehat{S}_f(\psi \| \varphi) := \lim_{\varepsilon \searrow 0} \widehat{S}_f(\psi + \varepsilon\omega \| \varphi + \varepsilon\omega) \in (-\infty, +\infty] \tag{4.2}$$

for any $\omega \in M_*^+$ with $\omega \sim \psi + \varphi$, where $\widehat{S}_f(\psi + \varepsilon\omega \| \varphi + \varepsilon\omega)$ is defined in Definition 4.1.

The most basic properties of $\widehat{S}_f(\psi \| \varphi)$ are the following:

Theorem 4.4 ([49])

(i) Monotonicity: *Let N be another von Neumann algebra and $\gamma : N \to M$ be a unital normal positive map. Then for every $\psi, \varphi \in M_*^+$,*

$$\widehat{S}_f(\psi \circ \gamma \| \varphi \circ \gamma) \le \widehat{S}_f(\psi \| \varphi).$$

(ii) Joint convexity: *For every $\psi_i, \varphi_i \in M_*^+$ and $\lambda_i \ge 0$ $(1 \le i \le n)$,*

$$\widehat{S}_f\left(\sum_{i=1}^n \lambda_i \psi_i \,\Big\|\, \sum_{i=1}^n \lambda_i \varphi_i\right) \le \sum_{i=1}^n \lambda_i \widehat{S}_f(\psi_i \| \varphi_i).$$

Our proofs of Lemma 4.2 and Theorem 4.4 in [49] go as follows: The properties in (i) and (ii) of Theorem 4.4 are first shown restricted to $(\psi, \varphi) \in (M_*^+ \times M_*^+)_\sim$, from which we can prove the convergence results in Lemma 4.2. Then from definition (4.2) the properties of Theorem 4.4 can extend to general $\psi, \varphi \in M_*^+$.

Another significant property of $\widehat{S}_f(\psi \| \varphi)$ is the joint lower semicontinuity, which will be given later after developing a general integral formula in the next section. Some other basic properties are summarized in the next proposition.

Proposition 4.5 ([49]) *Let $\psi, \varphi \in M_*^+$.*

(1) *Let $\widetilde{f}$ be the transpose of f (see Proposition 2.3 (5)). Then*

$$\widehat{S}_{\widetilde{f}}(\psi \| \varphi) = \widehat{S}_f(\varphi \| \psi).$$

(2) *Let $\psi_i, \varphi_i \in M_*^+$ $(i = 1, 2)$. If $s(\psi_1) \vee s(\varphi_1) \perp s(\psi_2) \vee s(\varphi_2)$, then*

$$\widehat{S}_f(\psi_1 + \psi_2 \| \varphi_1 + \varphi_2) = \widehat{S}_f(\psi_1 \| \varphi_1) + \widehat{S}_f(\psi_2 \| \varphi_2).$$

(3) *When $(\psi, \varphi) \in (M_*^+ \times M_*^+)_\le$,*

$$\widehat{S}_f(\psi \| \varphi) = \lim_{\varepsilon \searrow 0} \widehat{S}_f(\psi + \varepsilon \varphi \| \varphi).$$

When $(\varphi, \psi) \in (M_*^+ \times M_*^+)_\le$,

$$\widehat{S}_f(\psi \| \varphi) = \lim_{\varepsilon \searrow 0} \widehat{S}_f(\psi \| \varphi + \varepsilon \psi).$$

(4) *If $s(\psi) \not\le s(\varphi)$ and $f'(\infty) = +\infty$, then $\widehat{S}_f(\psi \| \varphi) = +\infty$. If $s(\varphi) \not\le s(\psi)$ and $f(0^+) = +\infty$, then $\widehat{S}_f(\psi \| \varphi) = +\infty$.*

(5) *If $f(0^+) < +\infty$, then*

$$\widehat{S}_f(\psi \| \varphi) = \lim_{\varepsilon \searrow 0} \widehat{S}_f(\psi \| \varphi + \varepsilon \psi),$$

and expression (4.1) holds, where $f(0) = f(0^+)$. Also, if $f'(\infty) < +\infty$, then

$$\widehat{S}_f(\psi \| \varphi) = \lim_{\varepsilon \searrow 0} \widehat{S}_f(\psi + \varepsilon \varphi \| \varphi).$$

Example 4.6 (1) Assume that $M = L^\infty(X, \mu)$ is an abelian von Neumann algebra on a σ-finite measure space $(X, \mathcal{X}, \mu)$. Let $\psi, \varphi \in M_*^+$, identified with functions in $L^1(X, \mu)_+$ so that $\psi(\phi) = \int_X \phi \psi \, d\mu$, $\varphi(\phi) = \int_X \phi \varphi \, d\mu$ for $\phi \in L^\infty(X, \mu)$. With $\omega = \psi + \varphi$ we have

$$\widehat{S}_f(\psi + \varepsilon \omega \| \varphi + \varepsilon \omega) = S_f(\psi + \varepsilon \omega \| \varphi + \varepsilon \omega),$$

see Example 2.2 (1). Take the limit of the above as $\varepsilon \searrow 0$ to see that $\widehat{S}_f(\psi \| \varphi)$ coincides with the classical f-divergence $S_f(\psi \| \varphi) = \int_X \varphi f(\psi / \varphi) \, d\mu$.

(2) Assume that $M = B(\mathcal{H})$ on a finite-dimensional Hilbert space $\mathcal{H}$. Let $\psi, \varphi \in B(\mathcal{H})_+^*$ with the density operators D_ψ, D_φ. When ψ, φ are faithful, (4.1) becomes

$$\widehat{S}_f(\psi \| \varphi) = \operatorname{Tr} D_\varphi f(D_\varphi^{-1/2} D_\psi D_\varphi^{-1/2}).$$

For general $\psi, \varphi \in B(\mathcal{H})_+^*$ let e be the support projection of $\psi + \varphi$. By Proposition 4.5 (2) we have

$$\begin{aligned}
\lim_{\varepsilon \searrow 0} \widehat{S}_f(\psi + \varepsilon I \| \varphi + \varepsilon I) &= \lim_{\varepsilon \searrow 0} \{\widehat{S}_f(\psi + \varepsilon e \| \varphi + \varepsilon e) + \widehat{S}_f(\varepsilon(I - e) \| \varepsilon(I - e))\} \\
&= \lim_{\varepsilon \searrow 0} \{\widehat{S}_f(\psi + \varepsilon e \| \varphi + \varepsilon e) + \varepsilon \operatorname{Tr}(I - e) f(1)\} \\
&= \lim_{\varepsilon \searrow 0} \widehat{S}_f(\psi + \varepsilon e \| \varphi + \varepsilon e),
\end{aligned}$$

so that $\widehat{S}_f(\psi\|\varphi)$ coincides with that defined in [52, Definition 3.21].

Example 4.7 (1) Consider a linear function $f(t) = a + bt$ with $a, b \in \mathbb{R}$. Let $(\psi, \varphi) \in (M_*^+ \times M_*^+)_\sim$ with $e = s(\psi) = s(\varphi)$. Let $A \in eMe$ be as in Lemma A.21 so that $\Psi = A\Phi$; then

$$\widehat{S}_{a+bt}(\psi\|\varphi) = \varphi(ae + bA^*A) = a\varphi(e) + b\|A\Phi\|^2 = a\varphi(1) + b\psi(1).$$

Hence by Definition 4.3 and (2.4) we have for every $\psi, \varphi \in M_*^+$,

$$\widehat{S}_{a+bt}(\psi\|\varphi) = S_{a+bt}(\psi\|\varphi) = a\varphi(1) + b\psi(1). \tag{4.3}$$

(2) Consider $f(t) = t^2$ and show that for every $\psi, \varphi \in M_*^+$,

$$\widehat{S}_{t^2}(\psi\|\varphi) = S_{t^2}(\psi\|\varphi). \tag{4.4}$$

Let $(\psi, \varphi) \in (M_*^+ \times M_*^+)_\sim$ and h_ψ, h_φ be the corresponding elements in $L^1(M)$ (Haagerup's L^1-space). With A in Lemma A.21 we have

$$\widehat{S}_{t^2}(\psi\|\varphi) = \mathrm{tr}\,(h_\varphi^{1/2}(A^*A)^2 h_\varphi^{1/2}) = \mathrm{tr}\,(h_\psi^{1/2}AA^* h_\psi^{1/2}) = \|A^* h_\psi^{1/2}\|_2^2.$$

On the other hand,

$$S_{t^2}(\psi\|\varphi) = \|\Delta_{\psi,\varphi} h_\varphi^{1/2}\|_2^2 = \|\Delta_{\psi,\varphi}^{1/2} h_\psi^{1/2}\|_2^2 = \|\Delta_{\psi,\varphi}^{1/2} A h_\varphi^{1/2}\|_2^2 = \|A^* h_\psi^{1/2}\|_2^2,$$

so (4.4) follows in this case. Therefore, in view of (4.2) and (2.16), this holds for all $\psi, \varphi \in M_*^+$. By (4.3) and (4.4) together, $\widehat{S}_f = S_f$ if f is a quadratic polynomial.

4.2 Further Properties of Maximal f-Divergences

Let $\psi, \varphi \in M_*^+$ and let $\Psi, \Omega \in \mathcal{P}$ be the vector representatives of ψ and $\omega := \psi + \varphi$. We have a unique $A \in s(\omega)Ms(\omega)$ such that $\Psi = A\Omega$. Define a positive operator $T_{\psi/\omega} := A^*A$ in $s(\omega)Ms(\omega)$ as in Definition 4.1. Since $0 \le T_{\psi/\omega} \le 1$, we write the spectral decomposition of $T_{\psi/\omega}$ as

$$T_{\psi/\omega} = \int_0^1 t\, dE_{\psi/\omega}(t). \tag{4.5}$$

In the next theorem we present a general integral formula of $\widehat{S}_f(\psi\|\varphi)$, which may serve as the second definition of maximal f-divergences.

Theorem 4.8 ([49]) *Let $\psi, \varphi \in M_*^+$ and $\omega := \psi + \varphi$ with the vector representative Ω. Then for every operator convex function f on $(0, \infty)$,*

$$\widehat{S}_f(\psi\|\varphi) = \int_0^1 (1-t)f\left(\frac{t}{1-t}\right) d\|E_{\psi/\omega}(t)\Omega\|_2^2, \tag{4.6}$$

where $(1-t)f\left(\frac{t}{1-t}\right)$ is understood as $f(0^+)$ for $t = 0$ and $f'(\infty)$ for $t = 1$.

Corollary 4.9 *If $f(0^+) < +\infty$ and $f'(+\infty) < +\infty$, then $\widehat{S}_f(\psi\|\varphi)$ is finite for every $\psi, \varphi \in M_*^+$.*

Proof If $f(0^+) < +\infty$ and $f'(+\infty) < +\infty$, then the function $(1-t)f\left(\frac{t}{1-t}\right)$ is bounded on $[0, 1]$. Hence the result is obvious from expression (4.6). □

The proof of Theorem 4.8 in [49] is based on the integral representation (2.5) of f. Here it is worth noting that the integrand of (4.6) enjoys a special theoretical meaning as follows. Let f be a real function on $(0, \infty)$ and g be given as $g(t) := (1-t)f\left(\frac{t}{1-t}\right)$, $t \in (0, 1)$. Then f is operator convex on $(0, \infty)$ if and only if g is operator convex on $(0, 1)$. Moreover, in this case, $g(0^+) = f(0^+)$ and $g(1^-) = f'(\infty)$. We include the proof of this fact in Appendix C.

Remark 4.10 Similarly to the definition of the standard f-divergence $S_f(\psi\|\varphi)$ in (2.3), one can write (4.6) as the sum of three terms

$$\widehat{S}_f(\psi\|\varphi) = \int_{(0,1)} (1-t)f\left(\frac{t}{1-t}\right) d\|E_{\psi/\omega}(t)\Omega\|_2^2$$
$$+ f(0^+)\langle\Omega, E_{\psi/\omega}(\{0\})\Omega\rangle + f'(\infty)\langle\Omega, E_{\psi/\omega}(\{1\})\Omega\rangle, \tag{4.7}$$

where $E_{\psi/\omega}(\{0\})$ and $E_{\psi/\omega}(\{1\})$ are the spectral projections of $T_{\psi/\omega}$ for $\{0\}$ and $\{1\}$. One might expect that the above boundary terms with $f(0^+)$ and $f'(\infty)$ are equal to the corresponding terms $f(0^+)\varphi(1 - s(\psi))$ and $f'(\infty)\psi(1 - s(\varphi))$, respectively, in (2.3). But it is not true as shown in [49, Example 4.6].

In the following we state the lower semicontinuity in the norm and the martingale convergence for $\widehat{S}_f(\psi\|\varphi)$.

Theorem 4.11 ([49]) *The function $(\psi, \varphi) \in M_*^+ \times M_*^+ \mapsto \widehat{S}_f(\psi\|\varphi)$ is jointly lower semicontinuous in the norm topology.*

Theorem 4.12 ([49]) *Let $\{M_\alpha\}$ be an increasing net of unital von Neumann subalgebras of M such that $\left(\bigcup_\alpha M_\alpha\right)'' = M$. Then for every $\psi, \varphi \in M_*^+$,*

$$\widehat{S}_f(\psi|_{M_\alpha}\|\varphi|_{M_\alpha}) \nearrow \widehat{S}_f(\psi\|\varphi).$$

The proofs of these properties are based on the integral formula (4.6). For each $n \in \mathbb{N}$ consider the approximation f_n of f in (2.6) with integral on the cut-off interval $[1/n, n]$. From (2.9) we have as $n \to \infty$,

$$(1-t)f_n\left(\frac{t}{1-t}\right) \nearrow (1-t)f\left(\frac{t}{1-t}\right), \qquad t \in [0, 1].$$

Hence from the monotone convergence theorem applied to the integral formula in (4.6) for f_n and f we have for every $\psi, \varphi \in M_*^+$,

$$\widehat{S}_f(\psi\|\varphi) = \lim_{n\to\infty} \widehat{S}_{f_n}(\psi\|\varphi) \quad \text{increasingly,}$$

so that $\widehat{S}_f(\psi\|\varphi) = \sup_{n\geq 1} \widehat{S}_{f_n}(\psi\|\varphi)$. Moreover, since

$$\lim_{t\searrow 0} f_n'(t) = \lim_{t\searrow 0} \frac{f_n(t) - f_n(0^+)}{t} > -\infty,$$

to prove Theorems 4.11 and 4.12, we may assume that $f(0^+) < +\infty$, $f'(\infty) < +\infty$ and $\lim_{t\searrow 0} f'(t) > -\infty$. See [49, Sec. V] for the details of the proofs of the theorems.

Problem 4.13 Theorem 2.7 (i) says that the standard f-divergence $S_f(\psi\|\varphi)$ is jointly lower semicontinuous in the $\sigma(M_*, M)$-topology. It follows from Theorem 4.12 that this property (stronger than Theorem 4.11) holds for $\widehat{S}_f(\psi\|\varphi)$ as well when M is injective, or equivalently, AFD, i.e., there is an increasing net $\{M_\alpha\}$ of finite-dimensional unital subalgebras of M such that $M = (\bigcup_\alpha M_\alpha)''$, see Appendix A.1. In fact, in this case, $\widehat{S}_f(\psi\|\varphi) = \sup_\alpha \widehat{S}_f(\psi|_{M_\alpha}\|\varphi|_{M_\alpha})$ by Theorem 4.12 and $(\psi, \varphi) \mapsto \widehat{S}_f(\psi|_{M_\alpha}\|\varphi|_{M_\alpha})$ is lower semicontinuous in the $\sigma(M_*, M)$-topology. However, it is unknown whether $\widehat{S}_f(\psi\|\varphi)$ is jointly lower semicontinuous in the $\sigma(M_*, M)$-topology for general M.

Another martingale convergence like Proposition 2.10 for S_f holds for $\widehat{S}_f$ as well.

Proposition 4.14 *Let $\{e_\alpha\}$ be an increasing net of projections in M such that $e_\alpha \nearrow 1$. Then for every $\psi, \varphi \in M_*^+$,*

$$\lim_\alpha \widehat{S}_f(e_\alpha \psi e_\alpha \| e_\alpha \varphi e_\alpha) = \widehat{S}_f(\psi\|\varphi),$$

where $e_\alpha \varphi e_\alpha$ is the restriction of φ to the reduced von Neumann algebra $e_\alpha M e_\alpha$.

4.3 Minimal Reverse Test

For every $\psi, \varphi \in M_*^+$ let $\omega := \psi + \varphi$ with the vector representative Ω of ω as well as Ψ of ψ. We have a positive operator $T_{\psi/\omega} = A^*A = \int_0^1 t\, dE_{\psi/\omega}(t)$ in $s(\omega)Ms(\omega)$ as in (4.5), where $A \in s(\omega)Ms(\omega)$ satisfies $\Psi = A\Omega$. Define a finite Borel measure ν on $[0, 1]$ by

$$\nu := \omega(E_{\psi/\omega}(\cdot)) = \|E_{\psi/\omega}(\cdot)\Omega\|^2, \tag{4.8}$$

and consider an abelian von Neumann algebra $L^\infty([0, 1], \nu) = L^1([0, 1], \nu)^*$.

Lemma 4.15 ([49]) *Define $\lambda_0 : M \to L^\infty([0, 1], \nu)$ by*

$$\lambda_0(x) = \frac{d\langle Jx\Omega, E_{\psi/\omega}(\cdot)\Omega\rangle}{d\nu} \ \ \textit{(the Radon-Nikodym derivative)}, \quad x \in M.$$

Then λ_0 is a unital normal positive map and its predual map $\lambda_{0*} : L^1([0,1], \nu) \to M_*$ is given by

$$\lambda_{0*}(\phi)(x) = \left\langle Jx\Omega, \left(\int_0^1 \phi\, dE_{\psi/\omega}\right)\Omega\right\rangle$$

for every $\phi \in L^\infty([0,1], \nu)$ $(\subset L^1([0,1], \nu))$ and $x \in M$.

The above λ_0 satisfies

$$\lambda_{0*}(\phi)(1) = \omega\left(\int_0^1 \phi\, dE_{\psi/\omega}\right) = \int_0^1 \phi\, d\nu$$

for all $\phi \in L^\infty([0,1], \nu)$ (hence all $\phi \in L^1([0,1], \nu)$). Moreover, for every $x \in M$,

$$\lambda_{0*}(1)(x) = \langle Jx\Omega, \Omega\rangle = \langle \Omega, x\Omega\rangle = \omega(x),$$
$$\lambda_{0*}(t)(x) = \langle Jx\Omega, T_{\psi,\omega}\Omega\rangle = \langle JxJ\Omega, A^*A\Omega\rangle$$
$$= \langle JxJA\Omega, A\Omega\rangle = \langle Jx\Psi, \Psi\rangle = \langle \Psi, x\Psi\rangle = \psi(x),$$

where t denotes the identity function $t \mapsto t$ on $[0,1]$. Hence $\lambda_{0*}(t) = \psi$ and $\lambda_{0*}(1 - t) = \omega - \psi = \varphi$.

Now, following Matsumoto's idea [79], we introduce a key notion below.

Definition 4.16 Let $(X, \mathcal{X}, \mu)$ be a σ-finite measure space and $\lambda : M \to L^\infty(X, \mu)$ be a positive linear map which is unital and normal. Then the predual map $\lambda_* : L^1(X, \mu) \to M_*$ is trace-preserving in the sense that $\int_X \phi\, d\mu = \lambda_*(\phi)(1)$ $(= \|\lambda_*(\phi)\|)$ for every $\phi \in L^1(X, \mu)_+$. We call a triplet (λ, p, q) of such a map λ and $p, q \in L^1(X, \mu)_+$ a *reverse test* for ψ, φ if $\lambda_*(p) = \psi$ and $\lambda_*(q) = \varphi$, i.e., $p \circ \lambda = \psi$ and $q \circ \lambda = \varphi$ with identification $L^1(X, \mu)_+ = L^\infty(X, \mu)_*^+$.

The next variational formula of $\widehat{S}_f(\psi\|\varphi)$ may serve as the third definition of maximal f-divergences.

Theorem 4.17 ([49]) *For every $\psi, \varphi \in M_*^+$,*

$$\widehat{S}_f(\psi\|\varphi) = \min\{S_f(p\|q) : (\lambda, p, q) \text{ a reverse test for } \psi, \varphi\}. \tag{4.9}$$

Moreover, $\lambda : M \to L^\infty(X, \mu)$ in (4.9) can be restricted to those with a standard Borel probability space $(X, \mathcal{X}, \mu)$ or more specifically to a Borel probability space on $[0,1]$.

In fact, $(\lambda_0, t, 1 - t)$ given in Lemma 4.15 is a reverse test, for which the equality $\widehat{S}_f(\psi\|\varphi) = S_f(t\|1 - t)$ holds since from (4.6) and (4.8) we have

$$\widehat{S}_f(\psi\|\varphi) = \int_0^1 (1 - t) f\left(\frac{t}{1 - t}\right) d\nu(t) = S_f(t\|1 - t).$$

Therefore, the reverse test $(\lambda_0, t, 1 - t)$ is a minimizer for expression (4.9), which is considered as the von Neumann algebra version of Matsumoto's *minimal* or *optimal reverse test* [79] for ψ, φ. Apply the monotonicity property of S_f (Theorem 2.7 (iv)) to this λ_0 (that is a unital and normal CP map) to have

$$S_f(\psi\|\varphi) \le S_f(t\|1 - t) = \widehat{S}_f(\psi\|\varphi),$$

so we have

Theorem 4.18 ([49]) *For every $\psi, \varphi \in M_*^+$,*

$$S_f(\psi\|\varphi) \le \widehat{S}_f(\psi\|\varphi).$$

When $\mathcal{H}$ is finite-dimensional, it is easy to verify that if $\psi, \varphi \in B(\mathcal{H})_+^*$ commute in the sense that the density operators D_ψ, D_φ commute, then $S_f(\psi\|\varphi) = \widehat{S}_f(\psi\|\varphi)$ for every operator convex (even simply convex) function on $(0, \infty)$. We extend this to the general von Neumann algebra setting. To do this, we first state the next lemma whose proof was given in [49, Appendix B].

Lemma 4.19 *For every $\psi, \varphi \in M_*^+$ the following conditions are equivalent:*

(i) $\psi \circ \varphi_t^{\psi + \varphi} = \psi$ *on $s(\psi + \varphi)Ms(\psi + \varphi)$ for all $t \in \mathbb{R}$ (note that the modular automorphism group $\sigma_t^{\psi + \varphi}$ is defined on $s(\psi + \varphi)Ms(\psi + \varphi)$);*
(ii) $\varphi \circ \varphi_t^{\psi + \varphi} = \psi$ *on $s(\psi + \varphi)Ms(\psi + \varphi)$ for all $t \in \mathbb{R}$;*
(iii) $h_\psi h_\varphi = h_\varphi h_\psi$, *where h_ψ, h_φ are the corresponding elements of Haagerup's $L^1(M)$ (see Appendix A.6).*

When $s(\psi) \le s(\varphi)$, the above are also equivalent to $\psi \circ \sigma_t^\varphi = \psi$ for all $t \in \mathbb{R}$.

More equivalent conditions on the commutativity of ψ, φ are provided in Proposition A.55 of Appendix A.7.

Proposition 4.20 ([49]) *If $\psi, \varphi \in M_*^+$ commute in the sense that the equivalent conditions of Lemma 4.19 hold, then*

$$S_f(\psi\|\varphi) = \widehat{S}_f(\psi\|\varphi)$$

for any operator convex function f on $(0, \infty)$.

Example 4.21 In [14] Belavkin and Staszewski introduced a kind of relative entropy for states on a C^*-algebra. In the von Neumann algebras case, their relative entropy is realized as $\widehat{S}_f$ for $f(t) = t \log t$ (see [49, Example 3.5] for a detailed explanation). So *Belavkin and Staszewski's relative entropy* $D_{\mathrm{BS}}(\psi\|\varphi)$ for every $\psi, \varphi \in M_*^+$ is defined to be

$$D_{\mathrm{BS}}(\psi\|\varphi) := \widehat{S}_{t \log t}(\psi\|\varphi).$$

Since $D(\psi\|\varphi) = S_{t \log t}(\psi\|\varphi)$, by Theorem 4.18 and Proposition 4.20 we have for every $\psi, \varphi \in M_*^+$,

$$D(\psi\|\varphi) \le D_{\mathrm{BS}}(\psi\|\varphi),$$

and $D(\psi\|\varphi) = D_{\mathrm{BS}}(\psi\|\varphi)$ if ψ, φ commute.

Remark 4.22 In Theorem 7.6 of Sec. 7.2 we will prove that the converse direction of Proposition 4.20 is also true (i.e., $\psi, \varphi \in M_*^+$ commute if $S_f(\psi\|\varphi) = \widehat{S}_f(\psi\|\varphi) < +\infty$) under $s(\psi) \leq s(\varphi)$ and a certain mild assumption on f. (This was shown in [52, Theorem 4.3] in the finite-dimensional case.) In particular, for $\psi, \varphi \in M_*^+$ with $D(\psi\|\varphi) < +\infty$, it will follow that ψ, φ commute if and only if $D(\psi\|\varphi) = D_{\mathrm{BS}}(\psi\|\varphi)$.

Chapter 5
Measured f-Divergences

5.1 Definition

Let f be a convex function on $(0, \infty)$, not necessarily operator convex unless we specify that. We use the convention in (2.2). Let M be a general von Neumann algebra. A *measurement* $\mathcal{M}$ in M is given by $\mathcal{M} = (A_j)_{1 \le j \le n}$ for some $n \in \mathbb{N}$, where $A_j \in M_+$ for $1 \le j \le n$ and $\sum_{j=1}^n A_j = 1$. The measurement $\mathcal{M}$ is identified with a unital positive map $\alpha : \mathbb{C}^n \to M$, determined by $\alpha(\delta_j) := A_j$, $1 \le j \le n$. For $\psi, \varphi \in M_*^+$ we write $\mathcal{M}(\psi) := \psi \circ \alpha$, i.e., $\mathcal{M}(\psi) := (\psi(A_j))_{1 \le j \le n}$, and similarly for $\mathcal{M}(\varphi)$. Then the classical f-divergence $S_f(\mathcal{M}(\psi) \| \mathcal{M}(\varphi))$ is defined by

$$S_f(\mathcal{M}(\psi) \| \mathcal{M}(\varphi)) := \sum_{j=1}^n \varphi(A_j) f\left(\frac{\psi(A_j)}{\varphi(A_j)}\right)$$

under the convention (2.2). A measurement $\mathcal{E} = (E_j)_{1 \le j \le n}$ in M is said to be *projective* if E_j's are orthogonal projections with $\sum_{j=1}^n E_j = 1$.

Definition 5.1 For $\psi, \varphi \in M_*^+$ we define the *measured f-divergence* and its *projective* variant as follows:

$$S_f^{\text{meas}}(\psi \| \varphi) := \sup\{S_f(\mathcal{M}(\psi) \| \mathcal{M}(\varphi)) : \mathcal{M} \text{ a measurement in } M\}, \tag{5.1}$$

$$S_f^{\text{pr}}(\psi \| \varphi) := \sup\{S_f(\mathcal{E}(\psi) \| \mathcal{E}(\varphi)) : \mathcal{E} \text{ a projective measurement in } M\}. \tag{5.2}$$

Obviously, $S_f^{\text{pr}}(\psi \| \varphi) \le S_f^{\text{meas}}(\psi \| \varphi)$. Moreover, $S_f^{\text{meas}}(\psi \| \varphi) \le S_f(\psi \| \varphi)$ whenever f is operator convex on $(0, \infty)$, due to the monotonicity of S_f, see Theorem 2.7 (iv).

Proposition 5.2 *For every $\psi, \varphi \in M_*^+$,*

$$S_f^{\text{meas}}(\psi\|\varphi) = \sup\{S_f(\psi\circ\alpha\|\varphi\circ\alpha) :$$

$$\alpha : \mathcal{A} \to M \text{ a unital normal positive map,}$$

$$\mathcal{A} \text{ a commutative von Neumann algebra}\}. \qquad (5.3)$$

Proof Since $S_f^{\text{meas}}(\psi\|\varphi) \le$ the RHS of (5.3) is obvious, we may prove that

$$S_f^{\text{meas}}(\psi\|\varphi) \ge S_f(\psi\circ\alpha\|\varphi\circ\alpha) \qquad (5.4)$$

for every $\alpha : \mathfrak{A} \to M$ as in (5.3). Let $e_0 := s(\psi\circ\alpha) \vee s(\varphi\circ\alpha)$ and define $\beta : e_0\mathfrak{A}e_0 \to M$ by $\beta(x) := x + \psi(x)(1 - e_0)$ for $x \in e_0\mathfrak{A}e_0$. Then β is a unital normal positive map and it is obvious that $S_f(\psi\circ\alpha\|\varphi\circ\alpha) = S_f(\psi\circ\alpha\circ\beta\|\varphi\circ\alpha\circ\beta)$. Replacing α with $\alpha\circ\beta : e_0\mathfrak{A}e_0 \to M$, we may assume that $\mathfrak{A} = L^\infty(X,\mathcal{X},\mu)$, where $(X,\mathcal{X},\mu)$ is a finite measure space. Identifying $\psi\circ\alpha, \varphi\circ\alpha$ as functions $\hat\psi, \hat\varphi \in L^1(X,\mathcal{X},\mu)_+$ we have

$$S_f(\psi\circ\alpha\|\varphi\circ\alpha)$$
$$= \int_{\{\hat\psi>0\}\cap\{\hat\varphi>0\}} \hat\varphi f(\hat\psi/\hat\varphi)\,d\mu + f(0^+)\int_{\{\hat\psi=0\}}\hat\varphi\,d\mu + f'(+\infty)\int_{\{\hat\varphi=0\}}\hat\psi\,d\mu.$$

Hence, to prove (5.4), it suffices to prove that

$$\int_{X_0} f\left(\frac{dp}{dq}\right) dq \le \sup\left\{\sum_{j=1}^n q(X_j)f\left(\frac{p(X_j)}{q(X_j)}\right) :\right.$$

$$\left.\{X_1,\ldots,X_n\} \text{ a finite measurable partition of } X_0\right\}, \quad (5.5)$$

where $X_0 := \{\hat\psi > 0\} \cap \{\hat\varphi > 0\}$, $dp := \hat\psi\,d\mu$, $dq := \hat\varphi\,d\mu$, and $dp/dq = \hat\psi/\hat\varphi$, the Radon-Nikodym derivative.

To show (5.5), we may and do assume that $f \ge 0$. Indeed, there exist $a, b \in \mathbb{R}$ and a convex function $f_0 \ge 0$ on $(0,\infty)$ such that $f(t) = f_0(t) + at + b$, $t \in (0,\infty)$. Then

$$\int_{X_0} f\left(\frac{dp}{dq}\right) dq = \int_{X_0} f_0\left(\frac{dp}{dq}\right) dq + ap(X_0) + bq(X_0),$$

$$\sum_{j=1}^n q(X_j)f\left(\frac{p(X_j)}{q(X_j)}\right) = \sum_{j=1}^n q(X_j)f_0\left(\frac{p(X_j)}{q(X_j)}\right) + ap(X_0) + bq(X_0),$$

so that it suffices to show (5.5) for f_0. For every $n \in \mathbb{N}$ consider a finite measurable partition $\{X_1,\ldots,X_{n2^n+1}\}$ given by

$$X_j := \left\{x \in X : \frac{j-1}{2^n} \le \frac{dp}{dq}(x) < \frac{j}{2^n}\right\}, \qquad 1 \le j \le n2^n,$$

$$X_{n2^n+1} := \left\{x \in X : \frac{dp}{dq}(x) \ge n2^n\right\},$$

and define a simple function $\kappa_n := \sum_{j=1}^{n2^n} \lambda_j 1_{X_j}$, where $\lambda_j := \inf\{f(t) : \frac{j-1}{2^n} \le t \le \frac{j}{2^n}\}$ (with $f(0) = f(0^+)$), $1 \le j \le n2^n$, and $\lambda_{n2^n+1} := \inf\{f(t) : t \ge n2^n\}$. Then, since f is convex and so continuous on $(0, \infty)$, it is immediate to see that $\kappa_n(x) \to f(\frac{dp}{dq}(x))$ for q-a.e. $x \in X$. By Fatou's lemma we have

$$\int_{X_0} f\left(\frac{dp}{dq}\right) dq \le \liminf_{n\to\infty} \int_{X_0} \kappa_n \, dq = \liminf_{n\to\infty} \sum_{j=1}^{n2^n+1} \lambda_j q(X_j)$$

$$\le \liminf_{n\to\infty} \sum_{j=1}^{n2^n+1} f\left(\frac{p(X_j)}{q(X_j)}\right) q(X_j)$$

$$\le \text{the RHS of (5.4)},$$

where the second inequality above follows since

$$\frac{p(X_j)}{q(X_j)} = \frac{1}{q(X_j)} \int_{X_j} \frac{dp}{dq} \, dq \in \begin{cases} [\frac{j-1}{2^n}, \frac{j}{2^n}], & 1 \le j \le n2^n, \\ [n2^n, +\infty), & j = n2^n + 1. \end{cases}$$

Hence (5.4) is obtained. $\square$

Let H be a self-adjoint operator affiliated with M and

$$H = \int_{-\infty}^{\infty} \lambda \, dE_H(\lambda)$$

be the spectral decomposition of H with the spectral resolution $\{E_H(\lambda) : \lambda \in \mathbb{R}\}$, so $E_H(\lambda) \in M$ for all $\lambda \in \mathbb{R}$, see the beginning of Appendix A.5. Write $\mathcal{E}_H(\psi)$ for the finite positive measures $d\psi(E_H(\lambda))$ on the Borel space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Then similarly to the proof of Proposition 5.2 we have the following:

Proposition 5.3 *For every $\psi, \varphi \in M_*^+$,*

$S_f^{\mathrm{pr}}(\psi\|\varphi)$

$\quad = \sup\{S_f(\mathcal{E}_H(\psi)\|\mathcal{E}_H(\varphi)) : H$ *a self-adjoint operator affiliated with M*$\}$. (5.6)

In this way, S_f^{pr} is the supremum of the classical f-divergences over the von Neumann (projective) measurements, while S_f^{meas} is that over all measurements of quantum-classical channels.

The next proposition summarizes basic properties of S_f^{meas} and S_f^{pr}.

Proposition 5.4

(1) Joint lower semicontinuity: $S_f^{\mathrm{meas}}(\psi\|\varphi)$ *and* $S_f^{\mathrm{pr}}(\psi\|\varphi)$ *are jointly lower semi-continuous in* $(\psi, \varphi) \in M_*^+ \times M_*^+$ *in the $\sigma(M_*, M)$-topology.*

(2) Joint convexity: $S_f^{\mathrm{meas}}(\psi\|\varphi)$ *and* $S_f^{\mathrm{pr}}(\psi\|\varphi)$ *are jointly convex and jointly subadditive on* $M_*^+ \times M_*^+$.

(3) Monotonicity: *Let $\gamma : N \to M$ be a unital normal positive map between von Neumann algebras. For every $\psi, \varphi \in M_*^+$,*

$$S_f^{\text{meas}}(\psi \circ \gamma \| \varphi \circ \gamma) \leq S_f^{\text{meas}}(\psi \| \varphi).$$

(4) *For every $\psi, \varphi, \omega \in M_*^+$,*

$$S_f^{\text{meas}}(\psi \| \varphi) = \lim_{\varepsilon \searrow 0} S_f^{\text{meas}}(\psi + \varepsilon\omega \| \varphi + \varepsilon\omega),$$
$$S_f^{\text{pr}}(\psi \| \varphi) = \lim_{\varepsilon \searrow 0} S_f^{\text{pr}}(\psi + \varepsilon\omega \| \varphi + \varepsilon\omega).$$

Proof Note that the classical f-divergence $S_f(a\|b) = \sum_{j=1}^n b_j f(a_j/b_j)$ for $a = (a_j)_{j=1}^n$, $b = (b_j)_{j=1}^n \in [0, \infty)^n$ is jointly lower semicontinuous and jointly convex. Hence (1) and (2) immediately follow from Definition 5.1.

(3) Let $\gamma : N \to M$ be as stated. For any measurement $\mathcal{M} = (B_j)_{j=1}^n$ in N, since $\gamma(\mathcal{M}) := (\gamma(B_j))_{j=1}^n$ is a measurement in M, we have

$$S_f(\mathcal{M}(\psi \circ \gamma) \| \mathcal{M}(\varphi \circ \gamma)) = S_f(\gamma(\mathcal{M})(\psi) \| \gamma(\mathcal{M})(\varphi)) \leq S_f^{\text{meas}}(\psi \| \varphi),$$

which implies that $S_f^{\text{meas}}(\psi \circ \gamma \| \varphi \circ \gamma) \leq S_f(\psi \| \varphi)$.

(4) From (1) it follows that

$$S_f^{\text{meas}}(\psi \| \varphi) \leq \liminf_{\varepsilon \searrow 0} S_f^{\text{meas}}(\psi + \varepsilon\omega \| \varphi + \varepsilon\omega).$$

On the other hand, by (2) we have

$$\begin{aligned}
S_f^{\text{meas}}(\psi + \varepsilon\omega \| \varphi + \varepsilon\omega) &\leq S_f^{\text{meas}}(\psi \| \varphi) + S_f^{\text{meas}}(\varepsilon\omega \| \varepsilon\omega) \\
&= S_f^{\text{meas}}(\psi \| \varphi) + \varepsilon\omega(1)f(1).
\end{aligned}$$

Hence the assertion for S_f^{meas} follows. The proof is similar for S_f^{pr}. $\square$

This is a good place for us to explain an abstract approach to quantum f-divergences. Assume that f is operator convex on $(0, \infty)$. We say that a function $S_f^q : M_*^+ \times M_*^+ \to (-\infty, \infty]$, where M varies over all von Neumann algebras, is a *monotone quantum f-divergence* if the following two properties are satisfied:

(a) $S_f^q(\psi \circ \gamma \| \varphi \circ \gamma) \leq S_f^q(\psi \| \varphi)$ for any unital normal CP map $\gamma : N \to M$ between von Neumann algebras and for every $\psi, \varphi \in M_*^+$.
(b) When M is an abelian von Neumann algebra with $M = L^\infty(X, \mu)$ on a σ-finite measure space $(X, \mathcal{X}\mu)$, $S_f^q(\psi \| \varphi)$ coincides with the classical f-divergence of $\psi, \varphi \in L^1(X, \mu)_+$ as in Example 2.2 (1).

If S_f^q is a monotone quantum f-divergence, then it is readily seen from Definition 5.1 and Theorem 4.17 that for every $\psi, \varphi \in M_*^+$,

$$S_f^{\text{meas}}(\psi \| \varphi) \leq S_f^q(\psi \| \varphi) \leq \widehat{S}_f(\psi \| \varphi),$$

which justifies the name maximal f-divergence for $\widehat{S}_f$ and suggests us to call S_f^{meas} the minimal f-divergence. Since the standard f-divergence S_f is typical among monotone quantum f-divergences, we have the following:

Proposition 5.5 *For every* $\psi, \varphi \in M_*^+$,

$$S_f^{\mathrm{meas}}(\psi\|\varphi) \leq S_f(\psi\|\varphi) \leq \widehat{S}_f(\psi\|\varphi). \tag{5.7}$$

5.2 Variational expressions

We consider the following classes of functions on $(0, \infty)$:

(a) $\mathcal{F}_{\mathrm{cv}}^{\nearrow}(0, \infty)$ is the set of non-decreasing convex real functions f on $(0, \infty)$ such that $f'(\infty) = +\infty$,
(b) $\mathcal{F}_{\mathrm{cc}}^{\nearrow}(0, \infty)$ is the set of non-decreasing concave real functions g on $(0, \infty)$ such that $g'(\infty) = 0$.

The following variational formulas were shown in [46, Appendix A], which provides a theory of conjugate functions (or the Legendre transform) on the half-line $(0, \infty)$.

Lemma 5.6 (1) *For every* $f \in \mathcal{F}_{\mathrm{cv}}^{\nearrow}(0, \infty)$ *define*

$$f^{\sharp}(t) := \sup_{s>0}\{st - f(s)\}, \qquad s \in (0, \infty).$$

Then $f^{\sharp} \in \mathcal{F}_{\mathrm{cv}}^{\nearrow}(0, \infty)$ *and* $f \mapsto f^{\sharp}$ *is an involutive bijection on* $\mathcal{F}_{\mathrm{cv}}^{\nearrow}(0, \infty)$, *i.e., a bijection such that* $f^{\sharp\sharp} = f$ *for all* $f \in \mathcal{F}_{\mathrm{cv}}^{\nearrow}(0, \infty)$. *Therefore,*

$$f(s) = \sup_{t>0}\{st - f^{\sharp}(t)\}, \qquad t \in (0, \infty),$$

and moreover $f^{\sharp}(0^+) = -f(0^+)$ $(\in \mathbb{R})$.
(2) *For every* $g \in \mathcal{F}_{\mathrm{cc}}^{\nearrow}(0, \infty)$ *define*

$$g^{\flat}(t) := \inf_{s>0}\{st - g(s)\}, \qquad s \in (0, \infty).$$

Then $g^{\flat} \in \mathcal{F}_{\mathrm{cc}}^{\nearrow}(0, \infty)$ *and* $g \mapsto g^{\flat}$ *is an involutive bijection on* $\mathcal{F}_{\mathrm{cc}}^{\nearrow}(0, \infty)$, *i.e., a bijection such that* $g^{\flat\flat} = g$ *for all* $g \in \mathcal{F}_{\mathrm{cc}}^{\nearrow}(0, \infty)$. *Therefore,*

$$g(s) := \inf_{t>0}\{st - g^{\flat}(t)\}, \qquad t \in (0, \infty),$$

and moreover $g^{\flat}(0^+) = -g(\infty)$ *and* $g(0^+) = -g^{\flat}(\infty)$.

For a, b with $-\infty \leq a < b \leq \infty$, we write $M_{(a,b)}$ for the set of self-adjoint operators $A \in M$ whose spectrum is included in (a, b). In particular, $M_{(0,\infty)}$ is the the set M_{++} of positive invertible operators in M, and $M_{(-\infty,\infty)}$ is the set M_{sa} of

self-adjoint operators in M. For $-\infty < a < b < \infty$ we also write $M_{[a,b]}$ for the set of self-adjoint operators $A \in M$ whose spectrum is in $[a,b]$, i.e., $a1 \leq A \leq b1$.

By Lemma 5.6 we can show variational expressions of S_f^{pr} in the next theorem.

Theorem 5.7 *Let* $\psi, \varphi \in M_*^+$.

(1) *Assume that* $f \in \mathcal{F}_{\mathrm{cv}}^{\nearrow}(0, \infty)$. *Let* $a := f'(0^+) \in [0, \infty)$ *and* $b := -f(0^+) \in (-\infty, \infty)$. *Then the inverse* $(f^{\sharp})^{-1} : (b, \infty) \to (a, \infty)$ *exists and*

$$S_f^{\mathrm{pr}}(\psi \| \varphi) = \sup_{A \in M_{++}} \{\psi(A) - \varphi(f^{\sharp}(A))\} \tag{5.8}$$

$$= \sup_{A \in M_{(a,\infty)}} \{\psi(A) - \varphi(f^{\sharp}(A))\} \tag{5.9}$$

$$= \sup_{A \in M_{(b,\infty)}} \{\psi((f^{\sharp})^{-1}(A)) - \varphi(A)\}. \tag{5.10}$$

(2) *Assume that* $g \in \mathcal{F}_{\mathrm{cc}}^{\nearrow}(0, \infty)$ *and* g *is non-constant. Let* $f := -g$, *and* $a := g'(0^+)$, $b := -g(0^+)$, $c := -g(\infty)$. *Then* $a \in (0, \infty]$, $-\infty \leq c < b \leq \infty$, *the inverse* $(g^{\flat})^{-1} : (c, b) \to (0, a)$ *exists, and*

$$S_f^{\mathrm{pr}}(\psi \| \varphi) = \sup_{A \in M_{++}} \{-\psi(A) + \varphi(g^{\flat}(A))\} \tag{5.11}$$

$$= \sup_{A \in M_{(0,a)}} \{-\psi(A) + \varphi(g^{\flat}(A))\} \tag{5.12}$$

$$= \sup_{A \in M_{(c,b)}} \{-\psi((g^{\flat})^{-1}(A)) + \varphi(A)\}. \tag{5.13}$$

Proof (1) Let $\mathcal{E} = (E_j)_{1 \leq j \leq n}$ be a projective measurement in M. For each j let us prove that

$$\varphi(E_j) f\left(\frac{\psi(E_j)}{\varphi(E_j)}\right) = \sup_{t > 0} \{t\psi(E_j) - f^{\sharp}(t)\varphi(E_j)\}. \tag{5.14}$$

When $\psi(E_j) > 0$ and $\varphi(E_j) > 0$, by Lemma 5.6 (1) the LHS of (5.14) is

$$\varphi(E_j) \sup_{t > 0} \left\{t\frac{\psi(E_j)}{\varphi(E_j)} - f^{\sharp}(t)\right\} = \sup_{t > 0} \{t\psi(E_j) - f^{\sharp}(t)\varphi(E_j)\}.$$

When $\psi(E_j) = 0$ and $\varphi(E_j) > 0$, the LHS of (5.14) is

$$f(0^+)\varphi(E_j) = -f^{\sharp}(0^+)\varphi(E_j) = \sup_{t > 0} \{-f^{\sharp}(t)\varphi(E_j)\}$$

by the convention (2.2) and $f(0^+) = -f^{\sharp}(0^+)$ stated in Lemma 5.6 (1). When $\psi(E_j) > 0$ and $\varphi(E_j) = 0$, the LHS of (5.14) is

$$f'(\infty)\psi(E_j) = +\infty = \sup_{t > 0} \{t\psi(E_j)\}.$$

When $\psi(E_j) = \varphi(E_j) = 0$, both sides of (5.14) are 0. Hence (5.14) holds in all the cases.

Summing up (5.14) for all j we have

$$\sum_{j=1}^{n} \varphi(E_j) f\left(\frac{\psi(E_j)}{\varphi(E_j)}\right) = \sup_{(t_j):t_j>0} \sum_{j=1}^{n} \{t_j \psi(E_j) - f^{\sharp}(t_j)\varphi(E_j)\}$$

$$= \sup_{(t_j):t_j>0} \left\{\psi\left(\sum_j t_j E_j\right) - \varphi\left(\sum_j f^{\sharp}(t_j)E_j\right)\right\}. \qquad (5.15)$$

Therefore, we have

$$S_f^{\mathrm{pr}}(\psi\|\varphi) = \sup_{(E_j)} \sup_{(t_j)} \left\{\psi\left(\sum_j t_j E_j\right) - \varphi\left(f^{\sharp}\left(\sum_j t_j E_j\right)\right)\right\}$$

$$= \sup_{A \in M_{++}} \{\psi(A) - \varphi(f^{\sharp}(A))\},$$

since $\psi(A) - \varphi(f^{\sharp}(A))$ for any $A \in M_{++}$ is approximated by $\psi(A_0) - \varphi(f^{\sharp}(A_0))$ for discrete $A_0 := \sum_{j=1}^{n} t_j E_j$. Hence (5.8) follows.

Since $f^{\sharp}(t) = -f(0^+)$ for $0 < t \leq a$, we can restrict t_j to $t_j > a$ in the supremum in (5.15), so the supremum in (5.8) is the same as that taken over $A \in M_{(a,\infty)}$. Hence (5.9) follows. By Lemma 5.6 (1) note that $f^{\sharp} \in \mathcal{F}_{\mathrm{cv}}^{\nearrow}(0,\infty)$ and $f^{\sharp}(0^+) = -f(0^+) = b$. We easily find that $f^{\sharp}(t) = b$ for $0 < t \leq a$ and $f^{\sharp}$ is strictly increasing from (a,∞) onto (b,∞). Hence the inverse $(f^{\sharp})^{-1} : (b,\infty) \to (a,\infty)$ exists. Since $X \mapsto (f^{\sharp})^{-1}(X)$ is a bijective map from $M_{(b,\infty)}$ onto $M_{(a,\infty)}$, we have (5.10) as well.

(2) Let $\mathcal{E} = (E_j)_{1 \leq j \leq n}$ be a projective measurement in M. For each j we prove that

$$\varphi(E_j) f\left(\frac{\psi(E_j)}{\varphi(E_j)}\right) = \sup_{t>0} \{-t\psi(E_j) + g^{\flat}(t)\varphi(E_j)\}. \qquad (5.16)$$

When $\psi(E_j) > 0$ and $\varphi(E_j) > 0$, by Lemma 5.6 (2) the LHS of (5.16) is

$$-\varphi(E_j) g\left(\frac{\psi(E_j)}{\varphi(E_j)}\right) = -\varphi(E_j) \inf_{t>0}\left\{\frac{\psi(E_j)}{\varphi(E_j)} t - g^{\flat}(t)\right\}$$

$$= \sup_{t>0} \{-t\psi(E_j) + g^{\flat}(t)\varphi(E_j)\}.$$

When $\psi(E_j) = 0$ and $\varphi(E_j) > 0$, the LHS of (5.16) is

$$-g(0^+)\varphi(E_j) = g^{\flat}(+\infty)\varphi(E_j) = \sup_{t>0}\{g^{\flat}(t)\varphi(E_j)\}$$

from (2.2) and $g(0^+) = -g^{\flat}(\infty)$ stated in Lemma 5.6 (2). When $\psi(E_j) > 0$ and $\varphi(E_j) = 0$, the LHS of (5.16) is

$$-g'(\infty)\psi(E_j) = 0 = \sup_{t>0}\{-t\psi(E_j)\}.$$

When $\psi(E_j) = \varphi(E_j) = 0$, both sides of (5.16) are 0. Hence (5.16) has been shown. Since g is non-constant, $a \in (0, \infty]$ and $-\infty \le c < b \le \infty$ are obvious. Moreover, when $a < \infty$, since $-b = g(0^+) > -\infty$ and $g^\flat(t) = -b$ for $t \ge a$, we find that the supremum in (5.16) is the same as that taken over $t \in (0, a)$. Thus, (5.11) and (5.12) follow as in the proof of (1). Moreover, it is easily verified that $g^\flat$ is strictly increasing from $(0, a)$ onto (c, b). Hence the inverse $(g^\flat)^{-1} : (c, b) \to (0, a)$ exists. Since $X \mapsto (g^\flat)^{-1}(X)$ is a bijective map from $M_{(c,b)}$ onto $M_{(0,a)}$, we have (5.13) as well.   □

Based on Theorem 5.7 we can show the equality $S_f^{\mathrm{pr}} = S_f^{\mathrm{meas}}$ in the following cases.

Theorem 5.8 *Let f be a convex function on $(0, \infty)$, and $\widetilde{f}$ be the transpose of f, i.e., $\widetilde{f}(t) := tf(t^{-1})$, $t > 0$. Let $g := -f$ and $\widetilde{g} := -\widetilde{f}$ (i.e., $\widetilde{g}$ is the transpose of g). Assume that one of the following conditions (1)–(4) is satisfied:*

(1) *$f \in \mathcal{F}_{\mathrm{cv}}^{\nearrow}(0, \infty)$ and either $f^\sharp$ is operator convex on $(f'(0^+), \infty)$ or $(f^\sharp)^{-1}$ is operator concave on $(-f(0^+), \infty)$,*
(2) *$\widetilde{f} \in \mathcal{F}_{\mathrm{cv}}^{\nearrow}(0, \infty)$ and either $\widetilde{f}^{\,\sharp}$ is operator convex on $(\widetilde{f}'(0^+), \infty)$ or $(\widetilde{f}^{\,\sharp})^{-1}$ is operator concave on $(-\widetilde{f}(0^+), \infty)$,*
(3) *$g \in \mathcal{F}_{\mathrm{cc}}^{\nearrow}(0, \infty)$ and either $g^\flat$ is operator concave on $(0, g'(0^+))$ or $(g^\flat)^{-1}$ is operator convex on $(-g(\infty), -g(0^+))$,*
(4) *$\widetilde{g} \in \mathcal{F}_{\mathrm{cc}}^{\nearrow}(0, \infty)$ and either $\widetilde{g}^{\,\flat}$ is operator concave on $(0, \widetilde{g}'(0^+))$ or $(\widetilde{g}^{\,\flat})^{-1}$ is operator convex on $(-\widetilde{g}(\infty), -\widetilde{g}(0^+))$.*

Then for every $\psi, \varphi \in M_*^+$,

$$S_f^{\mathrm{pr}}(\psi\|\varphi) = S_f^{\mathrm{meas}}(\psi\|\varphi).$$

Proof Since $S_f^{\mathrm{pr}}(\psi\|\varphi) \le S_f^{\mathrm{meas}}(\psi\|\varphi)$ holds trivially, we may prove the converse inequality.

Case (1). Assume that $f \in \mathcal{F}_{\mathrm{cv}}^{\nearrow}(0, \infty)$ and $f^\sharp$ is operator convex on $(f'(0^+), \infty)$. For any measurement $(A_j)_{1 \le j \le n}$ in M, as in the proof of Theorem 5.7 (1) with replacing $\psi(A_j), \varphi(A_j)$ in place of $\psi(E_j), \varphi(E_j)$, we have, with $a := f'(0^+)$,

$$\sum_j \varphi(A_j)f\left(\frac{\psi(A_j)}{\varphi(A_j)}\right) = \sup_{(t_j):t_j>a}\left\{\psi\left(\sum_j t_j A_j\right) - \varphi\left(\sum_j f^\sharp(t_j)A_j\right)\right\}$$

$$\le \sup_{(t_j):t_j>a}\left\{\psi\left(\sum_j t_j A_j\right) - \varphi\left(f^\sharp\left(\sum_j t_j A_j\right)\right)\right\}$$

$$= \sup_{A \in M_{(a,+\infty)}}\{\psi(A) - \varphi(f^\sharp(A))\} = S_f^{\mathrm{pr}}(\psi\|\varphi).$$

In the above, the last equality holds by Theorem 5.7 (1), and the inequality is seen from the operator convexity of $f^\sharp$ on (a, ∞) as follows:

$$f^\sharp\left(\sum_j t_j A_j\right) = f^\sharp\left(\sum_j A_j^{1/2} t_j A_j^{1/2}\right) \le \sum_j A_j^{1/2} f^\sharp(t_j) A_j^{1/2} = \sum_j f^\sharp(t_j) A_j$$

by [22, Theorem 2.1] (also [45, Theorem 2.5.7]). Therefore, we have $S_f^{\mathrm{meas}}(\psi\|\varphi) \le S_f^{\mathrm{pr}}(\psi\|\varphi)$.

Next, assume that $(f^\sharp)^{-1}$ is operator concave on $(-f(0^+), \infty)$. We have, with $b := -f(0^+)$,

$$
\begin{aligned}
\sum_j \varphi(A_j) f\left(\frac{\psi(A_j)}{\varphi(A_j)}\right) &= \sup_{(t_j): t_j > a}\left\{\psi\left(\sum_j t_j A_j\right) - \varphi\left(\sum_j f^\sharp(t_j) A_j\right)\right\} \\
&= \sup_{(s_j): s_j > b}\left\{\psi\left(\sum_j (f^\sharp)^{-1}(s_j) A_j\right) - \varphi\left(\sum_j s_j A_j\right)\right\} \\
&\le \sup_{(s_j): s_j > b}\left\{\psi\left((f^\sharp)^{-1}\left(\sum_j s_j A_j\right)\right) - \varphi\left(\sum_j s_j A_j\right)\right\} \\
&= \sup_{A \in M_{(b,+\infty)}}\{\psi(f^\sharp)^{-1}(A) - \varphi A\} = S_f^{\mathrm{pr}}(\psi\|\varphi).
\end{aligned}
$$

In the above, the last equality holds by Theorem 5.7 (1), and the inequality follows from the operator concavity of $(f^\sharp)^{-1}$ on (b, ∞).

Case (2). We readily find that

$$S_f^{\mathrm{pr}}(\psi\|\varphi) = S_{\widetilde{f}}^{\mathrm{pr}}(\varphi\|\psi), \qquad S_f^{\mathrm{meas}}(\psi\|\varphi) = S_{\widetilde{f}}^{\mathrm{meas}}(\varphi\|\psi).$$

Hence the result is immediate from case (1).

Case (3). Assume that $g \in \mathcal{F}_{\mathrm{cc}}^{\nearrow}(0, \infty)$. Since the assertion is obvious when g is a constant function, we may assume that g is non-constant; then $g'(0^+) > 0$ and $g(0^+) < g(\infty)$. Assume that $g^\flat$ is operator concave on $(0, g'(0^+))$. For any measurement $(A_j)_{1 \le j \le n}$ in M, as in the proof of Theorem 5.7 (2), we have, with $a := g'(0^+)$,

$$
\begin{aligned}
\sum_j \varphi(A_j) f\left(\frac{\psi(A_j)}{\varphi(A_j)}\right) &= \sup_{(t_j): 0 < t_j < a}\left\{-\psi\left(\sum_j t_j A_j\right) + \varphi\left(\sum_j g^\flat(t_j) A_j\right)\right\} \\
&\le \sup_{(t_j): 0 < t_j < a}\left\{-\psi\left(\sum_j t_j A_j\right) + \varphi\left(g^\flat\left(\sum_j t_j A_j\right)\right)\right\} \\
&= \sup_{A \in M_{(0,a)}}\{-\psi(A) + \varphi(g^\flat(A))\} = S_f^{\mathrm{pr}}(\psi\|\varphi).
\end{aligned}
$$

In the above, the last equality is due to Theorem 5.7 (2), and the inequality follows from the operator concavity of $g^\flat$ on $(0, a)$. When $(g^\flat)^{-1}$ is operator convex on $(-g(\infty), -g(0^+))$, the proof is similar to the second part of the proof of case (1).

Case (4). Immediate from case (3) as in case (2). $\qquad\square$

Example 5.9 (Rényi divergence) Consider

$$f_\alpha(t) := t^\alpha \in \mathcal{F}_{\mathrm{cv}}^{\nearrow}(0,\infty) \quad \text{for } \alpha \in (1,\infty),$$

$$g_\alpha(t) := t^\alpha \in \mathcal{F}_{\mathrm{cc}}^{\nearrow}(0,\infty). \quad f_\alpha := -g_\alpha \quad \text{for } \alpha \in (0,1).$$

(1) Assume that $\alpha \in (1,\infty)$. We have $f_\alpha(0^+) = f_\alpha'(0^+) = 0$ and

$$f_\alpha^\sharp(t) = (\alpha - 1)\left(\frac{t}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}, \quad (f_\alpha^\sharp)^{-1}(t) = \alpha\left(\frac{t}{\alpha-1}\right)^{\frac{\alpha-1}{\alpha}}, \qquad t \in (0,\infty).$$

Hence $(f_\alpha^\sharp)^{-1}$ is operator concave on $(0,\infty)$. By Theorem 5.7 (1) and case (1) of Theorem 5.8, for every $\psi, \varphi \in M_*^+$ we have

$$\begin{aligned}
S_{f_\alpha}^{\mathrm{pr}}(\psi\|\varphi) = S_{f_\alpha}^{\mathrm{meas}}(\psi\|\varphi) &= \sup_{A \in M_{++}} \left\{ \psi(A) - (\alpha-1)\varphi\left(\left(\frac{A}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}\right) \right\} \\
&= \sup_{A \in M_{++}} \left\{ \alpha\psi\left(\left(\frac{A}{\alpha-1}\right)^{\frac{\alpha-1}{\alpha}}\right) - \varphi(A) \right\} \\
&= \sup_{A \in M_{++}} \{ \alpha\psi(A^{\frac{\alpha-1}{\alpha}}) - (\alpha-1)\varphi(A) \}. \qquad (5.17)
\end{aligned}$$

(2) Assume that $\alpha \in (0,1)$. We have $g_\alpha(0^+) = 0$, $g_\alpha(\infty) = \infty$, $g_\alpha'(0^+) = \infty$, and

$$g_\alpha^\flat(t) = (\alpha - 1)\left(\frac{t}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}, \quad (g_\alpha^\flat)^{-1}(-t) = \alpha\left(\frac{t}{1-\alpha}\right)^{\frac{\alpha-1}{\alpha}}, \qquad t \in (0,\infty).$$

Note that $\widetilde{g}_\alpha(t) = t^{1-\alpha} \in \mathcal{F}_{\mathrm{cc}}^{\nearrow}(0,\infty)$ with $\widetilde{g}_\alpha(0^+) = 0$, $\widetilde{g}_\alpha(\infty) = \infty$ and $\widetilde{g}_\alpha'(0^+) = \infty$. When $\alpha \in [1/2, 1)$, $\widetilde{g}_\alpha^\flat(t) = -\alpha\left(\frac{t}{1-\alpha}\right)^{\frac{\alpha-1}{\alpha}}$ is operator concave on $(0,\infty)$. On the other hand, when $\alpha \in (0, 1/2]$, $(\widetilde{g}_\alpha^\flat)^{-1}(-t) = (1-\alpha)\left(\frac{t}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}$ is operator convex on $(0,\infty)$. Hence by Theorem 5.7 (2) and case (4) of Theorem 5.8, we have

$$\begin{aligned}
S_{f_\alpha}^{\mathrm{pr}}(\psi\|\varphi) = S_{f_\alpha}^{\mathrm{meas}}(\psi\|\varphi) &= \sup_{A \in M_{++}} \left\{ -\psi(A) - (1-\alpha)\varphi\left(\left(\frac{A}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}\right) \right\} \\
&= \sup_{A \in M_{++}} \left\{ -\alpha\psi\left(\left(\frac{A}{1-\alpha}\right)^{\frac{\alpha-1}{\alpha}}\right) - \varphi(A) \right\} \\
&= -\inf_{A \in M_{++}} \{ \alpha\psi(A^{\frac{\alpha-1}{\alpha}}) + (1-\alpha)\varphi(A) \}. \qquad (5.18)
\end{aligned}$$

The variational formulas in (5.17) and (5.18) for $S_{f_\alpha}^{\mathrm{pr}}$ were given in [16, Lemma 3, (20)] in the finite-dimensional setting.

(3) For each $\alpha \in (0,\infty) \setminus \{1\}$ and $\psi, \varphi \in M_*^+$ with $\psi \neq 0$, apart from the α-Rényi divergence $D_\alpha(\psi\|\varphi)$ (Sec. 3.1) and the sandwiched α-Rényi divergence $\widetilde{D}_\alpha(\psi\|\varphi)$ (Sec. 3.3), we define the *measured α-Rényi divergence* by

$$D_\alpha^{\mathrm{meas}}(\psi\|\varphi) := \sup\{ D_\alpha(\mathcal{M}(\psi)\|\mathcal{M}(\varphi)) : \mathcal{M} \text{ a measurement in } M \}.$$

Alternatively,

$$D_\alpha^{\mathrm{meas}}(\psi\|\varphi) = \frac{1}{\alpha-1}\log\frac{Q_\alpha^{\mathrm{meas}}(\psi\|\varphi)}{\psi(1)}$$

with

$$Q_\alpha^{\mathrm{meas}}(\psi\|\varphi) := \begin{cases} S_{f_\alpha}^{\mathrm{meas}}(\psi\|\varphi) & \text{if } \alpha \in (1,\infty), \\ -S_{f_\alpha}^{\mathrm{meas}}(\psi\|\varphi) & \text{if } \alpha \in (0,1). \end{cases}$$

The projective version $D_\alpha^{\mathrm{pr}}(\psi\|\varphi)$ is similarly defined (see Definition 5.1). But from (5.17) and (5.18) above we find that for all $\alpha \in (0,\infty) \setminus \{1\}$ and all $\psi,\varphi \in M_*^+$,

$$D_\alpha^{\mathrm{pr}}(\psi\|\varphi) = D_\alpha^{\mathrm{meas}}(\psi\|\varphi),$$

which was given in [16, Theorem 4] in the finite-dimensional setting. By Remark 3.17 note that the (projectively) measured versions of $\widetilde{D}_\alpha$ are the same as D_α^{meas} and D_α^{pr}.

By Proposition 5.4(1), $D_\alpha^{\mathrm{pr}} = D_\alpha^{\mathrm{meas}}$ is jointly lower semicontinuous in the $\sigma(M_*, M)$-topology for all $\alpha \in (0,\infty) \setminus \{1\}$, while the same lower semicontinuity is unknown for D_α and $\widetilde{D}_\alpha$, see Theorems 3.2(4) and 3.15(3). When $0 < \alpha \le 1/2$, $D_\alpha^{\mathrm{meas}} \le D_\alpha$ follows from Theorem 3.2(7). When $\alpha \in [1/2,\infty) \setminus \{1\}$, $D_\alpha^{\mathrm{meas}} \le \widetilde{D}_\alpha$ follows from Theorem 3.15(4). Combining this with the above $D_\alpha^{\mathrm{pr}} = D_\alpha^{\mathrm{meas}}$ and Theorem 3.15(6) we have

$$D_\alpha^{\mathrm{pr}}(\psi\|\varphi) = D_\alpha^{\mathrm{meas}}(\psi\|\varphi) \le \widetilde{D}_\alpha(\psi\|\varphi) \le D_\alpha(\psi\|\varphi) \qquad (5.19)$$

for all $\psi,\varphi \in M_*^+$ and all $\alpha \in [1/2,\infty) \setminus \{1\}$.

Example 5.10 (Relative entropy) Consider

$$g_0(t) := \log t \in \mathcal{F}_{\mathrm{cc}}^{\nearrow}(0,\infty), \quad f_0 := -g_0.$$

Recall the relative entropy $D(\psi\|\varphi) = S_{f_0}(\varphi\|\psi) = S_{\widetilde{f_0}}(\psi\|\varphi)$ for $\psi,\varphi \in M_*^+$. We define the (projectively) *measured relative entropies* $D^{\mathrm{meas}}(\psi\|\varphi)$ and $D^{\mathrm{pr}}(\psi\|\varphi)$ by Definition 5.1. We have $g_0(0^+) = -\infty$, $g_0(+\infty) = \infty$, $g_0'(0^+) = \infty$, and

$$g_0^\flat(t) = 1 + \log t, \quad t \in (0,\infty); \qquad (g_0^\flat)^{-1}(t) = e^{t-1}, \quad t \in (-\infty,\infty).$$

Since $g_0^\flat$ is operator concave on $(0,+\infty)$, by Theorem 5.7(2) and case (3) of Theorem 5.8 we have

$$D^{\mathrm{pr}}(\psi\|\varphi) = D^{\mathrm{meas}}(\psi\|\varphi) = \sup_{A \in M_{++}} \{-\varphi(A) + \psi(1 + \log A)\}$$
$$= \sup_{H \in M_{\mathrm{sa}}} \{\psi(H) + \psi(1) - \varphi(e^H)\}. \qquad (5.20)$$

The above first variational formula and $D^{\mathrm{pr}} = D^{\mathrm{meas}}$ were given in [16, Lemma 1 and Theorem 2] in the finite-dimensional setting.

It is worthwhile to compare (5.20) with other variational expressions for the relative entropy. For simplicity assume that $\psi,\varphi \in M_*^+$ are such that ψ is a state

and φ is faithful. For each $H \in M_{\text{sa}}$ let φ^H be the *perturbed functional* (a faithful functional in M_*^+) introduced in [5, 30]. Then the variational expression of $D(\psi \| \varphi)$ due to Petz [94] is

$$D(\psi \| \varphi) = \sup_{H \in M_{\text{sa}}} \{\psi(H) - \log \varphi^H(1)\}. \tag{5.21}$$

From Araki's *Golden-Thompson inequality* [6] we have

$$\varphi^H(1) \le \varphi(e^H). \tag{5.22}$$

(Indeed, in the finite-dimensional case, when φ has the density matrix e^{H_0}, the density of the perturbed φ^H is e^{H_0+H}, so (5.22) reduces to the usual Golden-Thompson inequality $\operatorname{Tr} e^{H_0+H} \le \operatorname{Tr}(e^{H_0}e^H)$.) Since

$$\log \varphi^H(1) \le \log \varphi(e^H) \le \varphi(e^H) - 1,$$

we conclude by (5.20)–(5.22) that

$$\begin{aligned}
D(\psi \| \varphi) &\ge \sup_{H \in M_{\text{sa}}} \{\psi(H) - \log \varphi(e^H)\} \\
&\ge \sup_{H \in M_{\text{sa}}} \{\psi(H) + 1 - \varphi(e^H)\} = D^{\text{pr}}(\psi \| \varphi).
\end{aligned}$$

Let E_H be the spectral resolution of $H \in M_{\text{sa}}$, and $dp(\lambda) := d\psi(E_H(\lambda))$, $dq(\lambda) := d\varphi(E_H(\lambda))$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ supported on $[-\|H\|, \|H\|]$, where p is a probability measure. By Proposition 5.3 we have

$$\begin{aligned}
D^{\text{pr}}(\psi \| \varphi) &\ge D(p \| q) \\
&\ge \int_{-\|H\|}^{\|H\|} \lambda \, dp(\lambda) - \log \int_{-\|H\|}^{\|H\|} e^\lambda \, dq(\lambda) \\
&= \psi(H) - \log \varphi(e^H),
\end{aligned}$$

where the above latter inequality is well-known for the classical relative entropy. Therefore, in addition to (5.20) we observe that

$$D^{\text{pr}}(\psi \| \varphi) = \sup_{H \in M_{\text{sa}}} \{\psi(H) - \log \varphi(e^H)\},$$

which was shown in the finite-dimensional case in [56, (1.7)] and [88, Proposition 7.13] but does not seem explicitly given so far in the von Neumann algebra setting.

Example 5.11 Here is one more example. Consider

$$g_\lambda(t) := \frac{t}{t + \lambda} \in \mathcal{F}_{\text{cc}}^{\nearrow}(0, \infty), \quad f_\lambda := -g_\lambda, \quad \text{where } 0 < \lambda < \infty.$$

We have $g_\lambda(0^+) = 0$, $g_\lambda(\infty) = 1$, $g_\lambda'(0^+) = 1/\lambda$, and

$$g_\lambda^\flat(t) = -\{1 - (\lambda t)^{1/2}\}^2, \qquad 0 < t \le 1/\lambda,$$

so that

$$(g_\lambda^\flat)^{-1}(t) = \frac{1}{\lambda}\{1 - (-t)^{1/2}\}^2, \qquad -1 < t < 0,$$

which is operator convex on $(-1, 0)$. Hence by Theorem 5.7 (2) and case (4) of Theorem 5.8 we have

$$S_{f_\lambda}^{\mathrm{pr}}(\psi\|\varphi) = S_{f_\lambda}^{\mathrm{meas}}(\psi\|\varphi) = \sup_{A \in M_{(0,1/\lambda)}} \left\{-\psi(A) - \varphi(\{1 - (\lambda A)^{1/2}\}^2)\right\}$$

$$= \sup_{A \in M_{(0,1)}} \left\{-\frac{1}{\lambda}\psi(\{1 - A^{1/2}\}^2) - \varphi(A)\right\}$$

$$= \sup\{-\psi(X^2) - \varphi(Y^2) : X, Y \in M_+, \lambda^{1/2}X + Y = 1\}.$$

5.3 Optimal Measurements

In the finite-dimensional case, it is known [52, Proposition 4.17] that the supremum in (5.1) is attained, and that if f is operator convex then the supremum in (5.2) is attained as well. This problem of attainability of (projective) measurements for (5.3) and (5.6) does not seem easy in the infinite-dimensional case. In this section we consider the attainability problem or the existence of optimal measurements, based on the variational expressions obtained in the previous sections.

The next lemma extending [49, Lemma 5.4] will be useful.

Lemma 5.12 *Let $-\infty < a < b < \infty$ and T, T_n ($n \in \mathbb{N}$) be self-adjoint bounded operators on a Hilbert space $\mathcal{H}$ such that $a\mathbb{1} \le T_n \le b\mathbb{1}$ for all $n \in \mathbb{N}$ and $T_n \to T$ in WOT (the weak operator topology).*

(1) *Assume that f is an operator convex function on $[a, b]$. Then any $\xi \in \mathcal{H}$,*

$$\langle \xi, f(T)\xi \rangle \le \liminf_{n\to\infty} \langle \xi, f(T_n)\xi \rangle.$$

Hence, the function $T \mapsto \langle \xi, f(T)\xi \rangle$ is lower semicontinuous in WOT on $B(\mathcal{H})_{[a,b]}$.

(2) *Assume that f is an operator concave function on $[a, b]$. Then for any $\xi \in \mathcal{H}$,*

$$\langle \xi, f(T)\xi \rangle \ge \limsup_{n\to\infty} \langle \xi, f(T_n)\xi \rangle.$$

Hence, the function $T \mapsto \langle \xi, f(T)\xi \rangle$ is upper semicontinuous in WOT on $B(\mathcal{H})_{[a,b]}$.

Proof The proofs of the two assertions are similar, so we prove (2) only. Let E_λ be a net of finite-dimensional orthogonal projections on $\mathcal{H}$ such that $E_\lambda \nearrow \mathbb{1}$. Set $T_{n,\lambda} := E_\lambda T_n E_\lambda + a(I - E_\lambda)$ and $T_\lambda := E_\lambda T E_\lambda + a(I - E_\lambda)$ for every $n \in \mathbb{N}$ and λ. For

each λ, since $\|T_{n,\lambda} - T_\lambda\| \to 0$ as $n \to \infty$, it follows that that $\|f(T_{n,\lambda}) - f(T_\lambda)\| \to 0$ as $n \to \infty$. By [22, Theorem 2.1] (also [45, Theorem 2.5.7]) we have

$$f(T_{n,\lambda}) \geq E_\lambda f(T_n) E_\lambda + f(a)(I - E_\lambda)$$

so that

$$\begin{aligned}
\langle \xi, f(T_\lambda)\xi \rangle &= \lim_{n\to\infty} \langle \xi, f(T_{n,\lambda})\xi \rangle \\
&\geq \limsup_{n\to\infty} \langle \xi, E_\lambda f(T_n) E_\lambda \rangle + f(a)\langle \xi, (I - E_\lambda)\xi \rangle. \quad (5.23)
\end{aligned}$$

Note that

$$\begin{aligned}
|\langle \xi, &E_\lambda f(T_n) E_\lambda \xi \rangle - \langle \xi, f(T_n)\xi \rangle| \\
&\leq |\langle E_\lambda \xi - \xi, f(T_n) E_\lambda \xi \rangle| + |\langle \xi, f(T_n)(E_\lambda \xi - \xi) \rangle| \leq 2K\|\xi\| \, \|E_\lambda \xi - \xi\|,
\end{aligned}$$

where $K := \sup_n \|f(T_n)\| < +\infty$. Therefore, $\langle \xi, E_\lambda f(T_n) E_\lambda \xi \rangle \to \langle \xi, f(T_n)\xi \rangle$ as $\lambda \to$ "∞" uniformly for n, so it follows from (5.23) that

$$\begin{aligned}
\limsup_{n\to\infty} \langle \xi, f(T_n)\xi \rangle &= \lim_\lambda \limsup_{n\to\infty} \langle \xi, E_\lambda f(T_n) E_\lambda \xi \rangle \\
&\leq \lim_\lambda \{\langle \xi, f(T_\lambda)\xi \rangle - f(a)\langle \xi, (I - E_\lambda)\xi \rangle\} \\
&= \lim_\lambda \langle \xi, f(T_\lambda)\xi \rangle = \langle \xi, f(T)\xi \rangle,
\end{aligned}$$

where the last equality follows since $T_\lambda \to T$ in SOT (the strong operator topology) implies that $f(T_\lambda) \to f(T)$ in SOT, see [65] (or [103, Theorem A.2]). $\square$

Theorem 5.13 *Assume that $f \in \mathcal{F}_{\mathrm{cv}}^{\nearrow}(0,\infty)$ and either $f^\sharp$ is operator convex on $(f'(0^+),\infty)$ or $(f^\sharp)^{-1}$ is operator concave on $(-f(0^+),\infty)$. Then for every $\psi, \varphi \in M_*^+$ such that $\psi \leq \lambda\varphi$ for some $\lambda \in (0,\infty)$, there exists an $H \in M_{\mathrm{sa}}$ such that*

$$S_f^{\mathrm{meas}}(\psi\|\varphi) = S_f^{\mathrm{pr}}(\psi\|\varphi) = S_f(\mathcal{E}_H(\psi)\|\mathcal{E}_H(\varphi)).$$

That is, the supremum of the expression in (5.3) is attained by a projective measurement induced by H.

Proof Let $f \in \mathcal{F}_{\mathrm{cv}}^{\nearrow}(0,\infty)$. Assume that $f^\sharp$ is operator convex on (a,∞), where $a := f'(0^+)$. From $(f^\sharp)'(\infty) = \infty$ (since $f^\sharp \in \mathcal{F}_{\mathrm{cv}}^{\nearrow}(0,\infty)$ by Lemma 5.6 (1)), there is a $t_0 \in (a,\infty)$ such that $(f^\sharp)'(t_0) > \lambda$. For any $A \in M_{(a,\infty)}$ with the spectral decomposition $A = \int_a^\infty t \, dE_A(t)$, define an $A_0 \in M_{(a,t_0)}$ by

$$A_0 := \int_a^{t_0} t \, dE_A(t) + t_0 E_A((t_0,\infty)).$$

We then have

$$\{\psi(A_0) - \varphi(f^\sharp(A_0))\} - \{\psi(A) - \varphi(f^\sharp(A))\}$$

$$= -\int_{(t_0,\infty)} (t - t_0)\, d\psi(E_A(t)) + \int_{(t_0,\infty)} (f^\sharp(t) - f^\sharp(t_0))\, d\varphi(E_A(t))$$

$$\geq \int_{(t_0,\infty)} (t - t_0)\left(\frac{f^\sharp(t) - f^\sharp(t_0)}{t - t_0} - \lambda\right) d\varphi(E_A(t))$$

thanks to $\psi \leq \lambda\varphi$. Since

$$\frac{f^\sharp(t) - f^\sharp(t_0)}{t - t_0} \geq (f^\sharp)'(t_0) > \lambda, \qquad t > t_0,$$

it follows that

$$\psi(A_0) - \varphi(f^\sharp(A_0)) \geq \psi(A) - \varphi(f^\sharp(A)).$$

By Theorem 5.7 (1) and case (1) of Theorem 5.8, this implies that

$$S_f^{\mathrm{meas}}(\psi\|\varphi) = S_t^{\mathrm{pr}}(\psi\|\varphi) = \sup_{A \in M_{[a,t_0]}} \{\psi(A) - \varphi(f^\sharp(A))\},$$

where $f^\sharp(0) := f^\sharp(0^+)$ when $a = 0$. Noting that $\varphi(f^\sharp(A)) = \langle\Phi, f^\sharp(A)\Phi\rangle$ where Φ is the representing vector of φ (in the standard form), we find by Lemma 5.12 (1) that the function $A \in M_{[a,t_0]} \mapsto \psi(A) - \varphi(f^\sharp(A))$ is upper semicontinuous in WOT. Since $M_{[a,t_0]}$ is compact in WOT, there exists an $H \in M_{[a,t_0]}$ such that

$$S^{\mathrm{pr}}(\psi\|\varphi) = \psi(H) - \varphi(f^\sharp(H)).$$

Letting $p := d\psi(E_H(t))$ and $q := d\varphi(E_H(t))$ we have

$$S^{\mathrm{pr}}(\psi\|\varphi) = \int_a^{t_0} t\, dp(t) - \int_a^{t_0} f^\sharp(t)\, dq(t)$$

$$= \int_a^{t_0} \left(t\,\frac{dp}{dq}(t) - f^\sharp(t)\right) dq(t)$$

$$\leq \int_a^{t_0} f\left(\frac{dp}{dq}\right) dq = S_f(p\|q) \leq S_f^{\mathrm{pr}}(\psi\|\varphi),$$

where the last inequality follows from Proposition 5.3. Thus, we obtain

$$S_f^{\mathrm{pr}}(\psi\|\varphi) = S_f(p\|q) = S_f(\mathcal{E}_H(\psi)\|\mathcal{E}_H(\varphi)),$$

as desired.

Next, assume that $(f^\sharp)^{-1}$ is operator concave on (b, ∞), where $b := -f(0^+)$. Note that $(f^\sharp)^{-1}(t) \nearrow \infty$ as $b < t \nearrow \infty$ by Theorem 5.7 (1) and

$$[(f^\sharp)^{-1}]'(t) = \frac{1}{(f^\sharp)'((f^\sharp)^{-1}(t))} \longrightarrow \frac{1}{(f^\sharp)'(\infty)} = 0 \quad \text{as } t \to \infty.$$

So there exists a $t_0 \in (b, +\infty)$ such that $[(f^\sharp)^{-1}]'(t_0) < 1/\lambda$. Now, the proof proceeds as in the previous case. For any $A \in M_{(b,\infty)}$ with the spectral decomposition $A = \int_b^\infty t \, dE_A(t)$, define an $A_0 \in M_{(b,t_0)}$ by

$$A_0 := \int_b^{t_0} t \, dE_A(t) + t_0 E_A((t_0, \infty)).$$

Then

$$\{\psi((f^\sharp)^{-1}(A_0)) - \varphi(A_0)\} - \{\psi((f^\sharp)^{-1}(A)) - \varphi(A)\}$$

$$= - \int_{(t_0,\infty)} ((f^\sharp)^{-1}(t) - (f^\sharp)^{-1}(t_0)) \, d\psi(E_A(t)) + \int_{(t_0,\infty)} (t - t_0) \, d\varphi(E_A(t))$$

$$\geq \int_{(t_0,\infty)} \lambda(t - t_0)\left(\frac{1}{\lambda} - \frac{(f^\sharp)^{-1}(t) - (f^\sharp)^{-1}(t_0)}{t - t_0}\right) d\varphi(E_A(t)) \geq 0,$$

since

$$\frac{(f^\sharp)^{-1}(t) - (f^\sharp)^{-1}(t_0)}{t - t_0} \leq [(f^\sharp)^{-1}]'(t_0) < \frac{1}{\lambda}, \qquad t > t_0.$$

Therefore,

$$\psi((f^\sharp)^{-1}(A_0)) - \varphi(A_0) \geq \psi((f^\sharp)^{-1}(A)) - \varphi(A).$$

By Theorem 5.7 (1) and case (1) of Theorem 5.8, this implies that

$$S_f^{\mathrm{meas}}(\psi\|\varphi) = S_f^{\mathrm{pr}}(\psi\|\varphi) = \sup_{A \in M_{[b,t_0]}} \{\psi((f^\sharp)^{-1}(A)) - \varphi(A)\},$$

where $(f^\sharp)^{-1}$ is operator concave on $[b, \infty)$ with $(f^\sharp)^{-1}(b) := a$. By Lemma 5.12 (2) there exists an $H_0 \in M_{[b,t_0]}$ such that $S_f^{\mathrm{pr}}(\psi\|\varphi) = \psi((f^\sharp)^{-1}(H_0)) - \varphi(H_0)$. Let $H := (f^\sharp)^{-1}(H_0)$; then $H \in M_{[a,t_1]}$ with $t_1 := (f^\sharp)^{-1}(t_0)$ and $H_0 = f^\sharp(H)$, so we have $S_f^{\mathrm{pr}}(\psi\|\varphi) = \psi(H) - \varphi(f^\sharp(H))$. Hence, similarly to the previous case, we have $S_f^{\mathrm{pr}}(\psi\|\varphi) = S_f(\mathcal{E}_H(\psi)\|\mathcal{E}_H(\varphi))$. $\qquad\qquad\square$

Theorem 5.14 *Assume that $g \in \mathcal{F}_{\mathrm{cc}}^{\nearrow}(0, \infty)$ and either $g^\flat$ is operator concave on $(0, g'(0^+))$ or $(g^\flat)^{-1}$ is operator convex on $(-g(\infty), -g(0^+))$. Let $f := -g$ and $\psi, \varphi \in M_*^+$. Assume that one of the following conditions (1)–(4) is satisfied:*

(1) $g(\infty) < \infty$, $g'(0^+) < \infty$ *(with no condition on ψ, φ),*
(2) $g(\infty) = \infty$, $g'(0^+) < \infty$, *and $\psi \leq \lambda\varphi$ for some $\lambda \in (0, \infty)$,*
(3) $g(\infty) < \infty$, $g'(0^+) = \infty$, *and $\varphi \leq \lambda\psi$ for some $\lambda \in (0, \infty)$,*
(4) $g(\infty) = \infty$, $g'(0^+) = \infty$, *and $\lambda^{-1}\varphi \leq \psi \leq \lambda\varphi$ for some $\lambda > 1$.*

Then there exists an $H \in M_{\mathrm{sa}}$ such that

$$S_f^{\mathrm{meas}}(\psi\|\varphi) = S_f^{\mathrm{pr}}(\psi\|\varphi) = S_f(\mathcal{E}_H(\psi)\|\mathcal{E}_H(\varphi)).$$

Proof Let $g \in \mathcal{F}_{\mathrm{cc}}^{\nearrow}(0, \infty)$ and $f := -g$. We may assume that g is non-constant; the conclusion is obviously true when g and hence f are constant.

Case (1). Assume that $g^{\flat}$ is operator concave on $(0, a)$, where $a := g'(0^+)$. Since $g^{\flat}(0^+) = -g(\infty) > -\infty$ (by Lemma 5.6 (2)) and $a = g'(0^+) < \infty$, $g^{\flat}$ is an operator concave function on $[0, a]$, where $g^{\flat}(0) = -g(\infty)$ and $g^{\flat}(a) = -g(0^+) < \infty$. By Theorem 5.7 (2) and case (3) of Theorem 5.8, we have

$$S_f^{\text{meas}}(\psi\|\varphi) = S_f^{\text{pr}}(\psi\|\varphi) = \sup_{A \in M_{[0,a]}} \{-\psi(A) + \varphi(g^{\flat}(A))\}.$$

Hence by Lemma 5.12 (2), there exists an $H \in M_{\text{sa}}$ such that $S_f^{\text{pr}}(\psi\|\varphi) = -\psi(H) + \varphi(g^{\flat}(H))$, and the remaining proof is similar to that in the proof of Theorem 5.13.

Next, assume that $(g^{\flat})^{-1}$ is operator convex on (c, b), where $b := -g(0^+)$ and $c := -g(\infty)$. From Theorem 5.7 (2) and case (3) of Theorem 5.8, $(g^{\flat})^{-1}$ extends to an operator convex function $[c, b] \to [0, a]$ and we have

$$S_f^{\text{meas}}(\psi\|\varphi) = S_f^{\text{pr}}(\psi\|\varphi) = \sup_{A \in M_{[c,b]}} \{-\psi((g^{\flat})^{-1}(A)) + \varphi(A)\}.$$

Hence the proof proceeds as above.

Case (2). Assume that $g^{\flat}$ is operator concave on $(0, a)$. Then $g^{\flat}(0^+) = -g(\infty) = -\infty$ while $a = g'(0^+) < \infty$. Since $(g^{\flat})'(0^+) = \infty$, there is a $t_0 \in (0, a)$ such that $(g^{\flat})'(t_0) > \lambda$. For any $A \in M_{(0,a)}$ with the spectral decomposition $A = \int_0^a t \, dE_A(t)$, define an $A_0 \in M_{[t_0,a]}$ by

$$A_0 := t_0 E_A((0, t_0)) + \int_{[t_0,a]} t \, dE_A(t).$$

Then

$$\{-\psi(A_0) + \varphi(g^{\flat}(A_0))\} - \{-\psi(A) + \varphi(g^{\flat}(A))\}$$

$$= -\int_{(0,t_0)} (t_0 - t) \, d\psi(E_A(t)) + \int_{(0,t_0)} (g^{\flat}(t_0) - g^{\flat}(t)) \, d\varphi(E_A(t))$$

$$\geq \int_{(0,t_0)} (t_0 - t) \left(\frac{g^{\flat}(t_0) - g^{\flat}(t)}{t_0 - t} - \lambda \right) d\varphi(E_A(t)) \geq 0,$$

since

$$\frac{g^{\flat}(t_0) - g^{\flat}(t)}{t_0 - t} \geq (g^{\flat})'(t_0) > \lambda, \qquad 0 < t < t_0.$$

Therefore, by Theorem 5.7 (2) and case (3) of Theorem 5.8, we have

$$S_f^{\text{meas}}(\psi\|\varphi) = S_f^{\text{pr}}(\psi\|\varphi) = \sup_{A \in M_{[t_0,a]}} \{-\psi(A) + \varphi(g^{\flat}(A))\}.$$

Since $g^{\flat}$ is operator concave on $[t_0, a]$, it follows from Lemma 5.12 (2) that the above supremum is attained by some $H \in M_{[t_0,a]}$, for which we have

$$S_f^{\text{pr}}(\psi\|\varphi) = S_f(\mathcal{E}_H(\psi)\|\mathcal{E}_H(\varphi))$$

similarly to the proof of Theorem 5.13.

Next, assume that $(g^b)^{-1}$ is operator convex on $(-\infty, b)$, where $b := -g(0^+)$. Note that

$$[(g^b)^{-1}]'(t) = \frac{1}{(g^b)'((g^b)^{-1}(t))} \longrightarrow \frac{1}{(g^b)'(0^+)} = 0 \quad \text{as } b > t \to -\infty,$$

since $g^b(0^+) = -g(\infty) = -\infty$ by Lemma 5.6 (2) and hence $(g^b)'(0^+) = \infty$. So there is a $t_0 \in (-\infty, b)$ such that $[(g^b)^{-1}]'(t_0) < 1/\lambda$. For any $A \in M_{(-\infty, b)}$ with $A = \int_{-\infty}^b t\, dE_A(t)$, define an $A_0 \in M_{[t_0, b]}$ by

$$A_0 := t_0 E_A((-\infty, t_0)) + \int_{t_0}^b t\, dE_A(t).$$

Then

$$\{-\psi((g^b)^{-1}(A_0) + \varphi(A_0)\} - \{-\psi((g^b)^{-1}(A)) + \varphi(A)\}$$
$$= -\int_{(-\infty, t_0)} ((g^b)^{-1}(t_0) - (g^b)^{-1}(t))\, d\psi(E_A(t)) + \int_{(-\infty, t_0)} (t_0 - t)\, d\varphi(E_A(t))$$
$$\geq \int_{(-\infty, t_0)} (t_0 - t)\left(\frac{1}{\lambda} - \frac{(g^b)^{-1}(t_0) - (g^b)^{-1}(t)}{t_0 - t}\right) d\psi(E_A(t)) \geq 0,$$

since

$$\frac{(g^b)^{-1}(t_0) - (g^b)^{-1}(t)}{t_0 - t} \leq [(g^b)^{-1}]'(t_0) < \frac{1}{\lambda}, \qquad -\infty < t < t_0.$$

Hence we have

$$S_f^{\text{meas}}(\psi\|\varphi) = S_f^{\text{pr}}(\psi\|\varphi) = \sup_{A \in M_{[t_0, b]}} \{-\psi((g^b)^{-1}(A)) + \varphi(A)\}$$
$$= -\psi((g^b)^{-1}(H_0)) + \varphi(H_0)$$

for some $H_0 \in M_{[t_0, b]}$. Letting $H := (g^b)^{-1}(H_0) \in M_{[t_1, a]}$ where $t_1 := g^b(t_0)$, we have

$$S_f^{\text{pr}}(\psi\|\varphi) = -\psi(H) + \varphi(g^b(H)) = S_f(\mathcal{E}_H(\psi)\|\mathcal{E}_H(\varphi))$$

as before.

Case (3). Assume that g^b is operator concave on $(0, \infty)$ and $g'(0^+) = \infty$. Since $g^b(0^+) = -g(\infty) > -\infty$, g^b extends to $[0, \infty)$ with $g^b(0) = -g(\infty)$. Since $(g^b)'(\infty) = 0$ (due to $g^b \in \mathcal{F}_{\text{cc}}^{\nearrow}(0, \infty)$), there is a $t_0 \in (0, \infty)$ such that $(g^b)'(t_0) < 1/\lambda$. For any $A \in M_{++}$ with $A = \int_0^\infty t\, dE_A(t)$, define an $A_0 \in M_{[0, t_0]}$ by

$$A_0 := \int_0^{t_0} t\, dE_A(t) + t_0 E_A((t_0, \infty)).$$

Then, from $\varphi \le \lambda\psi$ we have

$$\{-\psi(A_0) + \varphi(g^b(A_0))\} - \{-\psi(A) + \varphi(g^b(A))\}$$

$$= \int_{(t_0,\infty)} (t - t_0)\, d\psi(E_A(t)) - \int_{(t_0,\infty)} (g^b(t) - g^b(t_0))\, d\varphi(E_A(t))$$

$$\ge \int_{(t_0,+\infty)} (t - t_0)\left(\frac{1}{\lambda} - \frac{g^b(t) - g^b(t_0)}{t - t_0}\right) d\varphi(E_A(t)) \ge 0,$$

since

$$\frac{g^b(t) - g^b(t_0)}{t - t_0} \le (g^b)'(t_0) < \frac{1}{\lambda}, \qquad t > t_0.$$

Hence by Theorem 5.7 (2) and case (3) of Theorem 5.8, we have

$$S_f^{\text{meas}}(\psi\|\varphi) = S_f^{\text{pr}}(\psi\|\varphi) = \sup_{A \in M_{[0,t_0]}} \{-\psi(A) + \varphi(g^b(A))\},$$

and the remaining proof is the same as before.

Next, assume that $(g^b)^{-1}$ is operator convex on (c, b), where $b := -g(0^+) \in (-\infty, \infty]$ and $c := -g(\infty) > -\infty$. Note that

$$[(g^b)^{-1}]'(t) = \frac{1}{(g^b)'((g^b)^{-1}(t))} \longrightarrow \frac{1}{(g^b)'(\infty)} = +\infty \quad \text{as } c < t \nearrow b,$$

since $(g^b)'(\infty) = 0$. So there is a $t_0 \in (c, b)$ such that $[(g^b)^{-1}]'(t_0) > \lambda$. For any $A \in M_{(c,b)}$ with $A = \int_c^b t\, dE_A(t)$, define an $A_0 \in M_{[c,t_0]}$ by

$$A_0 := \int_c^{t_0} t\, dE_A(t) + t_0 E_A((t_0, b)).$$

Then

$$\{-\psi((g^b)^{-1}(A_0) + \varphi(A_0)\} - \{-\psi((g^b)^{-1}(A)) + \varphi(A)\}$$

$$= \int_{(t_0,b)} ((g^b)^{-1}(t) - (g^b)^{-1}(t_0))\, d\psi(E_A(t)) + \int_{(t_0,b)} (t - t_0)\, d\varphi(E_A(t))$$

$$\ge \int_{(t_0,b)} (t_0 - t)\left(\frac{(g^b)^{-1}(t) - (g^b)^{-1}(t_0)}{t - t_0} - \lambda\right)(t - t_0)\, d\psi(E_A(t)) \ge 0,$$

since

$$\frac{(g^b)^{-1}(t) - (g^b)^{-1}(t_0)}{t - t_0} \ge [(g^b)^{-1}]'(t_0) > \lambda, \qquad t_0 < t < b.$$

Since $(g^b)^{-1}$ is operator convex on $[c, t_0]$ where $(g^b)^{-1}(c) = 0$, we have

$$S_f^{\text{meas}}(\psi\|\varphi) = S_f^{\text{pr}}(\psi\|\varphi) = \sup_{A \in M_{[c,t_0]}} \{-\psi((g^b)^{-1}(A)) + \varphi(A)\}$$

$$= -\psi((g^b)^{-1}(H_0)) + \varphi(H_0)$$

for some $H_0 \in M_{[c,t_0]}$, and the remaining proof is the same as before.

Case (4). When $g^\flat$ is operator concave on $(0, \infty)$, we may combine the proofs of cases (2) and (3). In this case, we can choose t_0, t_1 with $0 < t_0 < t_1 < \infty$ such that $(g^\flat)'(t_0) > \lambda$ and $(g^\flat)'(t_1) < 1/\lambda$. Then for any $A \in M_{++}$, defining an $A_0 \in M_{[t_0,t_1]}$ by

$$A_0 := t_0 E_A((0, t_0)) + \int_{[t_0,t_1]} t \, dE_A(t) + t_1 E_A((t_1, \infty)),$$

the proof proceeds as before, while we omit the details. When $(g^\flat)^{-1}$ is operator convex on $(-\infty, b)$, where $b := -g(0^+) \in (-\infty, \infty]$. The proof is again the combination of those of cases (2) and (3); we omit the details. $\qquad\qquad\square$

Remark 5.15 As in Theorem 5.8, when the transpose $\widetilde{g}$ is in $\mathcal{F}_{cc}^{\nearrow}(0, \infty)$ and the assumption in Theorem 5.14 is satisfied for $\widetilde{g}$, the same attainability result for S_f^{pr} as in Theorem 5.14 holds under exchanging the roles of ψ, φ.

Example 5.16 (1) When $1 < \alpha < \infty$, consider $f_\alpha(t) := t^\alpha \in \mathcal{F}_{cv}^{\nearrow}(0, \infty)$. From Example 5.9 (1), $(f^\sharp)^{-1}$ is operator concave on $(0, \infty)$. Hence by Theorem 5.13, if $\psi \le \lambda \varphi$ for some $\lambda > 0$, then there exists an $H \in M_{sa}$ such that

$$S_{f_\alpha}^{meas}(\psi \| \varphi) = S_{f_\alpha}^{pr}(\psi \| \varphi) = S_{f_\alpha}(\mathcal{E}_H(\psi) \| \mathcal{E}_H(\varphi))$$

so that

$$D_\alpha^{meas}(\psi \| \varphi) = D_\alpha^{pr}(\psi \| \varphi) = D_\alpha(\mathcal{E}_H(\psi) \| \mathcal{E}_H(\varphi)).$$

(2) When $0 < \alpha < 1$, consider $g_\alpha(t) := t^\alpha \in \mathcal{F}_{cc}^{\nearrow}(0, \infty)$ and $f_\alpha := -g_\alpha$. Then $g_\alpha(0^+) = 0$, $g_\alpha(\infty) = \infty$, and $g'_\alpha(0^+) = \infty$. From Example 5.9 (2), when $1/2 \le \alpha < 1$, $\widetilde{g}_\alpha^\flat$ is operator concave on $(0, \widetilde{g}'_\alpha(0^+)) = (0, \infty)$. When $0 < \alpha \le 1/2$, $(\widetilde{g}_\alpha^\flat)^{-1}$ is operator convex on $(-\widetilde{g}_\alpha(\infty), -\widetilde{g}_\alpha(0^+)) = (-\infty, 0)$. Hence by case (4) of Theorem 5.14 (and Remark 5.15), if $\lambda^{-1}\varphi \le \psi \le \lambda\varphi$ for some $\lambda > 1$, there exists an $H \in M_{sa}$ such that

$$S_{f_\alpha}^{meas}(\psi \| \varphi) = S_{f_\alpha}^{pr}(\psi \| \varphi) = S_{f_\alpha}(\mathcal{E}_H(\psi) \| \mathcal{E}_H(\varphi))$$

so that

$$D_\alpha^{meas}(\psi \| \varphi) = D_\alpha^{pr}(\psi \| \varphi) = D_\alpha(\mathcal{E}_H(\psi) \| \mathcal{E}_H(\varphi)).$$

(3) Consider $g_0(t) := \log t \in \mathcal{F}_{cc}^{\nearrow}(0, \infty)$ and $f_0 := -g_0$. Then $g_0(0^+) = -\infty$, $g_0(\infty) = \infty$, and $g'_0(0^+) = \infty$. From Example 5.10, $g_0^\flat$ is operator concave on $(0, g'_0(0^+)) = (0, \infty)$. Hence by case (4) of Theorem 5.14, for every $\psi, \varphi \in M_*^+$ with $\lambda^{-1}\varphi \le \psi \le \lambda\varphi$, there exists an $H \in M_{sa}$ such that

$$D^{meas}(\psi \| \varphi) = D^{pr}(\psi \| \varphi) = D(\mathcal{E}_H(\psi) \| \mathcal{E}_H(\varphi)).$$

(4) For each $\lambda \in (0, \infty)$, consider $g_\lambda(t) := \frac{t}{t+\lambda} \in \mathcal{F}_{cc}^{\nearrow}(0, \infty)$ and $f_\lambda := -g_\lambda$. Then $g_\lambda(0^+) = 0$, $g_\lambda(\infty) = 1$, and $g'_\lambda(0^+) = 1/\lambda$. From Example 5.11, $(g_\lambda^\flat)^{-1}$ is operator convex on $(-g_\lambda(\infty), -g_\lambda(0^+)) = (-1, 0)$. Hence by case (1) of Theorem 5.14, for every $\psi, \varphi \in M_*^+$ (with no condition), there exists an $H \in M_{sa}$ such that

$$S_{f_\lambda}^{\text{meas}}(\psi\|\varphi) = S_{f_\lambda}^{\text{pr}}(\psi\|\varphi) = S_{f_\lambda}(\mathcal{E}_H(\psi)\|\mathcal{E}_H(\varphi)).$$

In the two examples below we will give the descriptions of optimal measurements for the particular Rényi divergences D_α when $\alpha = 2$ and $1/2$. Similar discussions on optimal measurements for those Rényi divergences in the finite-dimensional case are found in [80]. In the following, let $\psi, \varphi \in M_*^+$ with $\psi(1) = 1$ for simplicity.

Example 5.17 (Optimal measurement for D_2) Let $f_2(t) := t^2$ and so $f_2^\sharp(t) = t^2/4$ as given in Example 5.9 (1). By (5.17) we have

$$Q_2^{\text{meas}}(\psi\|\varphi) = S_{f_2}^{\text{meas}}(\psi\|\varphi) = \sup_{x\in M_+} \{2\psi(x) - \varphi(x^2)\}.$$

Assume that $\psi \le \lambda\varphi$ for some $\lambda > 0$. In this case, we may assume that φ is faithful. Then by the linear Radon-Nikodym theorem due to Sakai [101, 1.24.4], there exists an $h_0 \in M_+$ such that

$$\psi(x) = \frac{1}{2}\varphi(h_0 x + x h_0), \qquad x \in M. \tag{5.24}$$

For every $h \in M_{sa}$ we find that

$$
\begin{aligned}
&2\psi(h_0 + h) - \varphi((h_0 + h)^2)\\
&\quad = 2\psi(h_0) - \varphi(h_0^2) + \{2\psi(h) - \varphi(h_0 h + h h_0)\} - \varphi(h^2)\\
&\quad \le 2\psi(h_0) - \varphi(h_0^2).
\end{aligned}
$$

Therefore,

$$Q_2^{\text{meas}}(\psi\|\varphi) = 2\psi(h_0) - \varphi(h_0^2) = 2\psi(h_0) - \psi(h_0) = \psi(h_0).$$

Consequently, the optimal measurement for D_2 is induced from the spectral resolution of h_0.

According to Kosaki [67, Theorem 1.6], a linear Radon-Nikodym derivative h_0 satisfying (5.24) exists if and only if $\widetilde{\psi} \le \lambda\varphi$ for some $\lambda > 0$, where

$$\widetilde{\psi} := \int_{-\infty}^{\infty} \frac{1}{\cosh \pi t}\, \psi \circ \sigma_t^\varphi \, dt \tag{5.25}$$

(of course, $\psi \le \lambda\varphi \implies \widetilde{\psi} \le \lambda\varphi$), and in this case, h_0 is determined as

$$h_0 = [D\widetilde{\psi} : D\varphi]_{-i/2}^* [D\widetilde{\psi} : D\varphi]_{-i/2},$$

where $[D\widetilde{\psi} : D\varphi]_t$, $t \in \mathbb{R}$, is the Connes cocycle derivative and $[D\widetilde{\psi} : D\varphi]_{-i/2}$ is the analytic continuation at $t = -i/2$, see Sec. A.7. Therefore,

$$D_2^{\text{meas}}(\psi\|\varphi) = \log \psi([D\widetilde{\psi} : D\varphi]_{-i/2}^* [D\widetilde{\psi} : D\varphi]_{-i/2}).$$

Assume that $M = B(\mathcal{H})$ on a finite-dimensional Hilbert space $\mathcal{H}$ and $\psi, \varphi \in B(\mathcal{H})^*_+$ have the density operators D_ψ, D_φ. Then we note that $\widetilde{\psi}$ in (5.25) has the density

$$D_{\widetilde{\psi}} = \int_{-\infty}^{\infty} \frac{1}{\cosh \pi t} D_\varphi^{-it} D_\psi D_\varphi^{it} \, dt.$$

Therefore,

$$h_0 = (D_{\widetilde{\psi}}^{1/2} D_\varphi^{-1/2})^* (D_{\widetilde{\psi}}^{1/2} D_\varphi^{-1/2}) = D_\varphi^{-1/2} D_{\widetilde{\psi}} D_\varphi^{-1/2}$$

$$= \int_{-\infty}^{\infty} \frac{1}{\cosh \pi t} D_\varphi^{-it} D_\varphi^{-1/2} D_\psi D_\varphi^{-1/2} D_\varphi^{-it} \, dt$$

and

$$Q_2^{\text{meas}}(\psi \| \varphi) = \operatorname{Tr} D_\psi h_0 = \int_{-\infty}^{\infty} \frac{1}{\cosh \pi t} \operatorname{Tr} D_\psi D_\varphi^{-it} D_\varphi^{-1/2} D_\psi D_\varphi^{-1/2} D_\varphi^{-it} \, dt.$$

Since (5.24) is nothing but the familiar Lyapunov equation $D_\psi = \frac{1}{2}(h_0 D_\varphi + D_\varphi h_0)$, the alternative well-known formula of h_0 is

$$h_0 = 2 \int_0^{\infty} e^{-t D_\varphi} D_\psi e^{-t D_\varphi} \, dt.$$

Example 5.18 (Optimal measurement for $D_{1/2}$) We first show that for every $\psi, \varphi \in M_*^+$ (with $\psi(1) = 1$),

$$D_{1/2}^{\text{meas}}(\psi \| \varphi) = \widetilde{D}_{1/2}(\psi \| \varphi) = -2 \log F(\psi, \varphi). \tag{5.26}$$

Indeed, the latter equality is (3.24). By Example 5.9 (2) we note that

$$Q_{1/2}^{\text{meas}}(\psi \| \varphi) = \frac{1}{2} \inf_{x \in M_{++}} \{\psi(x^{-1}) + \varphi(x)\}.$$

On the other, if $\psi \leq \lambda \varphi$ for some $\lambda > 0$, then (3.27) in Lemma 3.18 gives

$$\widetilde{Q}_{1/2}(\psi \| \varphi) = \frac{1}{2} \inf_{x \in M_{++}} \{\psi(x^{-1}) + \varphi(x)\}$$

so that $D_{1/2}^{\text{meas}}(\psi \| \varphi) = \widetilde{D}_{1/2}(\psi \| \varphi)$. For general ψ, φ it follows from Proposition 5.4 (4) and Theorem 3.15 (3) that

$$D_{1/2}^{\text{meas}}(\psi \| \varphi) = \lim_{\varepsilon \searrow 0} D_{1/2}^{\text{meas}}(\psi + \varepsilon(\psi + \varphi) \| \varphi + \varepsilon(\psi + \varphi))$$

$$= \lim_{\varepsilon \searrow 0} \widetilde{D}_{1/2}(\psi + \varepsilon(\psi + \varphi) \| \varphi + \varepsilon(\psi + \varphi)) = \widetilde{D}_{1/2}(\psi \| \varphi).$$

Therefore, (5.26) has been shown, which also says that the measured version of the fidelity F is the same as F itself and $F(\psi, \varphi) = \frac{1}{2} \sup_{x \in M_{++}} \{\psi(x^{-1}) + \varphi(x)\}$. This and other similar expressions for F were given in [2, 3].

Now, assume that $h_\psi^\delta \le \lambda h_\varphi^\delta$ for some $\delta, \lambda > 0$, or equivalently, $s(\psi) \le s(\varphi)$ and Connes' cocycle derivative $[D\psi : D\varphi]_t$ has the analytic continuation to the strip $-\delta/2 \le \mathrm{Im}\, z \le 0$, see Lemma A.58 of Appendix A.7. Then by the $p = 1$ case of Lemma A.59 (Kosaki's extension [72] of Sakai's quadratic Radon-Nikodym theorem), there exists a unique $k_1 \in (s(\varphi)Ms(\varphi))_+$ such that $h_\psi = k_1 h_\varphi k_1$, i.e., $\psi(x) = \varphi(k_1 x k_1)$ for all $x \in M$. Moreover, in this case, k_1 satisfies $(h_\varphi^{1/2} h_\psi h_\varphi^{1/2})^{1/2} = h_\varphi^{1/2} k_1 h_\varphi^{1/2}$ so that

$$F(\psi, \varphi) = \mathrm{tr}\,(h_\varphi^{1/2} h_\psi h_\varphi^{1/2})^{1/2} = \mathrm{tr}\, h_\varphi k_1 = \varphi(k_1). \qquad (5.27)$$

With the spectral decomposition $k_1 = \int_0^r t \, dE_{k_1}(t)$ with $r := \|k_1\|$, let $p := d\psi(E_{k_1}(t))$ and $q := d\varphi(E_{k_1}(t))$. Then $p \ll q$ (absolutely continuous) since $s(\psi) \le s(\varphi)$, and we have

$$Q_{1/2}^{\mathrm{meas}}(\psi\|\varphi) \le \int_0^r \left(\frac{dp}{dq}\right)^{1/2} dq = \lim_{\varepsilon \searrow 0} \int_\varepsilon^r \left(\frac{dp}{dq}\right)^{1/2} dq$$

$$\le \lim_{\varepsilon \searrow 0} \frac{1}{2} \int_\varepsilon^r \left(t^{-1} \frac{dp}{dq}(t) + t\right) dq(t)$$

$$= \lim_{\varepsilon \searrow 0} \frac{1}{2} \left(\int_\varepsilon^r t^{-1} \, dp(t) + \int_0^r t \, dq(t)\right).$$

Furthermore, since $\int_0^r t \, dq(t) = \varphi(k_1)$ and

$$\int_\varepsilon^r t^{-1} \, dp(t) = \psi\left(\int_\varepsilon^r t^{-1} \, dE_{k_1}(t)\right) = \varphi\left(k_1\left(\int_\varepsilon^r t^{-1} \, dE_{k_1}(t)\right)k_1\right)$$

$$= \varphi\left(\int_\varepsilon^r t \, dE_{k_1}(t)\right) \longrightarrow \varphi(k_1) \quad \text{as } \varepsilon \searrow 0,$$

it follows that

$$D_{1/2}^{\mathrm{meas}}(\psi\|\varphi) \ge D_{1/2}(p\|q) = -2\log \int_0^r \left(\frac{dp}{dq}\right)^{1/2} dq \ge -2\log \varphi(k_1). \qquad (5.28)$$

By (5.26)–(5.28) we find that

$$D_{1/2}^{\mathrm{meas}}(\psi\|\varphi) = -2\log F(\psi, \varphi) = -2\log \varphi(k_1)$$

and the optimal measurement for $D_{1/2}$ is induced from the spectral resolution of k_1.

In the finite-dimensional case with the density operators D_ψ, D_φ with φ faithful, the optimal measurement is given by the spectral decomposition of

$$k_1 := D_\varphi^{-1/2}(D_\varphi^{1/2} D_\psi D_\varphi^{1/2})^{1/2} D_\varphi^{-1/2} = D_\psi \# D_\varphi^{-1},$$

the geometric mean of D_ψ and D_φ^{-1}, as explained in [87, Sec. 9.2.2].

Problem 5.19 (1) An example of operator convex function on $(0, \infty)$ for which $S_f^{\mathrm{pr}}(\psi \| \varphi) < S_f^{\mathrm{meas}}(\psi \| \varphi)$ for some $\psi, \varphi \in M_*^+$ is not known. Such an example might exist in the finite-dimensional setting if any.

(2) In the attainability results in Theorems 5.13 and 5.14 (as well as Examples 5.17 and 5.18) we have proved the existence of von Neumann measurements induced by bounded self-adjoint operators in M, under such an assumption as $\psi \leq \lambda \varphi$ or more strongly $\lambda^{-1} \varphi \leq \psi \leq \lambda \varphi$ for some $\lambda > 0$. These dominance assumptions are a bit too strict in the infinite-dimensional setting, so it is desirable to remove them by considering measurements induced by unbounded self-adjoint operators. However, the problem seems difficult because the compactness argument has been used in the proof of the theorems.

(3) It is also interesting to find a more explicit form of optimal measurements, as in Examples 5.17 and 5.18, in the attainability results of Theorems 5.13 and 5.14. Since an optimal measurement is given as the maximizer of a certain strictly convex function on a certain set $M_{[a,b]}$, the maximizer might explicitly be specified.

Chapter 6
Reversibility and Quantum Divergences

6.1 Petz' Recovery Map

Let M and N be von Neumann algebras, whose standard forms are $(M, \mathcal{H}, J, \mathcal{P})$ and $(N, \mathcal{H}_0, J_0, \mathcal{P}_0)$, respectively. For convenience, we first summarize basic properties of positive linear maps between von Neumann algebras, though those have already been used in previous chapters. For a positive linear map $\gamma : N \to M$, γ is *unital* if $\gamma(1) = 1$. The γ is *normal* if $\gamma(y_j) \nearrow \gamma(y)$ for any net $\{y_j\}$ in N_+ with $y_j \nearrow y \in N_+$, or equivalently, γ is continuous with respect to the σ-weak topologies (i.e., the $\sigma(N, N_*)$, $\sigma(M, M_*)$-topologies) on N and M. We call γ a *Schwarz map* if $\gamma(y)^*\gamma(y) \le \gamma(y^*y)$ for all $y \in N$. The γ is said to be *n-positive* if $\gamma \otimes \mathrm{id}_n : N \otimes \mathbb{M}_n \to M \otimes \mathbb{M}_n$ is positive, where $\mathbb{M}_n$ is the $n \times n$ matrix algebra and $(\gamma \otimes \mathrm{id}_n)([y_{ij}]) := [\gamma(y_{ij})]$ for $[y_{ij}]_{i,j=1}^n \in \mathbb{M}_n(N) = N \otimes \mathbb{M}_n$. Furthermore, γ is said to be *completely positive (CP)* if it is *n*-positive for all $n \ge 1$. As is well-known,

$$\mathrm{CP} \implies 2\text{-positive} \implies \text{Schwarz map} \implies \text{simply positive.}$$

In quantum information in the finite-dimensional setting, quantum channels (implementing quantum operations) are typically *TPCP* (trace-preserving CP) maps, whose dual maps are unital CP maps. But we may consider, depending on problems, weaker notions of positivity such as Schwarz positivity or 2-positivity or even plain positivity for quantum operations. In the von Neumann algebra setting, normality is essential for a channel $\gamma : N \to M$, that means that γ is the dual of the predual map $\gamma_* : M_* \to N_*$.

The next theorem was proved in [1, 93] under the assumption that both of φ and $\varphi \circ \gamma$ are faithful, but the faithfulness of φ is not essential.

Theorem 6.1 ([1, 93]) *Let* $\gamma : N \to M$ *be a unital normal positive map, and* $\varphi \in M_*^+$. *Assume that* $\varphi \circ \gamma$ *is faithful. Let* $\Phi \in \mathcal{P}$ *and* $\Phi_0 \in \mathcal{P}_0$ *be the vector representatives of* φ *and* $\varphi \circ \gamma$. *Then there exists a unique unital normal positive map* $\beta : M \to N$ *such that*

$$\langle J_0 \beta(x)\Phi_0, y\Phi_0\rangle = \langle Jx\Phi, \gamma(y)\Phi\rangle \tag{6.1}$$

for all $x \in M$ and all $y \in N$.

Proof (Theorem 6.1) First, assuming that for each $x \in M$, $\beta(x) \in N$ exists so that (6.1) holds for all $y \in N$, we prove that $\beta : M \to N$ is a unital normal positive map. Since Φ_0 is a cyclic vector for N (i.e., $N\Phi_0 := \{y\Phi_0 : y \in N\}$ is dense in $\mathcal{H}_0$), it is obvious that $\beta(x) \in N$ is uniquely determined by (6.1) and $\beta : M \to N$ is a linear map. Since $\langle J\Phi, \gamma(y)\Phi\rangle = (\varphi \circ \gamma)(y) = \langle J_0\Phi_0, y\Phi_0\rangle$ for all $y \in N$, we have $\beta(1) = 1$. For every $x \in M_+$ and $y \in N$, it follows from (6.1) that

$$\langle (J_0\beta(x)J_0)y\Phi_0, y\Phi_0\rangle = \langle J_0\beta(x)\Phi_0, y^*y\Phi_0\rangle = \langle Jx\Phi, \gamma(y^*y)\Phi\rangle$$
$$= \langle (JxJ)\gamma(y^*y)^{1/2}\Phi, \gamma(y^*y)^{1/2}\Phi\rangle \ge 0,$$

implying that $J_0\beta(x)J_0 \ge 0$ and so $\beta(x) \in N_+$. Hence β is a unital positive map. To prove the normality of β, let $\{x_j\} \subset M_+$ be a net with $x_j \nearrow x \in M_+$; then $\beta(x_j) \nearrow y_0$ for some $y_0 \in N_+$. For every $y \in N$ we have

$$\langle J_0 y_0 \Phi_0, y\Phi_0\rangle = \lim_j \langle J_0\beta(x_j)\Phi_0, y\Phi_0\rangle = \lim_j \langle Jx_j\Phi, \gamma(y)\Phi\rangle = \langle Jx\Phi, \gamma(y)\Phi\rangle,$$

which implies that $y_0 = \beta(x)$ so that $\beta(x_j) \nearrow \beta(x)$. Hence β is normal.

It remains to prove, for each $x \in M$, the existence of $\beta(x)$ satisfying (6.1) for all $y \in N$. For this, by linearity we may assume that $x \in M_+$. Define $\psi_x \in N_*$ by $\psi_x(y) := \langle Jx\Phi, \gamma(y)\Phi\rangle$, $y \in N$. For every $y \in N_+$ one has

$$\psi_x(y) = \langle \gamma(y)^{1/2}(JxJ)\gamma(y)^{1/2}\Phi, \Phi\rangle \le \|x\|\langle\gamma(y)\Phi, \Phi\rangle = \|x\|(\varphi \circ \gamma)(y)$$

so that $0 \le \psi_x \le \|x\|(\varphi \circ \gamma)$. Hence by Lemma A.21 of Appendix A.3 there exists a $B \in N$ such that $\Psi_x = B\Phi_0$, where $\Psi_x \in \mathcal{P}$ is the vector representative of ψ_x. Since $\Psi_x = J_0\Psi_x = J_0B\Phi_0 = J_0BJ_0\Phi_0$, for every $y \in N$,

$$\psi_x(y) = \langle B\Phi_0, yJ_0BJ_0\Phi_0\rangle = \langle J_0B^*J_0B\Phi_0, y\Phi_0\rangle$$
$$= \langle J_0B^*\Psi_x, y\Phi_0\rangle = \langle J_0B^*B\Phi_0, y\Phi_0\rangle.$$

Letting $\beta(x) := B^*B \in N_+$ one obtains $\psi_x(y) = \langle J_0\beta(x)\Phi_0, y\Phi_0\rangle$, $y \in N$, which gives (6.1).   □

Definition 6.2 Under the assumption of $\gamma : N \to M$ and $\varphi \in M_*^+$ in Theorem 6.1, we refer to $\beta : M \to N$ satisfying (6.1) as *Petz' recovery map* of γ with respect to φ, and denote it by γ_φ^*. The γ_φ^* is defined by (6.1) independently of choice of the standard forms of M, N due to Theorem A.19 of Appendix A.3.

Proposition 6.3 *Let $\gamma : N \to M$ and $\varphi \in M_*^+$ be as in Theorem 6.1.*

(1) $\varphi \circ \gamma \circ \gamma_\varphi^* = \varphi$.
(2) *Assume that φ is faithful as well as $\varphi \circ \gamma$. Then $(\gamma_\varphi^*)_{\varphi\circ\gamma}^* = \gamma$, i.e., γ is Petz' recovery map of γ_φ^* with respect to $\varphi \circ \gamma$.*

(3) *If γ is 2-positive (resp., CP), then γ_φ^* is 2-positive (resp., CP).*

Proof We write β for γ_φ^*.

(1) Letting $y = 1$ in (6.1) gives

$$\varphi(x) = \langle Jx\Phi, \Phi\rangle = \langle J_0\beta(x)\Phi_0, \Phi_0\rangle = \varphi \circ \gamma(\beta(x)), \qquad x \in M,$$

so that $\varphi = \varphi \circ \gamma \circ \beta$.

(2) immediately follows from (1) and (6.1) since the roles of γ and β are symmetric in (6.1).

(3) Assume that γ is 2-positive. Consider the von Neumann algebra $M^{(2)} := M \otimes \mathbb{M}_2$, the tensor product of M with the 2×2 matrix algebra $\mathbb{M}_2$, whose standard form is described in Example A.20 of Appendix A.3. In fact, the standard representation of $M^{(2)}$ is given in (A.5) on $\mathcal{H}^{(2)} := \mathcal{H} \oplus \mathcal{H} \oplus \mathcal{H} \oplus \mathcal{H}$ with the standard involution $J^{(2)}$ in (A.7). Similarly, we take the standard $\mathcal{H}_0^{(2)}$ and $J_0^{(2)}$ for $N^{(2)} := N \otimes \mathbb{M}_2$. Set $\varphi^{(2)} \in (M^{(2)})_*^+$ by $\varphi^{(2)}\left(\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}\right) := \varphi(x_{11}) + \varphi(x_{22})$. Since γ is 2-positive, we have $\varphi^{(2)} \circ (\gamma \otimes \mathrm{id}_2) \in (N^{(2)})_*^+$, which is given as $\varphi^{(2)} \circ (\gamma \otimes \mathrm{id}_2)\left(\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}\right) = \varphi \circ \gamma(y_{11}) + \varphi \circ \gamma(y_{22})$. Moreover, note that the vector representatives of $\varphi^{(2)}$ and $\varphi^{(2)} \circ (\gamma \otimes \mathrm{id}_2)$ are $\Phi^{(2)} = \begin{bmatrix} \Phi & 0 \\ 0 & \Phi \end{bmatrix}$ $(= \Phi \oplus 0 \oplus 0 \oplus \Phi)$ and $\Phi_0^{(2)} = \begin{bmatrix} \Phi_0 & 0 \\ 0 & \Phi_0 \end{bmatrix}$ $(= \Phi_0 \oplus 0 \oplus 0 \oplus \Phi_0)$, respectively, see Example A.20. Then for $x = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \in M^{(2)}$ and $y = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \in N^{(2)}$, with $\beta = \gamma_\varphi^*$, we have

$$\begin{aligned} \langle J_0^{(2)}(\beta \otimes \mathrm{id}_2)(x)\Phi_0^{(2)}, y\Phi_0^{(2)}\rangle &= \sum_{i,j=1}^2 \langle J_0\beta(x_{ij})\Phi_0, y_{ij}\Phi_0\rangle \\ &= \sum_{i,j=1}^2 \langle Jx_{ij}\Phi, \gamma(y_{ij})\Phi\rangle \\ &= \langle J^{(2)}x\Phi^{(2)}, (\gamma \otimes \mathrm{id}_2)(y)\Phi^{(2)}i\rangle \end{aligned}$$

so that $\beta \otimes \mathrm{id}_2$ is Petz' recovery map of $\gamma \otimes \mathrm{id}_2$ with respect to $\varphi^{(2)}$. Hence $\beta \otimes \mathrm{id}_2$ is positive (as seen from the proof of Theorem 6.1), i.e., β is 2-positive. For any $n \geq 1$, if γ is n-positive, then one can prove that β is so, in a similar way by taking the standard forms of $M \otimes \mathbb{M}_n$ and $N \otimes \mathbb{M}_n$ (while the details are omitted). Thus, if γ is CP, then so is β. $\qquad\square$

Remark 6.4 Petz' recovery map γ_φ^* of a Schwarz map γ is not necessarily a Schwarz map, as shown by [58, Proposition 2] even in the finite-dimensional setting. This is the reason why we often need to assume in the reversibility problem that a quantum operation is 2-positive rather than a Schwarz map.

Let M_0 be a von Neumann subalgebra of M and $(M_0, \mathcal{H}_0, J_0, \mathcal{P}_0)$ be the standard form of M_0. For a faithful $\varphi \in M_*^+$ let $\Phi_0 \in \mathcal{P}_0$ be the vector representative of $\varphi|_{M_0}$ as well as $\Phi \in \mathcal{P}$ of φ. Then the *generalized conditional expectation* $E_\varphi = E_\varphi^{M_0} : M \to M_0$ (due to Accardi and Cecchini [1]) is given as

$$E_\varphi(x) := J_0 V_\varphi^* J x J V_\varphi J_0, \qquad x \in M,$$

that is also defined by

$$E_\varphi(x)\Phi_0 = J_0 V_\varphi^* J x \Phi, \qquad x \in M,$$

where $V_\varphi : \mathcal{H}_0 \to \mathcal{H}$ is the isometry defined by $V_\varphi y \Phi_0 := y\Phi$, $y \in M_0$. In this case, since

$$\langle J_0 E_\varphi(x)\Phi_0, y\Phi_0 \rangle = \langle J x \Phi, V_\varphi y \Phi_0 \rangle = \langle J x \Phi, y\Phi \rangle, \qquad x \in M, \ y \in M_0,$$

we have the following:

Proposition 6.5 *Let M_0 be a von Neumann subalgebra of M. If $\varphi \in M_*^+$ is faithful, then Petz' recovery map of the inclusion map $M_0 \hookrightarrow M$ with respect to φ is the generalized conditional expectation $E_\varphi : M \to M_0$.*

We can extend Petz' recovery map γ_φ^* to arbitrary $\varphi \in M_*^+$ as follows.

Proposition 6.6 *Let γ be as in Theorem 6.1. For every $\varphi \in M_*^+$ let $\Phi \in \mathcal{P}$ and $\Phi_0 \in \mathcal{P}_0$ be the vector representatives of φ and $\varphi \circ \gamma$. Let $e := s(\varphi) \in M$ and $e_0 := s(\varphi \circ \gamma) \in N$. Then there exists a unique unital (i.e., $\beta(1) = e_0$) normal positive map $\beta : M \to e_0 N e_0$ such that (6.1) holds for all $x \in M$ and all $y \in e_0 N e_0$. Moreover, $\beta(1 - e) = 0$ and $\beta(x) = \beta(exe)$ for all $x \in M$.*

Proof Let $e_0' := J_0 e_0 J_0 = s_{N'}(\varphi \circ \gamma)$ and $q := e_0 e_0'$. By Proposition A.16 note that $e_0 N e_0 \cong q N q$ by $y \mapsto yq$ ($y \in e_0 N e_0$) and the standard form of $q N q$ is $(q N q, q \mathcal{H}_0, q J_0 q, q \mathcal{P}_0)$. Moreover, $\Phi_0 \in q\mathcal{P}_0$. Although $\gamma|_{e_0 N e_0} : e_0 N e_0 \to M$ is not necessarily unital, we can apply the proof of Theorem 6.1 to $\gamma|_{e_0 N e_0}$, so we have a unique normal map $\beta : M \to e_0 N e_0$ such that

$$\langle (q J_0 q)(\beta(x)q)\Phi_0, (yq)\Phi_0 \rangle = \langle J x \Phi, \gamma(y)\Phi \rangle$$

for all $x \in M$ and $y \in e_0 N e_0$. Since $q J_0 q = e_0 e_0' J_0$, the above LHS is equal to

$$\langle J_0 \beta(x)\Phi_0, e_0 e_0' y\Phi_0 \rangle = \langle J_0 \beta(x)\Phi_0, y\Phi_0 \rangle.$$

Hence β is determined by (6.1) for all $x \in M$ and $y \in e_0 N e_0$. Then $\beta(1) = e_0$ follows since $\langle J\Phi, \gamma(y)\Phi \rangle = (\varphi \circ \gamma)(y) = \langle J_0 e_0 \Phi_0, y\Phi_0 \rangle$ for $y \in e_0 N e_0$. Moreover, for every $x \in M$, since $J x (1 - e)\Phi = 0$, one has $\beta(x(1 - e)) = 0$ and $\beta((1 - e)x) = \beta(x^*(1 - e))^* = 0$. Therefore, $\beta(1 - e) = 0$ and $\beta(x) = \beta(exe)$ for all $x \in M$. $\square$

We denote $\beta : M \to e_0 N e_0$ in the above proposition by γ_φ^*, which may be called again Petz' recovery map of γ with respect to φ. This extended version will be used later in Chap. 8. If γ is 2-positive, then so is γ_φ^*, as in Proposition 6.3 (3).

Example 6.7 Here we consider the finite-dimensional case $M = B(\mathcal{H})$ and $N = B(\mathcal{K})$ with $\dim \mathcal{H}, \dim \mathcal{K} < \infty$. Let $T : B(\mathcal{H}) \to B(\mathcal{K})$ be a trace-preserving positive map, so the dual map $\gamma = T^* : B(\mathcal{K}) \to B(\mathcal{H})$ is a unital positive map. For every $\varphi \in B(\mathcal{H})^*_+$ with density D_φ, the density of $\varphi \circ \gamma \in B(\mathcal{K})^*_+$ is $T(D_\varphi)$ and $E_0 = s(\varphi \circ \gamma)$ is the support projection of $T(D_\varphi)$. Let $\beta : B(\mathcal{H}) \to E_0 B(\mathcal{K}) E_0$ be as given in Proposition 6.6. Recall that the standard representation of $B(\mathcal{H})$ is the left multiplication on $B(\mathcal{H})$ with the Hilbert-Schmidt inner product and $J = {}^*$. Hence for every $X \in B(\mathcal{H})$, $\beta(X)$ is given as

$$\operatorname{Tr} T(D_\varphi)^{1/2} \beta(X) T(D_\varphi)^{1/2} Y = \operatorname{Tr} T(D_\varphi^{1/2} X D_\varphi^{1/2}) Y, \quad Y \in E_0 B(\mathcal{K}) E_0.$$

Since $E_0 T(D_\varphi^{1/2}) E_0 = T(D_\varphi^{1/2})$ and $E_0 T(D_\varphi^{1/2} X D_\varphi^{1/2}) E_0 = T(D_\varphi^{1/2} X D_\varphi^{1/2})$, we have

$$T(D_\varphi)^{1/2} \beta(X) T(D_\varphi)^{1/2} = T(D_\varphi^{1/2} X D_\varphi^{1/2}). \tag{6.2}$$

Therefore,

$$\beta(X) = T(D_\varphi)^{-1/2} T(D_\varphi^{1/2} X D_\varphi^{1/2}) T(D_\varphi)^{-1/2},$$

where $T(D_\varphi)^{-1/2}$ is defined under restriction to the support $E_0 \mathcal{K}$. This expression of β is the same as in [52, (3.19)].

6.2 A Technical Lemma

The next notion of multiplicative domains plays an important role not only in operator algebras but also in quantum information.

Definition 6.8 Let $\gamma : N \to M$ be a unital normal Schwarz map between von Neumann algebras. The *multiplicative domain* of γ is defined as

$$\mathcal{M}_\gamma := \{ y \in N : \gamma(y^* y) = \gamma(y)^* \gamma(y), \ \gamma(yy^*) = \gamma(y)\gamma(y^*) \}.$$

The next result was first shown by Choi [22] when γ is a unital 2-positive map. See [53, Lemma 3.9] for the case of a Schwarz map.

Proposition 6.9 *Let $\gamma : N \to M$ be as in Definition 2.1. For any $y \in N$, $\gamma(y^* y) = \gamma(y)^* \gamma(y)$ if and only if $\gamma(zy) = \gamma(z)\gamma(y)$ for all $z \in N$. Consequently,*

$$\mathcal{M}_\gamma = \{ y \in N : \gamma(zy) = \gamma(z)\gamma(y), \ \gamma(yz) = \gamma(y)\gamma(z) \text{ for all } z \in N \},$$

and it is a von Neumann subalgebra of N.

When γ is a unital normal Schwarz map from M into itself, we also consider the fixed-point set

$$\mathcal{F}_\gamma := \{ x \in M : \gamma(x) = x \}.$$

The set $\mathcal{F}_\gamma$ is not generally a subalgebra of M, and there are no general inclusion relations between $\mathcal{F}_\gamma$ and $\mathcal{M}_\gamma$, see [52, Appendix B]. But we have the next result,

which seems first observed in [74] (see also [11, Theorem 2.3], [19, Lemma 3.4] and [58, Theorem 1 (i)]). We give a proof for convenience.

Lemma 6.10 *Let γ be a unital normal Schwarz map from M into itself. Assume that there exists a faithful φ such that $\varphi \circ \gamma = \varphi$. Then*

$$\mathcal{F}_\gamma = \{x \in M : \gamma(xz) = x\gamma(z),\ \gamma(zx) = \gamma(z)x \text{ for all } z \in M\} \subset \mathcal{M}_\gamma,$$

and hence $\mathcal{F}_\gamma$ is a von Neumann subalgebra of M.

Proof It suffices to prove the first equality. The inclusion $\supset$ is obvious. Conversely, let $x \in \mathcal{F}_\gamma$. Then $x^*x = \gamma(x)^*\gamma(x) \leq \gamma(x^*x)$ and $\varphi(\gamma(x^*x) - x^*x) = \varphi(x^*x) - \varphi(x^*x) = 0$, implying $\gamma(x^*x) = x^*x = \gamma(x)^*\gamma(x)$. Similarly, $\gamma(xx^*) = xx^* = \gamma(x)\gamma(x)^*$. Therefore, $x \in \mathcal{M}_\gamma$ and by Proposition 6.9 we have $\gamma(xz) = x\gamma(z)$ and $\gamma(zx) = \gamma(z)x$ for all $z \in M$. $\square$

The next lemma due to Petz [91] will play a crucial role in the next section.

Lemma 6.11 ([91]) *Let $\gamma : N \to M$ be a unital normal 2-positive map, and $\varphi \in M_*^+$. Assume that both of φ and $\varphi \circ \gamma$ are faithful. Let σ_t^φ be the modular automorphism group of M associated with φ and $\sigma_t^{\varphi \circ \gamma}$ be that of N associated with $\varphi \circ \gamma$ (see Appendix A.2). Then for any $v \in N$ the following conditions are equivalent:*

  (i) *$\gamma(v^*v) = \gamma(v)^*\gamma(v)$ and $\gamma(\sigma_t^{\varphi \circ \gamma}(v)) = \sigma_t^\varphi(\gamma(v))$ for all $t \in \mathbb{R}$;*
  (ii) *$\gamma_\varphi^* \circ \gamma(v) = v$.*

Moreover, $N_1 := \mathcal{F}_{\gamma_\varphi^* \circ \gamma}$ and $M_1 := \mathcal{F}_{\gamma \circ \gamma_\varphi^*}$ are von Neumann subalgebras of N and M, respectively, and $\gamma|_{N_1}$ is an isomorphism from N_1 onto M_1, whose inverse is $\gamma_\varphi^*|_{M_1}$.

Proof (i) $\Longrightarrow$ (ii). Let $\Phi \in \mathcal{P}$ and $\Phi_0 \in \mathcal{P}_0$ be as in Theorem 6.1. Recall that the operator $x\Phi \mapsto x^*\Phi$ ($x \in M$) is closable and its closure has the polar decomposition $J\Delta^{1/2}$. Similarly, the closure of $y\Phi_0 \mapsto y^*\Phi_0$ ($y \in N$) is $J_0\Delta_0^{1/2}$. Then $\sigma_t^\varphi(x) = \Delta^{it}x\Delta^{-it}$ ($x \in M$) and $\sigma_t^{\varphi \circ \gamma}(y) = \Delta_0^{it}y\Delta_0^{-it}$ ($y \in N$), see Sec. A.2. Since γ is a Schwarz map, one has

$$\|\gamma(y)\Phi\|^2 = \varphi(\gamma(y)^*\gamma(y)) \leq \varphi \circ \gamma(y^*y) = \|y\Phi_0\|^2, \qquad y \in N,$$

so that a contraction $V_\varphi : \mathcal{H}_0 \to \mathcal{H}$ can be defined by extending $y\Phi_0 \mapsto \gamma(y)\Phi$ ($y \in N$). By the second assumption of (i) one has $\gamma(\sigma_t^{\varphi \circ \gamma}(v^*)) = \sigma_t^\varphi(\gamma(v^*))$ so that $V_\varphi \sigma_t^{\varphi \circ \gamma}(v^*)\Phi_0 = \sigma_t^\varphi(\gamma(v^*))\Phi$ for all $t \in \mathbb{R}$. Recall that $\sigma_z^{\varphi \circ \gamma}(v^*)$ and $\sigma_z^\varphi(\gamma(v^*))$ are continuous (in the weak operator topology) on $-1/2 \leq \text{Im } z \leq 0$ and analytic in $-1/2 < \text{Im } z < 0$. Hence the analytic continuation of those at $z = -i/2$ gives

$$V_\varphi \Delta_0^{1/2}v^*\Phi_0 = \Delta^{1/2}\gamma(v^*)\Phi,$$

that is,

$$V_\varphi J_0 v\Phi_0 = J\gamma(v)\Phi. \tag{6.3}$$

which implies that

$$\|V_\varphi J_0 v \Phi_0\|^2 = \|\gamma(v)\Phi\|^2 = \varphi(\gamma(v)^*\gamma(v)) = \varphi(\gamma(v^*v)) = \|v\Phi_0\|^2 = \|J_0 v \Phi_0\|^2,$$

where we have used the first assumption of (i). Letting $\zeta_0 := J_0 v \Phi_0 \in \mathcal{H}_0$, since V_φ is a contraction, one has

$$\|V_\varphi^* V_\varphi \zeta_0 - \zeta_0\|^2 = \|V_\varphi^* V_\varphi \zeta_0\|^2 - 2\|V_\varphi \zeta_0\|^2 + \|\zeta_0\|^2 = \|V_\varphi^* V_\varphi \zeta_0\|^2 - \|\zeta_0\|^2 \le 0$$

so that $V_\varphi^* V_\varphi \zeta_0 = \zeta_0$. This and (6.3) give

$$J_0 v \Phi_0 = V_\varphi^* V_\varphi J_0 v \Phi_0 = V_\varphi^* J\gamma(v)\Phi. \tag{6.4}$$

Therefore,

$$\langle J_0 v \Phi_0, y\Phi_0 \rangle = \langle V_\varphi^* J\gamma(v)\Phi, y\Phi_0 \rangle = \langle J\gamma(v)\Phi, \gamma(y)\Phi \rangle, \qquad y \in N, \tag{6.5}$$

which, compared with (6.1), is equivalent to equality in (ii).

   Before proving (ii) $\Longrightarrow$ (i), let us prove the latter assertion. By Proposition 6.3 (1) and (2),

$$\varphi \circ (\gamma \circ \gamma_\varphi^*) = \varphi, \qquad \varphi \circ \gamma \circ (\gamma_\varphi^* \circ \gamma) = \varphi \circ \gamma. \tag{6.6}$$

Since γ is 2-positive, so is γ_φ^* by Proposition 6.3 (3). Hence $\gamma \circ \gamma_\varphi^*$ and $\gamma_\varphi^* \circ \gamma$ are 2-positive. Since φ and $\varphi \circ \gamma$ are faithful by assumption, it follows from Lemma 6.10 that the fixed-point sets M_1 and N_1 are von Neumann subalgebras of M and N, respectively. Since $\gamma(y) = \gamma \circ \gamma_\varphi^* \circ \gamma(y)$ for all $y \in N_1$ and $\gamma_\varphi^*(x) = \gamma_\varphi^* \circ \gamma \circ \gamma_\varphi^*(x)$ for all $x \in M_1$, it is immediate to see that $\gamma|_{N_1}$ maps N_1 to M_1 bijectively and $\gamma_\varphi^*|_{M_1}$ is the inverse of $\gamma|_{N_1}$. Next, assume that $v \in N_1$. Then

$$\begin{aligned}
\langle J_0 v \Phi_0, y\Phi_0 \rangle &= \langle J_0 \gamma_\varphi^*(\gamma(v))\Phi_0, y\Phi_0 \rangle \\
&= \langle J\gamma(v)\Phi, \gamma(y)\Phi \rangle = \langle V_\varphi^* J\gamma(v)\Phi, y\Phi_0 \rangle, \qquad y \in N,
\end{aligned}$$

so that (6.5) holds, which in turn gives $J_0 v \Phi_0 = V_\varphi^* J\gamma(v)\Phi$ as in (6.4). Hence

$$\varphi \circ \gamma(v^*v) = \|v\Phi_0\|^2 = \|V_\varphi^* J\gamma(v)\Phi\|^2 \le \|\gamma(v)\Phi\|^2 = \varphi(\gamma(v)^*\gamma(v)).$$

Since γ is a Schwarz map and φ is faithful, $\gamma(v^*v) = \gamma(v)^*\gamma(v)$ holds. Since $v^* \in N_1$, $\gamma(vv^*) = \gamma(v)\gamma(v)^*$ holds as well. Therefore, $v \in \mathcal{M}_\gamma$ so that $\gamma(yv) = \gamma(y)\gamma(v)$ and $\gamma(vy) = \gamma(v)\gamma(y)$ for all $y \in N$, showing that $\gamma|_{N_1}$ is an isomorphism.

   (ii) $\Longrightarrow$ (i). Assume (ii), i.e., $v \in N_1$. The first equality in (i) was already shown above. As for the second, since $\varphi \circ \gamma|_{N_1} = (\varphi|_{M_1}) \circ (\gamma|_{N_1})$, we find by using the KMS condition (see Appendix A.2) that

$$\gamma(\sigma_t^{\varphi \circ \gamma|_{N_1}}(y)) = \sigma_t^{\varphi|_{M_1}}(\gamma(y)), \qquad y \in N_1, \ t \in \mathbb{R}. \tag{6.7}$$

The mean ergodic theorem says that the norm-limit

$$E(x) := \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} (\gamma \circ \gamma_\varphi^*)^k(x)$$

exists for every $x \in M$ and E is a norm one projection from M onto M_1. By Lemma 6.10 with (6.6) it is immediate that E is a conditional expectation onto M_1 such that $\varphi \circ E = \varphi$. (This is also a consequence of Tomiyama's theorem [112].) The normality of E is also immediate from $\varphi \circ E = \varphi$ since φ is faithful. Therefore, we have $\sigma_t^{\varphi|M_1} = \sigma_t^\varphi|_{M_1}$ ($t \in \mathbb{R}$) by Takesaki's theorem (see Theorem A.10 of Appendix A.2). Similarly, replacing φ and $\gamma \circ \gamma_\varphi^*$ in the above argument with $\varphi \circ \gamma$ and $\gamma_\varphi^* \circ \gamma$, we have $\sigma_t^{\varphi \circ \gamma|N_1} = \sigma_t^{\varphi \circ \gamma}|_{N_1}$ ($t \in \mathbb{R}$). Inserting these into (6.7) gives the second equality in (i). □

Remark 6.12 In the above proof of Lemma 6.11 we need the 2-positivity of γ to guarantee γ_φ^* being a Schwarz map (see Remark 6.4).

6.3 Preservation of Connes' Cocycle Derivatives

Throughout this section, let f be an operator convex function on $(0, \infty)$, having the integral expression (2.5). We write μ_f for the representing measure in (2.5) to specify its dependence on f, and denote its support by supp μ_f, i.e., the topological support of μ_f consisting of $t \in [0, \infty)$ such that $\mu_f((t - \varepsilon, t + \varepsilon)) > 0$ for any $\varepsilon > 0$. Let M, N be von Neumann algebras and $\gamma : N \to M$ be a unital normal *Schwarz* map. Furthermore, let $\psi, \varphi \in M_*^+$ and assume that both of φ and $\varphi \circ \gamma$ are faithful (but ψ is general). There are two main ingredients of the discussions below. One is the standard f-divergence $S_f(\psi\|\varphi)$ and another is Connes' cocycle derivative $[D\psi : D\varphi]_t$ ($t \in \mathbb{R}$), whose concise survey is in Appendix A.7.

In the section we prove two lemmas stating that the preservation of a standard f-divergence implies that of Connes' cocycle derivatives. Those lemmas are the most substantial parts of the proof of the reversibility/sufficiency theorems in the next Sec. 6.4.

Lemma 6.13 *Assume that* supp μ_f *has a limit point in* $(0, \infty)$. *If*

$$S_f(\psi \circ \gamma\|\varphi \circ \gamma) = S_f(\psi\|\varphi) < +\infty, \tag{6.8}$$

then we have

$$\gamma([D(\psi \circ \gamma) : D(\varphi \circ \gamma)]_t) = [D\psi : D\varphi]_t, \qquad t \in \mathbb{R}. \tag{6.9}$$

Proof We are working in the standard forms $(M, \mathcal{H}, J, \mathcal{P})$ and $(N, \mathcal{H}_0, J_0, \mathcal{P}_0)$ as before. Write $\psi_0 := \psi \circ \gamma$ and $\varphi_0 := \varphi \circ \gamma$. Let $\Psi, \Phi \in \mathcal{P}$ and $\Psi_0, \Phi_0 \in \mathcal{P}_0$ be the vector representatives of ψ, φ and ψ_0, φ_0, respectively. Below we divide the proof into several steps.

Step 1. A contraction $V_\varphi : \mathcal{H}_0 \to \mathcal{H}$ is defined by extending $y\Phi_0 \mapsto \gamma(y)\Phi$ ($y \in N$) as in the first part of the proof of Lemma 6.11. For every $y \in N$ one has

$$\|\Delta_{\psi,\varphi}^{1/2} V_\varphi(y\Phi_0)\|^2 = \|\Delta_{\psi,\varphi}^{1/2} \gamma(y)\Phi\|^2 = \|\gamma(y)^*\Psi\|^2 = \psi(\gamma(y)\gamma(y)^*)$$

$$\leq \psi(\gamma(yy^*)) = \psi_0(yy^*) = \|y^*\Psi_0\|^2 = \|\Delta_{\psi_0,\varphi_0}^{1/2}(y\Phi_0)\|^2. \quad (6.10)$$

Note that the operator $\Delta_{\psi,\varphi}^{1/2} V_\varphi|_{N\Phi_0}$ is closable. Indeed, let $y_n \in N$ ($n \in \mathbb{N}$) be such that $y_n\Phi_0 \to 0$ and $\Delta_{\psi,\varphi}^{1/2} V_\varphi(y_n\Phi_0) = \Delta_{\psi,\varphi}^{1/2}\gamma(y_n)\Phi \to \zeta \in \mathcal{H}$. Since

$$\|\gamma(y_n)\Phi\|^2 = \varphi(\gamma(y_n)^*\gamma(y_n)) \leq \varphi(\gamma(y_n^*y_n)) = \|y_n\Phi_0\|^2,$$

we have $\gamma(y_n)\Phi \to 0$. Since $\Delta_{\psi,\varphi}^{1/2}$ is a closed operator, $\zeta = 0$ follows. Thus, $\Delta_{\psi,\varphi}^{1/2} V_\varphi|_{N\Phi_0}$ has the closure T and (6.10) means that $\|T(y\Phi_0)\| \leq \|\Delta_{\psi_0,\varphi_0}^{1/2}(y\Phi_0)\|$ for all $y \in N$. Since $N\Phi_0$ is a core of $\Delta_{\psi_0,\varphi_0}^{1/2}$, we see by Proposition B.2 of Appendix B that

$$T^*T \leq \Delta_{\psi_0,\varphi_0} \quad (6.11)$$

in the sense of Proposition B.1. Let $\Delta_{\psi,\varphi} = \int_0^\infty t \, dE_{\psi,\varphi}(t)$ be the spectral decomposition of $\Delta_{\psi,\varphi}$, and for each $n \in \mathbb{N}$ let $H_n := \int_0^n t \, dE_{\psi,\varphi}(t)$. Since $N\Phi_0$ is a core of T and

$$\|H_n^{1/2} V_\varphi(y\Phi_0)\| = \|H_n^{1/2}(\gamma(y)\Phi)\| \leq \|\Delta_{\psi,\varphi}^{1/2}(\gamma(y)\Phi)\| = \|T(y\Phi_0)\|, \quad y \in N,$$

we have $V_\varphi^* H_n V_\varphi \leq T^*T$ by Proposition B.2 again. This and (6.11) show that $V_\varphi^* H_n V_\varphi \leq \Delta_{\psi_0,\varphi_0}$. Therefore, for every $s > 0$,

$$(s1 + \Delta_{\psi_0,\varphi_0})^{-1} \leq (s1 + V_\varphi^* H_n V_\varphi)^{-1} \leq V_\varphi^*(s1 + H_n)^{-1}V_\varphi,$$

where the latter inequality above is a consequence of Hansen's inequality [43] since $t \mapsto (s+t)^{-1}$ is an operator monotone decreasing function on $[0, \infty)$. Letting $n \to \infty$ gives

$$(s1 + \Delta_{\psi_0,\varphi_0})^{-1} \leq V_\varphi^*(s1 + \Delta_{\psi,\varphi})^{-1}V_\varphi, \qquad s > 0. \quad (6.12)$$

Step 2. Write expression (2.5) as

$$f(t) = a + b(t - 1) + cf_2(t) + \int_{[0,\infty)} g_s(t) \, d\mu_f(s), \qquad t \in (0, \infty), \quad (6.13)$$

where $f_2(t) := (t - 1)^2$ and $g_s(t) := (t - 1)^2(t + s)^{-1}$ for $s \in [0, \infty)$. Note that

$$S_f(\psi\|\varphi) = (a-b)\varphi(1) + b\psi(1) + cS_{f_2}(\psi\|\varphi)$$
$$+ \int_{[0,\infty)} S_{g_s}(\psi\|\varphi)\,d\mu_f(s) < +\infty, \tag{6.14}$$

$$S_f(\psi_0\|\varphi_0) = (a-b)\varphi_0(1) + b\psi_0(1) + cS_{f_2}(\psi_0\|\varphi_0)$$
$$+ \int_{[0,\infty)} S_{g_s}(\psi_0\|\varphi_0)\,d\mu_f(s), \tag{6.15}$$

and

$$\varphi_0(1) = \varphi(1), \qquad \psi_0(1) = \psi(1),$$
$$S_{f_2}(\psi_0\|\varphi_0) \le S_{f_2}(\psi\|\varphi), \qquad S_{g_s}(\psi_0\|\varphi_0) \le S_{g_s}(\psi\|\varphi) \tag{6.16}$$

by the monotonicity property in Theorem 2.7 (iv). Therefore, the assumption (6.8) implies that

$$S_{h_s}(\psi_0\|\varphi_0) = S_{h_s}(\psi\|\varphi) \quad \text{for } \mu_f\text{-a.e. } s \in [0,\infty).$$

Write

$$g_s(t) = \frac{((t+s)-(1+s))^2}{t+s} = t - (2+s) + (1+s)^2 h_s(t), \qquad t \in (0,\infty),$$

where $h_s(t) := (t+s)^{-1}$ for $s \in [0,\infty)$. Since

$$S_{g_s}(\psi\|\varphi) = \psi(1) - (2+s)\varphi(1) + (1+s)^2 S_{h_s}(\psi\|\varphi), \tag{6.17}$$
$$S_{g_s}(\psi_0\|\varphi_0) = \psi_0(1) - (2+s)\varphi_0(1) + (1+s)^2 S_{h_s}(\psi_0\|\varphi_0), \tag{6.18}$$

we have

$$S_{\varphi_s}(\psi_0\|\varphi_0) = S_{\varphi_s}(\psi\|\varphi) \quad \text{for } \mu_f\text{-a.e. } s \in [0,\infty),$$

that is,

$$\langle \Phi_0, (s1 + \Delta_{\psi_0,\varphi_0})^{-1}\Phi_0 \rangle = \langle \Phi, (s1 + \Delta_{\psi,\varphi})^{-1}\Phi \rangle \quad \text{for } \mu_f\text{-a.e. } s \in [0,\infty). \tag{6.19}$$

We easily see that

$$\langle \Phi, (z1 + \Delta_{\psi,\varphi})^{-1}\Phi \rangle = \int_0^\infty \frac{1}{z+t}\,d\|E_{\psi,\varphi}(t)\Phi\|^2$$

is analytic in $\{z \in \mathbb{C} : \operatorname{Re} z > 0\}$, and similarly for $\langle \Phi_0, (z1 + \Delta_{\psi_0,\varphi_0})^{-1}\Phi_0 \rangle$. From the assumption on $\operatorname{supp}\mu_f$ and (6.19) it follows that the set of $s > 0$ for which equality in (6.19) holds has a limit point in $(0,\infty)$. Thus, the coincidence theorem for analytic functions yields that for all $s > 0$,

$$\langle \Phi_0, (s1 + \Delta_{\psi_0,\varphi_0})^{-1}\Phi_0 \rangle = \langle \Phi, (s1 + \Delta_{\psi,\varphi})^{-1}\Phi \rangle$$

and hence

$$\langle \Phi_0, [V_\varphi^*(s1 + \Delta_{\psi,\varphi})^{-1}V_\varphi - (s1 + \Delta_{\psi_0,\varphi_0})^{-1}]\Phi_0\rangle = 0,$$

thanks to $V_\varphi \Phi_0 = \Phi$. By (6.12) this implies that

$$V_\varphi^*(s1 + \Delta_{\psi,\varphi})^{-1}V_\varphi \Phi_0 = (s1 + \Delta_{\psi_0,\varphi_0})^{-1}\Phi_0,$$

and therefore

$$V_\varphi^*(s1 + \Delta_{\psi,\varphi})^{-1}\Phi = (s1 + \Delta_{\psi_0,\varphi_0})^{-1}\Phi_0, \qquad s > 0. \tag{6.20}$$

Step 3. Let $C_0[0, \infty)$ denote the Banach space of continuous complex functions ϕ on the locally compact space $[0, \infty)$ vanishing at infinity (i.e., $\phi(\infty) = 0$) with the sup-norm. Noting that $h_s \in C_0[0, \infty)$ for $s \in (0, \infty)$, we define $\mathfrak{A}$ to be the norm-closed complex linear span of h_s, $s \in (0, \infty)$, in $C_0[0, \infty)$. Since

$$h_{s_1} h_{s_2} = \frac{1}{s_2 - s_1}(h_{s_1} - h_{s_2}) \quad \text{for } s_1 \neq s_2,$$

$$h_s^2 = \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon}(h_s - h_{s+\varepsilon}) \text{ (in the norm)} \quad \text{for } s > 0,$$

we see that for every $c_i, c_j' \in \mathbb{C}$ and $s_i, s_j' \in (0, \infty)$,

$$\left(\sum_{i=1}^m c_i h_{s_i}\right)\left(\sum_{j=1}^n c_j' h_{s_j'}\right) = \sum_{i,j} c_i c_j' h_{s_i} h_{s_j'} \in \mathfrak{A}.$$

This means that $\mathfrak{A}$ is a closed subalgebra of $C_0[0, \infty)$. Obviously, $h_s(t) > 0$ for all $t \in [0, \infty)$, and $h_s(t_1) \neq h_s(t_2)$ for every $t_1, t_2 \in [0, \infty)$ with $t_1 \neq t_2$. Hence the Stone-Weierstrass theorem implies that $\mathfrak{A} = C_0[0, \infty)$. Since

$$\phi \in C_0[0, \infty) \longmapsto V_\varphi^*\phi(\Delta_{\psi,\varphi})\Phi, \; \phi(\Delta_{\psi_0,\varphi_0})\Phi_0 \in \mathcal{H}_0$$

are bounded linear maps with respect to the sup-norm on $C_0[0, \infty)$ and the norm on $\mathcal{H}_0$, it follows from (6.20) that

$$V_\varphi^*\phi(\Delta_{\psi,\varphi})\Phi = \phi(\Delta_{\psi_0,\varphi_0})\Phi_0, \qquad \phi \in C_0[0, \infty), \tag{6.21}$$

Step 4. For every $s > 0$ one has by (6.20)

$$\begin{aligned}
\|V_\varphi^*(s1 + \Delta_{\psi,\varphi})^{-1}\Phi\|^2 &= \langle(s1 + \Delta_{\psi_0,\varphi_0})^{-1}\Phi_0, (s1 + \Delta_{\psi_0,\varphi_0})^{-1}\Phi_0\rangle \\
&= \langle(s1 + \Delta_{\psi_0,\varphi_0})^{-2}\Phi_0, \Phi_0\rangle \\
&= \langle V_\varphi^*(s1 + \Delta_{\psi,\varphi})^{-2}\Phi, \Phi_0\rangle \quad \text{(by (6.21))} \\
&= \langle(s1 + \Delta_{\psi,\varphi})^{-2}\Phi, \Phi\rangle \quad \text{(since } V_\varphi \Phi_0 = \Phi) \\
&= \|(s1 + \Delta_{\psi,\varphi})^{-1}\Phi\|^2.
\end{aligned}$$

Letting $\zeta := (s1 + \Delta_{\psi,\varphi})^{-1}\Phi \in \mathcal{H}$, since the above means $\|V_\varphi^* \zeta\|^2 = \|\zeta\|^2$, one has

$$\|V_\varphi V_\varphi^* \zeta - \zeta\|^2 = \|V_\varphi V_\varphi^* \zeta\|^2 - 2\|V_\varphi^* \zeta\|^2 + \|\zeta\|^2 = \|V_\varphi V_\varphi^* \zeta\|^2 - \|\zeta\|^2 \le 0$$

so that $V_\varphi V_\varphi^* \zeta = \zeta$, i.e.,

$$V_\varphi V_\varphi^* (s1 + \Delta_{\psi,\varphi})^{-1}\Phi = (s1 + \Delta_{\psi,\varphi})^{-1}\Phi.$$

By (6.20) this implies that

$$V_\varphi (s1 + \Delta_{\psi_0,\varphi_0})^{-1}\Phi_0 = (s1 + \Delta_{\psi,\varphi})^{-1}\Phi, \qquad s > 0.$$

Hence, similarly to Step 3, we have

$$V_\varphi \phi(\Delta_{\psi_0,\varphi_0})\Phi_0 = \phi(\Delta_{\psi,\varphi})\Phi, \qquad \phi \in C_0[0,\infty). \tag{6.22}$$

Step 5. For each $t \in \mathbb{R}$ set

$$\phi(x) := x^{it} \quad (x > 0), \qquad \phi(0) := 0. \tag{6.23}$$

For $n \in \mathbb{N}$ let

$$\kappa_n(x) := \begin{cases} nx & (0 \le x \le 1/n), \\ 1 & (1/n \le x \le n), \\ 1 - n(x-n) & (n \le x \le n + \frac{1}{n}), \\ 0 & (x \ge n + \frac{1}{n}), \end{cases} \tag{6.24}$$

and $\phi_n(x) := \kappa_n(x)\phi(x)$ for $x \in [0,\infty)$. Let $s(\psi), s(\psi_0)$ be the support projections of ψ, ψ_0, respectively. Note that the support projection of $\Delta_{\psi,\varphi}$ is $s(\psi)$ and that of $\Delta_{\psi_0,\varphi_0}$ is $s(\psi_0)$, since φ, φ_0 are faithful. Then it is immediate to verify that $\phi_n \in C_0[0,\infty)$ and

$$\phi_n(\Delta_{\psi,\varphi}) \longrightarrow \phi(\Delta_{\psi,\varphi}) = s(\psi)\Delta_{\psi,\varphi}^{it} \quad \text{strongly,}$$

$$\phi_n(\Delta_{\psi_0,\varphi_0}) \longrightarrow \phi(\Delta_{\psi_0,\varphi_0}) = s(\psi_0)\Delta_{\psi_0,\varphi_0}^{it} \quad \text{strongly,}$$

By (6.21), $V_\varphi \phi_n(\Delta_{\psi_0,\varphi_0})\Phi_0 = \phi_n(\Delta_{\psi,\varphi})\Phi$. Letting $n \to \infty$ gives

$$V_\varphi (s(\psi_0)\Delta_{\psi_0,\varphi_0}^{it}\Phi_0) = s(\psi)\Delta_{\psi,\varphi}^{it}\Phi.$$

Since $\Delta_\varphi^{-it}\Phi = \Phi$ and $\Delta_{\varphi_0}^{-it}\Phi_0 = \Phi_0$, where $\Delta_\varphi = \Delta_{\varphi,\varphi}$ is the modular operator associated with φ, we have

$$V_\varphi (s(\psi_0)\Delta_{\psi_0,\varphi_0}^{it}\Delta_{\varphi_0}^{-it}\Phi_0) = s(\psi)\Delta_{\psi,\varphi}^{it}\Delta_\varphi^{-it}\Phi. \tag{6.25}$$

From Proposition A.53 and Remark A.54 we write

$$[D\psi : D\varphi]_t = s(\psi)\Delta_{\psi,\varphi}^{it}\Delta_{\varphi}^{-it}, \quad [D\psi_0 : \varphi_0]_t = s(\psi_0)\Delta_{\psi_0,\varphi_0}^{it}\Delta_{\varphi_0}^{-it}, \qquad t \in \mathbb{R}.$$

(Normally, $s(\psi)$ is not written in the above, since $\Delta_{\psi,\varphi}^{it}$ is defined with restriction on $s(\psi)\mathcal{H}$.) Hence (6.25) means that

$$V_{\varphi}([D\psi_0 : D\varphi_0]_t\Phi_0) = [D\psi : D\varphi]_t\Phi,$$

that is,

$$\gamma([D\psi_0 : D\varphi_0]_t)\Phi = [D\psi : D\varphi]_t\Phi.$$

Since Φ is separating for M (i.e., $x \in M$, $x\Phi = 0 \implies x = 0$), we arrive at (6.9). $\square$

Lemma 6.14 *Assume that γ is 2-positive. If* (6.9) *holds, then we have*

$$\gamma_{\varphi}^*([D\psi : D\varphi]_t) = [D(\psi \circ \gamma) : D(\varphi \circ \gamma)]_t, \qquad t \in \mathbb{R}, \qquad (6.26)$$

$$\gamma \circ \gamma_{\varphi}^*([D\psi : D\varphi]_t) = [D\psi : D\varphi]_t, \qquad t \in \mathbb{R}, \qquad (6.27)$$

$$\psi \circ \gamma \circ \gamma_{\varphi}^* = \psi. \qquad (6.28)$$

Proof We write $\varphi_0 := \varphi \circ \gamma$, $\psi_0 := \psi \circ \gamma$, $u_t := [D\psi : D\varphi]_t$ and $v_t := [D\psi_0 : D\varphi_0]_t$ ($t \in \mathbb{R}$). The assumption says that $\gamma(v_t) = u_t$ for all $t \in \mathbb{R}$. From properties of Connes' cocycle derivatives in Theorem A.50 we have

$$\varphi(\gamma(v_t)^*\gamma(v_t)) = \varphi(u_t^*u_t) = \varphi(\sigma_t^{\varphi}(u_0)) = \varphi(u_0) = \varphi(\gamma(v_0)) = \varphi_0(v_0),$$
$$\varphi(\gamma(v_t^*v_t)) = \varphi_0(v_t^*v_t) = \varphi_0(\sigma_t^{\varphi_0}(v_0)) = \varphi_0(v_0), \qquad t \in \mathbb{R}.$$

Since γ is a Schwarz map and φ is faithful, the above implies that

$$\gamma(v_t^*v_t) = \gamma(v_t)^*\gamma(v_t), \qquad t \in \mathbb{R}. \qquad (6.29)$$

From the cocycle property of v_t, u_t and Proposition 6.9 we find that

$$\gamma(\sigma_s^{\varphi_0}(v_t)) = \gamma(v_s^*v_{s+t}) = \gamma(v_s)^*\gamma(v_{s+t})$$
$$= u_s^*u_{s+t} = \sigma_s^{\varphi}(u_t) = \sigma_s^{\varphi}(\gamma(v_t)), \qquad s, t \in \mathbb{R}. \qquad (6.30)$$

By (6.29) and (6.30) it follows from Lemma 6.11 that $v_t \in N_1$ so that $u_t = \gamma(v_t) \in M_1$ and $\gamma_{\varphi}^*(u_t) = v_t$ for all $t \in \mathbb{R}$, where N_1, M_1 are as given in Lemma 6.11. Therefore, $\gamma \circ \gamma_{\varphi}^*(u_t) = \gamma(v_t) = u_t$ for all $t \in \mathbb{R}$, so (6.26) and (6.27) have been verified. Moreover, since $\gamma_{\varphi}^*(u_t^*u_t) = \gamma_{\varphi}^*(u_t)^*\gamma_{\varphi}^*(u_t)$ and $\gamma_{\varphi}^*(u_tu_t^*) = \gamma_{\varphi}^*(u_t)\gamma_{\varphi}^*(u_t)^*$ by Lemma 6.11 again together with Lemma 6.10, note that $u_t \in \mathcal{M}_{\gamma_{\varphi}^*}$ as well as $v_t \in \mathcal{M}_{\gamma}$ for all $t \in \mathbb{R}$ (see Definition 6.8).

Now, let Φ, Φ_0 be the vector representatives of φ, φ_0, respectively, as in Theorem 6.1. For every $x \in M$, $y \in N$ and $s, t \in \mathbb{R}$, thanks to the identities $\Delta_{\psi_0,\varphi_0}^{it} = v_t\Delta_{\varphi_0}^{it}$ and $\Delta_{\psi,\varphi}^{it} = u_t\Delta_{\varphi}^{it}$ (Proposition A.53 and Remark A.54), we obtain

$$\langle J_0 y \Delta^{is}_{\psi_0,\varphi_0} \Phi_0, \gamma^*_\varphi(x) \Delta^{it}_{\psi_0,\varphi_0} \Phi_0 \rangle = \langle J_0 y v_s \Phi_0, \gamma^*_\varphi(x) v_t \Phi_0 \rangle$$

$$= \langle J_0 y v_s \Phi_0, \gamma^*_\varphi(x) \gamma^*_\varphi(u_t) \Phi_0 \rangle$$

$$= \langle J_0 y v_s \Phi_0, \gamma^*_\varphi(x u_t) \Phi_0 \rangle \quad (\text{since } u_t \in \mathcal{M}_{\gamma^*_\varphi})$$

$$= \langle J_0 \gamma^*_\varphi(x u_t) \Phi_0, y v_s \Phi_0 \rangle$$

$$= \langle J x u_t \Phi, \gamma(y v_s) \Phi \rangle \quad (\text{see } (6.1))$$

$$= \langle J \gamma(y v_s) \Phi, x u_t \Phi \rangle$$

$$= \langle J \gamma(y) u_s \Phi, x u_t \Phi \rangle \quad (\text{since } v_s \in \mathcal{M}_\gamma)$$

$$= \langle J \gamma(y) \Delta^{is}_{\psi,\varphi} \Phi, x \Delta^{it}_{\psi,\varphi} \Phi \rangle.$$

By analytic continuation of the above both sides twice at $s = -i/2$ and then at $t = -i/2$, it follows that

$$\langle J_0 y \Delta^{1/2}_{\psi_0,\varphi_0} \Phi_0, \gamma^*_\varphi(x) \Delta^{1/2}_{\psi_0,\varphi_0} \Phi_0 \rangle = \langle J \gamma(y) \Delta^{1/2}_{\psi,\varphi} \Phi, x \Delta^{1/2}_{\psi,\varphi} \Phi \rangle,$$

that is,

$$\langle J_0 y \Psi_0, \gamma^*_\varphi(x) \Psi_0 \rangle = \langle J \gamma(y) \Psi, x \Psi \rangle, \qquad x \in M, \ y \in N, \tag{6.31}$$

where Ψ, Ψ_0 are the vector representatives of ψ, ψ_0, respectively. Taking $y = 1$ in (6.31) yields $\psi_0 \circ \gamma^*_\varphi(x) = \psi(x)$ for all $x \in M$, showing (6.28). $\square$

Remark 6.15 Assume in Lemma 6.14 that ψ is faithful too. Then from Theorem A.50 one has $u_0 = s(\psi) = 1$ and $s(\psi_0) = v_0 = \gamma^*_\varphi(u_0) = 1$. Hence $\psi_0 = \psi \circ \gamma$ is also faithful automatically. Therefore, equality (6.31) means that γ^*_φ is Petz' recovery map of γ with respect to ψ, so we have $\gamma^*_\psi = \gamma^*_\varphi$.

6.4 Reversibility via Standard f-Divergences

Throughout the section, assume that $\gamma : N \to M$ is a unital normal 2-*positive* map between von Neumann algebras. For given $\psi, \varphi \in M^+_*$ we say that γ is *reversible* for $\{\psi, \varphi\}$ if there exists a unital normal 2-positive map $\beta : M \to N$ such that

$$\psi \circ \gamma \circ \beta = \psi \quad \text{and} \quad \varphi \circ \gamma \circ \beta = \varphi.$$

When M_0 is a von Neumann subalgebra of M, we say that M_0 is *sufficient* for $\{\psi, \varphi\}$ if the inclusion map $M_0 \hookrightarrow M$ is reversible for $\{\psi, \varphi\}$, i.e., there exists a unital normal 2-positive map $\beta : M \to M_0$ such that $\psi \circ \beta = \psi$ and $\varphi \circ \beta = \varphi$.

Let $\Psi, \Phi \in \mathcal{P}$ be the vector representatives of ψ, φ, respectively, in the standard form $(M, \mathcal{H}, J, \mathcal{P})$. The *transition probability* of $\psi, \varphi \in M^+_*$ is

$$P(\psi, \varphi) := \langle \Psi, \Phi \rangle = -S_{-t^{1/2}}(\psi \| \varphi), \tag{6.32}$$

that already appeared in (3.25). Note that $P(\psi, \varphi)$ has the (reverse) monotonicity property under unital normal Schwarz maps (by Theorem 2.7 (iv)).

The next theorem was originally proved by Petz [91, Theorem 5] under the assumption that φ, $\varphi \circ \gamma$, ψ and $\psi \circ \gamma$ are all faithful. The faithfulness assumption of ψ and $\psi \circ \gamma$ was later removed in Jenčová and Petz [62, Theorem 3].

Theorem 6.16 ([91, 62]) *Let $\psi, \varphi \in M_*^+$ and assume that φ and $\varphi \circ \gamma$ are faithful. Then the following conditions (i)–(iv) are equivalent:*

(i) $\psi \circ \gamma \circ \gamma_\varphi^* = \psi$ ($\varphi \circ \gamma \circ \gamma_\varphi^* = \varphi$ is automatic by Proposition 6.3 (1));
(ii) γ is reversible for $\{\psi, \varphi\}$;
(iii) $P(\psi, \varphi) = P(\psi \circ \gamma, \varphi \circ \gamma)$;
(iv) $[D\psi : D\varphi]_t = \gamma([D(\psi \circ \gamma) : D(\varphi \circ \gamma)]_t)$ *for all $t \in \mathbb{R}$.*

Moreover, if ψ and $\psi \circ \gamma$ are faithful too, then the above (i)–(iv) are equivalent to the following:

(v) $[D\psi : D\varphi]_t = \gamma \circ \gamma_\varphi^*([D\psi : D\varphi]_t)$ *for all $t \in \mathbb{R}$;*
(vi) $\gamma_\psi^* = \gamma_\varphi^*$.

Proof (i) $\Longrightarrow$ (ii) is obvious. In view of (6.32), (ii) $\Longrightarrow$ (iii) is immediate from the monotonicity property of the standard f-divergence.

(iii) $\Longrightarrow$ (iv). Condition (iii) is rephrased as

$$S_{-t^{1/2}}(\psi \| \varphi) = S_{-t^{1/2}}(\psi \circ \gamma \| \varphi \circ \gamma),$$

whose value is finite. Recall the well-known integral expression of $-t^{1/2}$

$$-t^{1/2} = -t + \frac{1}{\pi} \int_{(0,\infty)} \left(\frac{t}{1+s} - \frac{t}{t+s} \right) s^{-1/2} \, ds,$$

which is rewritten[1] in the form of (2.5) as

$$-t^{1/2} = -1 - \frac{1}{2}(t-1) + \frac{1}{\pi} \int_{(0,\infty)} \frac{(t-1)^2}{t+s} \frac{s^{1/2}}{(1+s)^2} \, ds,$$

so that the support of the representing measure of $-t^{1/2}$ is $(0, \infty)$. Hence (iii) $\Longrightarrow$ (iv) is a special case of Lemma 6.13.

(iv) $\Longrightarrow$ (i) is contained in Lemma 6.14.

Next, assume that ψ and $\psi \circ \gamma$ are also faithful. (iv) $\Longrightarrow$ (v) is contained in Lemma 6.14, and (iv) $\Longrightarrow$ (vi) was shown in the proof of Lemma 6.14 as noted in Remark 6.15. (vi) $\Longrightarrow$ (i) is obvious by Proposition 6.3 (1).

---

[1] This can be checked by a direct computation

$$\int_0^\infty \left[\left(\frac{t}{1+s} - \frac{t}{t+s} \right) s^{-1/2} - \frac{(t-1)^2}{t+s} \frac{s^{1/2}}{(1+s)^2} \right] ds$$

$$= (t-1) \int_0^\infty \frac{ds}{(1+s)^2 s^{1/2}} = (t-1) \int_0^\infty \frac{2ds}{(1+s^2)^2} = \frac{\pi}{2}(t-1).$$

It remains to prove that (v) $\Longrightarrow$ (iii). Let $\mathfrak{M}$ be the von Neumann subalgebra of M generated by $\{[D\psi : D\varphi]_t : t \in \mathbb{R}\}$. From the cocycle property of $[D\psi : D\varphi]_t$ (see Theorem A.50), $\mathfrak{M}$ is globally invariant under the modular automorphism σ^φ. Hence by Theorem A.10 (Appendix A.2) says that there exists the conditional expectation E from M onto $\mathfrak{M}$ with respect to φ. Furthermore, since $\sigma_t^\varphi|_{\mathfrak{M}} = \sigma_t^{\varphi|_{\mathfrak{M}}}$ fo all $t \in \mathbb{R}$ by Theorem A.10, $[D\psi : D\varphi]_t$ is a $\sigma^{\varphi|_{\mathfrak{M}}}$-cocycle in $\mathfrak{M}$. Therefore, Connes' inverse theorem (see Theorem A.57) implies that there exists a (unique faithful) normal semifinite weight $\tilde{\psi}$ on $\mathfrak{M}$ such that

$$[D\psi : D\varphi]_t = [D\tilde{\psi} : D(\varphi|_{\mathfrak{M}})]_t, \qquad t \in \mathbb{R}. \tag{6.33}$$

On the other hand, by Proposition A.52[2] one has

$$[D(\tilde{\psi} \circ E) : D\varphi]_t = [D(\tilde{\psi} \circ E) : D((\varphi|_{\mathfrak{M}}) \circ E)]_t$$
$$= [D\tilde{\psi} : D(\varphi|_{\mathfrak{M}})]_t. \qquad t \in \mathbb{R}. \tag{6.34}$$

Combining (6.33) and (6.34) implies by Proposition A.51 (3)[3] that $\psi = \tilde{\psi} \circ E$ and hence $\psi \circ E = \psi$.

Now, let $N_1 := \mathcal{F}_{\gamma_\varphi^* \circ \gamma}$ and $M_1 := \mathcal{F}_{\gamma \circ \gamma_\varphi^*}$ (see Lemma 6.11) and assume (v), that is, $\mathfrak{M} \subset M_1$. Then one has

$$\begin{aligned} P(\psi|_{M_1}, \varphi|_{M_1}) &\leq P(\psi|_{\mathfrak{M}}, \varphi|_{\mathfrak{M}}) \\ &\leq P(\psi, \varphi) \quad (\text{since } \psi = (\psi|_{\mathfrak{M}}) \circ E, \varphi = (\varphi|_{\mathfrak{M}}) \circ E) \tag{6.35} \\ &\leq P(\psi \circ \gamma, \varphi \circ \gamma) \\ &\leq P(\psi \circ \gamma|_{N_1}, \varphi \circ \gamma|_{N_1}) \\ &= P(\psi|_{M_1}, \varphi|_{M_1}), \end{aligned}$$

where the last equality holds since $\gamma|_{N_1} : N_1 \to M_1$ is an isomorphism (Lemma 6.11). Therefore, (iii) follows. $\qquad\square$

The following is the specialization of Theorem 6.16 to the case of a von Neumann subalgebra. The corollary was first given in [91] under the assumption that both of φ, ψ are faithful. The faithfulness assumption of ψ was removed in [62, Theorem 1].

Corollary 6.17 ([91, 62]) *Let M_0 be a von Neumann subalgebra of M, and $\psi, \varphi \in M_*^+$ with φ faithful. Then the following conditions (i)–(iv) are equivalent:*

(i) *$\psi \circ E_\varphi = \psi$, where $E_\varphi : M \to M_0$ is the generalized conditional expectation with respect to φ (see the paragraph just before Proposition 6.5);*
(ii) *M_0 sufficient for $\{\psi, \varphi\}$;*
(iii) *$P(\psi, \varphi) = P(\psi|_{M_0}, \varphi|_{M_0})$;*
(iv) *$[D\psi : D\varphi]_t = [D(\psi|_{M_0}) : D(\varphi|_{M_0})]_t$ for all $t \in \mathbb{R}$;*

---

[2] The proposition holds also for faithful normal semifinite weights.

[3] This again holds also for faithful normal semifinite weights.

Moreover, if ψ is faithful too, then the above (i)–(iv) are also equivalent to the following:

(v) $[D\psi : D\varphi]_t \in M_0$ *for all* $t \in \mathbb{R}$;
(vi) $E_\psi = E_\varphi$, *where* $E_\psi : M \to M_0$ *is the generalized conditional expectation with respect to* ψ.

Proof In view of Proposition 6.5, all the conditions of the corollary, except (v), exactly correspond to those in Theorem 6.16 in this specialized case. Condition (v) of Theorem 6.16 means that $[D\psi, D\varphi]_t = E_\varphi([D\psi, D\varphi]_t)$ for all $t \in \mathbb{R}$, which obviously implies (v) of the corollary. Conversely, assume that ψ is also faithful and (v) of the corollary holds. Let $\mathfrak{M}$ be the von Neumann subalgebra of M generated by $\{[D\psi : d\varphi]_t : t \in \mathbb{R}\}$; then $\mathfrak{M} \subset M_0$ by (v). In a similar way to the proof of (v) $\Longrightarrow$ (iii) of Theorem 6.16, there exists the conditional expectation E from M to $\mathfrak{M}$ with respect to φ, and we have $\psi \circ E = \psi$. We hence have

$$P(\psi|_{\mathfrak{M}}, \varphi|_{\mathfrak{M}}) \leq P(\psi, \varphi) \leq P(\psi|_{M_0}, \varphi|_{M_0}) \leq P(\psi|_{\mathfrak{M}}, \varphi|_{\mathfrak{M}}),$$

where the above first inequality is as in (6.35). Therefore, (iii) of the corollary follows.   □

Problem 6.18 Although [62, Theorem 1] asserts (without proof) that (v) of Corollary 6.17 is also equivalent to (i)–(iv) even when ψ is not faithful, it is unknown to us whether this is really the case. In the proof of (v) $\Longrightarrow$ (iii) of Theorem 6.16 and Corollary 6.17, we have used the result that if $E : M \to M_0$ is a faithful normal conditional expectation from M onto its von Neumann subalgebra M_0 and φ_1, φ_2 are faithful normal semifinite weights on M_0, then

$$[D(\varphi_1 \circ E) : D(\varphi_2 \circ E)]_t = [D\varphi_1 : D\varphi_2]_t, \qquad t \in \mathbb{R}. \tag{6.36}$$

The above question holds true if (6.36) were true even when φ_1 is not faithful.

We now present the main reversibility theorem via standard f-divergences.

Theorem 6.19 *Let $\psi, \varphi \in M_*^+$ with $\psi \neq 0$ and assume that $s(\psi) \leq s(\varphi)$. Then the following conditions are equivalent:*

(i) γ *is reversible for* $\{\psi, \varphi\}$;
(ii) $S_f(\psi\|\varphi) = S_f(\psi \circ \gamma\|\varphi \circ \gamma)$ *for every operator convex function on* $(0, \infty)$;
(iii) $S_f(\psi\|\varphi) = S_f(\psi \circ \gamma\|\varphi \circ \gamma) < +\infty$ *for some operator convex function on* $(0, \infty)$ *such that* supp μ_f *has a limit point in* $(0, \infty)$;
(iv) $D_\alpha(\psi\|\varphi) = D_\alpha(\psi \circ \gamma\|\varphi \circ \gamma)$ *for some* $\alpha \in (0, 1)$, *where $D_\alpha(\psi\|\varphi)$ is the α-Rényi divergence (see Sec. 3.1);*
(v) $P(\psi, \varphi) = P(\psi \circ \gamma, \varphi \circ \gamma)$.

Proof (i) $\Longrightarrow$ (ii) is immediate from the monotonicity property of S_f. (ii) $\Longrightarrow$ (iii) is obvious. When $\alpha \in (0, 1)$, note that $-t^\alpha$ is an operator convex function on $(0, \infty)$, the support of whose representing measure is $(0, \infty)$ as that of $-t^{1/2}$ mentioned in the

proof of (iii) $\implies$ (iv) of Theorem 6.16. By the definition of D_α note that condition (iv) is equivalent to $S_{-f_\alpha}(\psi\|\varphi) = S_{-f_\alpha}(\psi \circ \gamma\|\varphi \circ \gamma)$, whose value is always finite. Hence it is clear that (ii) $\implies$ (iv) and (iv) $\implies$ (iii). Moreover, (v) is equivalent to (iv) with $\alpha = 1/2$. After all, it only remains to prove that (iii) implies (i).

(iii) $\implies$ (i). Write $\psi_0 := \psi \circ \gamma$, $\varphi_0 := \varphi \circ \gamma$, $e := s(\varphi)$ ($\in M$) and $e_0 := s(\varphi_0)$ ($\in N$). Define

$$\hat{\gamma} := e\gamma(\cdot)e|_{e_0 N e_0} : e_0 N e_0 \longrightarrow eMe.$$

Since

$$\varphi(\gamma(e_0)) = \varphi_0(e_0) = \varphi_0(1) = \varphi(\gamma(1)) = \varphi(1),$$

one has $\varphi(1 - \gamma(e_0)) = 0$ so that $e(1 - \gamma(e_0))e = 0$. Hence $\hat{\gamma}(e_0) = e$, i.e., $\hat{\gamma}$ is unital. It is clear that $\hat{\gamma}$ is normal and 2-positive. Since $s(\psi) \leq s(\varphi) = e$ and $s(\psi_0) \leq s(\varphi_0) = e_0$, note that

$$\psi_0|_{e_0 N e_0} = (\psi|_{eMe}) \circ \hat{\gamma}, \qquad \varphi_0|_{e_0 N e_0} = (\varphi|_{eMe}) \circ \hat{\gamma}. \tag{6.37}$$

Hence, for any operator convex function on $(0, \infty)$, we have

$$S_f(\psi_0\|\varphi_0) = S_f(\psi_0|_{e_0 N e_0}\|\varphi_0|_{e_0 N e_0}) = S_f((\psi|_{eMe}) \circ \hat{\gamma}\|(\varphi|_{eMe}) \circ \hat{\gamma}), \tag{6.38}$$

as well as

$$S_f(\psi\|\varphi) = S_f(\psi|_{eMe}\|\varphi|_{eMe}). \tag{6.39}$$

Now, assume (iii). Then by (6.38) and (6.39), for f in condition (iii) we obtain

$$S_f(\psi|_{eMe}\|\varphi|_{eMe}) = S_f((\psi|_{eMe}) \circ \hat{\gamma}\|(\varphi|_{eMe}) \circ \hat{\gamma}) < +\infty.$$

Since $\varphi|_{eMe}$ and $(\varphi|_{eMe}) \circ \hat{\gamma}$ are faithful, it follows from Lemma 6.13 and (iv) $\implies$ (i) of Theorem 6.16 that

$$\psi|_{eMe} = (\psi|_{eMe}) \circ \hat{\gamma} \circ \hat{\beta}, \qquad \varphi|_{eMe} = (\varphi|_{eMe}) \circ \hat{\gamma} \circ \hat{\beta}, \tag{6.40}$$

where $\hat{\beta}$ is Petz' recovery map of $\hat{\gamma}$ with respect to $\varphi|_{eMe}$. Choose a normal state ρ on $(1 - e)M(1 - e)$ and define $\beta : M \to N$ by

$$\beta(x) := \hat{\beta}(exe) + \rho((1 - e)x(1 - e))(1 - e_0), \qquad x \in M.$$

Then it is clear that β is a unital normal 2-positive map. Furthermore, for every $x \in M$, we find that for every $x \in M$,

$$\begin{aligned} \psi_0 \circ \beta(x) &= (\psi_0|_{e_0 N e_0}) \circ \hat{\beta}(exe) + \rho((1 - e)x(1 - e))\psi_0(1 - e_0) \\ &= (\psi|_{eMe}) \circ \hat{\gamma} \circ \hat{\beta}(exe) \qquad \text{(by (6.37))} \\ &= (\psi|_{eMe})(exe) = \psi(x) \qquad \text{(by (6.40))}, \end{aligned}$$

and similarly $\varphi_0 \circ \beta(x) = \varphi(x)$. Thus, γ is reversible for $\{\psi, \varphi\}$, i.e., (i) holds. $\quad\square$

For every $\psi, \varphi \in M_*^+$, it is easy to see that γ is reversible for $\{\psi, \varphi\}$ if and only if γ is reversible for $\{\psi, \psi + \varphi\}$. Hence the next theorem follows from Theorem 6.19. This way of presentation of the reversibility theorem is given in [62, Theorem 3] for a family $\{\psi_\theta\}$ in M_*^+.

Theorem 6.20 *For every $\psi, \varphi \in M_*^+$ with $\psi \neq 0$ the following conditions are equivalent:*

(i) *γ is reversible for $\{\psi, \varphi\}$;*
(ii) *$S_f(\psi \| \psi + \varphi) = S_f(\psi \circ \gamma \| (\psi + \varphi) \circ \gamma)$ for every operator convex function on $(0, \infty)$;*
(iii) *$S_f(\psi \| \psi + \varphi) = S_f(\psi \circ \gamma \| (\psi + \varphi) \circ \gamma) < +\infty$ for some operator convex function on $(0, \infty)$ such that $\mathrm{supp}\, \mu_f$ has a limit point in $(0, \infty)$;*
(iv) *$D_\alpha(\psi \| \psi + \varphi) = D_\alpha(\psi \circ \gamma \| (\psi + \varphi) \circ \gamma)$ for some $\alpha \in (0, 1)$;*
(v) *$P(\psi, \psi + \varphi) = P(\psi \circ \gamma, (\psi + \varphi) \circ \gamma)$.*

Remark 6.21 When $M = B(\mathcal{H})$ with $\dim \mathcal{H} < \infty$, the assumption on $\mathrm{supp}\, \mu_f$ in condition (iii) of Theorems 6.19 and 6.20 can be relaxed to $|\mathrm{supp}\, \mu_f| \geq (\dim \mathcal{H})^2$, see [53, Theorem 5.1]. But the support condition on μ_f cannot completely be removed, so f cannot be a general nonlinear operator convex function in (iii) of the theorems. In fact, some counter-examples in the finite-dimensional case are known when $f(t) = t^2$ (see [64], [52, Example 4.8]) and when $f(t) = (1 + t)^{-1}$ (see [58, Example 1]), in which S_f is preserved under a CP map γ and yet the reversibility fails to hold.

Problem 6.22 It is interesting to find whether Theorem 6.19 holds true or not without the support assumption $s(\psi) \leq s(\varphi)$. In particular, assume that $s(\psi)$ and $s(\varphi)$ are orthogonal. In this case, condition (iv) (also (v)) implies that $s(\psi \circ \gamma)$ and $s(\varphi \circ \gamma)$ are also orthogonal. Then one can easily see that γ is reversible for $\{\psi, \varphi\}$. So it seems that the problem is not so hopeless.

6.5 Reversibility via Sandwiched Rényi Divergences

The reversibility via the sandwiched Rényi divergence (discussed in Sec. 3.3) in the von Neumann algebra setting has recently been obtained by Jenčová [59, 60]. Those reversibility results are reported below without proofs.

Theorem 6.23 ([59]) *Let $\gamma : N \to M$ be a unital normal 2-positive map between von Neumann algebras. Let $\psi, \varphi \in M_*^+$ with $\psi \neq 0$, and let $\alpha \in (1, \infty)$. If*

$$\widetilde{D}_\alpha(\psi \circ \gamma \| \varphi \circ \gamma) = \widetilde{D}_\alpha(\psi \| \varphi) < +\infty$$

(hence $s(\psi) \leq s(\varphi)$), then $\psi \circ \gamma \circ \gamma_\varphi^* = \psi$, where γ_φ^* is Petz' recovery map of γ with respect to φ (note that $\varphi \circ \gamma \circ \gamma_\varphi^* = \varphi$ automatically), and hence γ is reversible for $\{\psi, \varphi\}$.

Theorem 6.24 ([60]) *Let $\gamma : N \to M$ be a unital normal completely positive map between von Neumann algebras. Let $\psi, \varphi \in M_*^+$ with $\psi \neq 0$ and $s(\psi) \leq s(\varphi)$. Let $\alpha \in (1/2, 1)$. If*

$$\widetilde{D}_\alpha(\psi \circ \gamma \| \varphi \circ \gamma) = \widetilde{D}_\alpha(\psi \| \varphi),$$

then $\psi \circ \gamma \circ \gamma_\varphi^* = \psi$ and hence γ is reversible for $\{\psi, \varphi\}$.

Remark 6.25 Recall the two limit cases $\widetilde{D}_{1/2}(\psi \| \varphi) = -2 \log(F(\psi, \varphi)/\psi(1))$ in (3.24) and $\lim_{\alpha \to \infty} \widetilde{D}_\alpha(\psi \| \varphi) = D_{\max}(\psi \varphi)$ in Theorem 3.15 (2). In these limit cases, the reversibility as in the above theorems for a CP map γ fails to hold even in the finite-dimensional case, see [83, Corollary A.9] (also [52, Remark 5.15]).

Chapter 7
Reversibility and Measurements

7.1 Approximation of Connes' Cocycle Derivatives

Throughout this chapter, let f be an operator convex function on $(0, \infty)$, and $\psi, \varphi \in M_*^+$ be as before. The aim of this section is to prove the next lemma, which will be used in the next section in the situation where M_k's are commutative von Neumann algebras. The lemma is the approximation version of Lemma 6.13, that is, the equalities in (6.8) and (6.9) are replaced with the convergences in (7.1) and (7.2).

Lemma 7.1 *Let $\alpha_k : M_k \to M$ ($k \in \mathbb{N}$) be a sequence of unital normal Schwarz maps with von Neumann algebras M_k. Assume that* $\operatorname{supp} \mu_f$ *has a limit point in* $(0, \infty)$ *and that $\psi, \varphi \in M_*^+$ and α_k ($k \in \mathbb{N}$) are all faithful. If*

$$\lim_k S_f(\psi \circ \alpha_k \| \varphi \circ \alpha_k) = S_f(\psi \| \varphi) < +\infty, \tag{7.1}$$

then we have

$$\lim_k \alpha_k([D(\psi \circ \alpha_k) : D(\varphi \circ \alpha_k)]_t) = [D\psi : D\varphi]_t \ \ strongly, \quad t \in \mathbb{R}. \tag{7.2}$$

Proof We are working in the standard forms $(M, \mathcal{H}, J, \mathcal{P})$ and $(M_k, \mathcal{H}_k, J_k, \mathcal{P}_k)$. Write $\varphi_k := \varphi \circ \alpha_k$ and $\psi_k := \psi \circ \alpha_k$, and let Φ, Φ_k be the vector representatives of φ, φ_k, respectively. The proof below is similar to that of Lemma 6.13, so we similarly divide the proof into several steps.

Step 1. For each k a contraction $V_k : \mathcal{H}_k \to \mathcal{H}$ is defined by extending $y\Phi_k \mapsto \alpha_k(y)\Phi$ ($y \in M_k$). As in Step 1 of the proof of Lemma 6.13, replacing $\gamma, N, \psi_0, \varphi_0$ with $\alpha_k, M_k, \psi_k, \varphi_k$, we have

$$(s1 + \Delta_{\psi_k, \varphi_k})^{-1} \leq V_k^*(s1 + \Delta_{\psi, \varphi})^{-1}V_k, \qquad s > 0. \tag{7.3}$$

Step 2. Writing f as in (6.13) we have equalities (6.14), (6.15) and inequalities (6.16), where ψ_0, φ_0 are replaced with ψ_k, φ_k for any k. Therefore, assumption (7.1) implies that

$$\int_{[0,\infty)} \{S_{g_s}(\psi\|\varphi) - S_{g_s}(\psi_k\|\varphi_k)\}\, d\mu_f(s) \longrightarrow 0$$

(as well as $S_{f_2}(\psi_k\|\varphi_k) \to S_{f_2}(\psi\|\varphi)$ if $c > 0$), and so we can choose a subsequence $\{k(l)\}$ of $\{k\}$ such that

$$S_{g_s}(\psi_{k(l)}\|\varphi_{k(l)}) \longrightarrow S_{g_s}(\psi\|\varphi) \quad \text{for } \mu_f\text{-a.e. } s \in [0,\infty).$$

Thus, we may assume that

$$S_{g_s}(\psi_k\|\varphi_k) \longrightarrow S_{g_s}(\psi\|\varphi) \quad \text{for } \mu_f\text{-a.e. } s \in [0,\infty). \tag{7.4}$$

(Indeed, suppose that (7.2) does not hold; then there are a $t_0 \in \mathbb{R}$ and a neighborhood $\mathcal{V}$ of $[D\psi : D\varphi]_{t_0}$ in the strong topology, and a subsequence $\{m_k\}$ of $\{k\}$ such that $\alpha_k([D(\psi \circ \alpha_{m_k}) : D(\varphi \circ \alpha_{m_k})]_t) \notin \mathcal{V}$ for all k. From the above argument we can choose a subsequence $\{m_{k(l)}\}$ of $\{m_k\}$ such that

$$S_{g_s}(\psi_{m_{k(l)}}\|\varphi_{m_{k(l)}}) \longrightarrow S_{g_s}(\psi\|\varphi) \quad \text{for } \mu_f\text{-a.e. } s \in [0,\infty). \tag{7.5}$$

Then a contradiction occurs since the proof below under (7.5) shows (7.2) for the subsequence $\{m_{k(l)}\}$.)

With $h_s(t) := (t + s)^{-1}$, $s \in [0,\infty)$, we have (6.17) and (6.18), where ψ_0, φ_0 are replaced with ψ_k, φ_k for any k, so that we have by (7.4)

$$S_{h_s}(\psi_k\|\varphi_k) \longrightarrow S_{h_s}(\psi\|\varphi) \quad \text{for } \mu_f\text{-a.e. } s \in [0,\infty),$$

that is,

$$\langle \Phi_k, (s1 + \Delta_{\psi_k,\varphi_k})^{-1}\Phi_k \rangle \longrightarrow \langle \Phi, (s1 + \Delta_{\psi,\varphi})^{-1}\Phi \rangle \quad \text{for } \mu_f\text{-a.e. } s \in [0,\infty). \tag{7.6}$$

Note that $\langle \Phi, (z1 + \Delta_{\psi,\varphi})^{-1}\Phi \rangle$ and $\langle \Phi_k, (z1 + \Delta_{\psi_k,\varphi_k})^{-1}\Phi_k \rangle$ are analytic in $\{z \in \mathbb{C} : \text{Re } z > 0\}$, as seen just after (6.19). From the assumption on supp μ_f and (7.6) it follows that the set of $s > 0$ for which convergence in (7.6) holds has a limit point in $(0,\infty)$. Moreover, it is clear that the set of analytic functions $\langle \Phi_k, (z1+\Delta_{\psi_k,\varphi_k})^{-1}\Phi_k \rangle$ on $\{z \in \mathbb{C} : \text{Re } z > 0\}$ is locally bounded, i.e., uniformly bounded on each compact subset of $\{z \in \mathbb{C} : \text{Re } z > 0\}$. Thus, Vitali's theorem for analytic functions (see, e.g., [100, Sec. 7.3, p. 156]) yields that for all $s > 0$,

$$\langle \Phi_k, (s1 + \Delta_{\psi_k,\varphi_k})^{-1}\Phi_k \rangle \longrightarrow \langle \Phi, (s1 + \Delta_{\psi,\varphi})^{-1}\Phi \rangle$$

and hence

$$\langle \Phi_k, [V_k^*(s1 + \Delta_{\psi,\varphi})^{-1}V_k - (s1 + \Delta_{\psi_k,\varphi_k})^{-1}]\Phi_k \rangle \longrightarrow 0,$$

thanks to $V_k\Phi_k = \Phi$. By (7.3) this implies that

$$\|V_k^*(s1 + \Delta_{\psi,\varphi})^{-1}V_k\Phi_k - (s1 + \Delta_{\psi_k,\varphi_k})^{-1}\Phi_k\| \longrightarrow 0,$$

and therefore

$$\|V_k^*(s1 + \Delta_{\psi,\varphi})^{-1}\Phi - (s1 + \Delta_{\psi_k,\varphi_k})^{-1}\Phi_k\| \longrightarrow 0, \qquad s > 0. \qquad (7.7)$$

Step 3. As shown in Step 3 of the proof of Lemma 6.13, note that the linear span of h_s, $s \in [0, \infty)$, is dense in $C_0[0, \infty)$ with the sup-norm. Since

$$\phi \in C_0[0, \infty) \longmapsto V_k^*\phi(\Delta_{\psi,\varphi})\Phi, \ \phi(\Delta_{\psi_k,\varphi_k})\Phi_k \in \mathcal{H}_k \quad (k \in \mathbb{N})$$

are uniformly bounded linear maps with respect to the sup-norm on $C_0[0, \infty)$ and the norm on $\mathcal{H}_k$, it follows from (7.7) that

$$\|V_k^*\phi(\Delta_{\psi,\varphi})\Phi - \phi(\Delta_{\psi_k,\varphi_k})\Phi_k\| \longrightarrow 0, \qquad \phi \in C_0[0, \infty). \qquad (7.8)$$

Step 4. For every $s > 0$ one has

$$\lim_k \|V_k^*(s1 + \Delta_{\psi,\varphi})^{-1}\Phi\|^2$$

$$= \lim_k \langle (s1 + \Delta_{\psi_k,\varphi_k})^{-1}\Phi_k, (s1 + \Delta_{\psi_k,\varphi_k})^{-1}\Phi_k \rangle \quad \text{(by (7.7))}$$

$$= \lim_k \langle (s1 + \Delta_{\psi_k,\varphi_k})^{-2}\Phi_k, \Phi_k \rangle$$

$$= \langle (s1 + \Delta_{\psi,\varphi})^{-2}\Phi, \Phi \rangle \quad \text{(by (7.8) and } V_k\Phi_k = \Phi)$$

$$= \|(s1 + \Delta_{\psi,\varphi})^{-1}\Phi\|^2.$$

Letting $\zeta := (s1 + \Delta_{\psi,\varphi})^{-1}\Phi \in \mathcal{H}$, since the above means that $\lim_k \|V_k^*\zeta\|^2 = \|\zeta\|^2$, one has

$$\|V_kV_k^*\zeta - \zeta\|^2 = \|V_kV_k^*\zeta\|^2 - 2\|V_k^*\zeta\|^2 + \|\zeta\|^2 \le 2\|\zeta\|^2 - 2\|V_k^*\zeta\|^2 \longrightarrow 0$$

so that $\|V_kV_k^*\zeta - \zeta\| \to 0$, i.e.,

$$\|V_kV_k^*(s1 + \Delta_{\psi,\varphi})^{-1}\Phi - (s1 + \Delta_{\psi,\varphi})^{-1}\Phi\| \longrightarrow 0.$$

By (7.7) this implies that

$$\|V_k(s1 + \Delta_{\psi_k,\varphi_k})^{-1}\Phi_k - (s1 + \Delta_{\psi,\varphi})^{-1}\Phi\| \longrightarrow 0, \qquad s > 0.$$

Hence, similarly to Step 3, we have

$$\|V_k\phi(\Delta_{\psi_k,\varphi_k})\Phi_k - \phi(\Delta_{\psi,\varphi})\Phi\| \longrightarrow 0, \qquad \phi \in C_0[0, \infty). \qquad (7.9)$$

Step 5. Here we may assume that φ is a state, so $\|\Phi\| = 1$ and $\|\Phi_k\| = 1$ for all k. For each $t \in \mathbb{R}$ set $\phi(x)$, $x \ge 0$, by (6.23). For $n \in \mathbb{N}$ with $n \ge 2$ set κ_n by (6.24) and χ_n by

$$\chi_n(x) := \begin{cases} 0 & (0 \le x \le 1/n), \\ nx - 1 & (1/n \le x \le 2/n), \\ 1 & (2/n \le x \le n - \frac{1}{n}), \\ n(n - x) & (n - \frac{1}{n} \le x \le n), \\ 0 & (x \ge n). \end{cases}$$

Moreover, let $\phi_n(x) := \kappa_n(x)\phi(x)$ for $x \in [0, \infty)$. Then it is obvious that $\kappa_n, \chi_n, \phi_n \in C_0[0, +\infty)$ and $\chi_n \le 1_{[1/n,n]} \le \kappa_n$. Since $s(\psi) = s(\varphi) = 1$ and $s(\psi_k) = s(\varphi_k) = 1_{M_k}$, note that $\Delta_{\psi,\varphi}$ and $\Delta_{\psi_k,\varphi_k}$ are non-singular and so $\|\chi_n(\Delta_{\psi,\varphi})\Phi - \Phi\| \to 0$ as $n \to \infty$. For any $\varepsilon \in (0, 1)$ one can choose an n_0 such that

$$\|\chi_{n_0}(\Delta_{\psi,\varphi})\Phi\| \ge 1 - \varepsilon. \tag{7.10}$$

By (7.9) there exists a k_0 such that

$$\|V_k \chi_{n_0}(\Delta_{\psi_k,\varphi_k})\Phi_k - \chi_{n_0}(\Delta_{\psi,\varphi})\Phi\| \le \varepsilon, \qquad k \ge k_0. \tag{7.11}$$

It then follows from (7.10) and (7.11) that for every $n \ge n_0$ and $k \ge k_0$,

$$\begin{aligned} \|\chi_n(\Delta_{\psi_k,\varphi_k})\Phi_k\| &\ge \|\chi_{n_0}(\Delta_{\psi_k,\varphi_k})\Phi_k\| \ge \|V_k \chi_{n_0}(\Delta_{\psi_k,\varphi_k})\Phi_k\| \\ &\ge \|\chi_{n_0}(\Delta_{\psi,\varphi})\Phi\| - \|V_k \chi_{n_0}(\Delta_{\psi_k,\varphi_k})\Phi_k - \chi_{n_0}(\Delta_{\psi,\varphi})\Phi\| \\ &\ge 1 - 2\varepsilon. \end{aligned}$$

Therefore, for every $n \ge n_0$ and $k \ge k_0$,

$$\begin{aligned} \|(1 - \kappa_n)(\Delta_{\psi_k,\varphi_k})\Phi_k\|^2 &\le \|(1 - E_{\psi_k\varphi_k}([1/n, n])\Phi_k\|^2 \\ &= 1 - \|E_{\psi_k,\varphi_k}([1/n, n])\Phi_k\|^2 \\ &\le 1 - \|\chi_n(\Delta_{\psi_k,\varphi_k})\Phi_k\|^2 \\ &\le 1 - (1 - 2\varepsilon)^2 \le 4\varepsilon, \tag{7.12} \end{aligned}$$

where $\Delta_{\psi_k,\varphi_k} = \int_0^\infty t \, dE_{\psi_k,\varphi_k}(t)$ is the spectral decomposition of $\Delta_{\psi_k,\varphi_k}$. Furthermore, one can choose an $n_1 \ge n_0$ such that $\|(1 - \kappa_{n_1})(\Delta_{\psi,\varphi})\Phi\| \le \varepsilon$. From this and (7.12) it follows that for every $k \ge k_0$,

$$\begin{aligned} \|V_k \Delta_{\psi_k,\varphi_k}^{it}\Phi_k - \Delta_{\psi,\varphi}^{it}\Phi\| &\le \|V_k \phi(\Delta_{\psi_k,\varphi_k})\Phi_k - V_k \phi_{n_1}(\Delta_{\psi_k,\varphi_k})\Phi_k\| \\ &\quad + \|V_k \phi_{n_1}(\Delta_{\psi_k,\varphi_k})\Phi_k - \phi_{n_1}(\Delta_{\psi,\varphi})\Phi\| \\ &\quad + \|\phi_{n_1}(\Delta_{\psi,\varphi})\Phi - \phi(\Delta_{\psi,\varphi})\Phi\| \\ &\le \|\phi(\Delta_{\psi_k,\varphi_k})(1 - \kappa_{n_1})(\Delta_{\psi_k,\varphi_k})\Phi_k\| \\ &\quad + \|V_k \phi_{n_1}(\Delta_{\psi_k,\varphi_k})\Phi_k - \phi_{n_1}(\Delta_{\psi,\varphi})\Phi\| \\ &\quad + \|\phi(\Delta_{\psi,\varphi})(1 - \kappa_{n_1})(\Delta_{\psi,\varphi})\Phi\| \\ &\le 2\varepsilon^{1/2} + \|V_k \phi_{n_1}(\Delta_{\psi_k,\varphi_k})\Phi_k - \phi_{n_1}(\Delta_{\psi,\varphi})\Phi\| + \varepsilon. \end{aligned}$$

Therefore, by (7.9) there exists a $k_1 \ge k_0$ such that

$$\|V_k \Delta^{it}_{\psi_k,\varphi_k} \Phi_k - \Delta^{it}_{\psi,\varphi} \Phi\| \le 2(\varepsilon^{1/2} + \varepsilon), \qquad k \ge k_1,$$

which implies that

$$\|V_k \Delta^{it}_{\psi_k,\varphi_k} \Phi_k - \Delta^{it}_{\psi,\varphi} \Phi\| \longrightarrow 0 \quad \text{as } k \to \infty.$$

Since $\Delta^{-it}_{\varphi_k} \Phi_k = \Phi_k$ and $\Delta^{-it}_{\varphi} \Phi = \Phi$, the above implies by Proposition A.53 (and Remark A.54) that

$$\|V_k [D\psi_k : D\varphi_k]_t \Phi_k - [D\psi : D\varphi]_t \Phi\| \longrightarrow 0 \quad \text{as } k \to \infty.$$

Since $[D\psi_k : D\varphi_k]_t \in M_k$ so that $V_k [D\psi_k : D\varphi_k]_t \Phi_k = \alpha_k([D\psi_k : D\varphi_k]_t)\Phi$, it follows that $\|\alpha_k([D\psi_k : D\varphi_k]_t)\Phi - [D\psi : D\varphi]_t \Phi\| \to 0$, so (7.2) follows since $\overline{M'\Phi} = \mathcal{H}$. $\qquad\qquad\square$

7.2 Reversibility via Measurements

In this section we will study reversibility via measurement procedure, i.e., quantum-classical channels described by unital normal positive maps $\alpha : \mathcal{A} \to M$ with commutative $\mathcal{A}$) and its relation with equalities between S_f, $\widehat{S}_f$ and S_f^{meas}, the standard, the maximal and the measured f-divergences discussed in Chaps. 2, 4 and 5. Recall that $S_f^{\text{meas}} \le S_f \le \widehat{S}_f$, see (5.7).

We consider the following conditions:

(i) ψ, φ commute (see Lemma 4.19 and Definition A.56).

(ii) *Existence of a sufficient commutative subalgebra:* There exist a commutative von Neumann subalgebra $\mathcal{A}$ of M and a normal conditional expectation $E : M \to \mathcal{A}$ such that $\psi \circ E = \psi$ and $\varphi \circ E = \varphi$.

(iii) *Reversibility via a measurement channel:* There exist unital normal positive maps $\alpha : \mathcal{A} \to M$ and $\beta : M \to \mathcal{A}$ with a commutative von Neumann algebra $\mathcal{A}$ such that

$$\psi = \psi \circ \alpha \circ \beta, \qquad \varphi = \varphi \circ \alpha \circ \beta.$$

(iv) *Approximate reversibility via measurement channels:* There exist unital normal positive maps $\alpha_k : \mathcal{A}_k \to M$ and $\beta_k : M \to \mathcal{A}_k$ ($k \in \mathbb{N}$) with commutative von Neumann algebras $\mathcal{A}_k$ such that

$$\psi \circ \alpha_k \circ \beta_k \to \psi, \qquad \varphi \circ \alpha_k \circ \beta_k \to \varphi \quad \text{in } \sigma(M_*, M).$$

(v) There exists a unital normal positive map $\alpha : \mathcal{A} \to M$ with a commutative von Neumann algebra $\mathcal{A}$ such that

$$S_f(\psi \circ \alpha \| \varphi \circ \alpha) = S_f(\psi \| \varphi).$$

(vi) $S_f^{\text{pr}}(\psi \| \varphi) = S_f(\psi \| \varphi)$ (see Definition 5.1).

(vii) $S_f^{\mathrm{meas}}(\psi\|\varphi) = S_f(\psi\|\varphi)$, that is, there exists a sequence of unital normal positive maps $\alpha_k : \mathcal{A}_k \to M$ with commutative von Neumann algebras $\mathcal{A}_k$ such that

$$\lim_k S_f(\psi \circ \alpha_k \| \varphi \circ \alpha_k) = S_f(\psi\|\varphi).$$

(viii) $S_f^{\mathrm{meas}}(\psi\|\varphi) = \widehat{S}_f(\psi\|\varphi)$.

(ix) $S_f(\psi\|\varphi) = \widehat{S}_f(\psi\|\varphi)$.

We first summarize in the next lemma an easy part of the proof of Theorem 7.3 below.

Lemma 7.2 *For the above conditions we have*

$$\text{(i)} \implies \text{(ii)} \implies \text{(iii)} \implies \text{(iv)} \implies \text{(vii)},$$
$$\text{(iii)} \implies \text{(v)} \implies \text{(vii)},$$
$$\text{(ii)} \implies \text{(vi)} \implies \text{(vii)},$$
$$\text{(iii)} \implies \text{(viii)} \implies \text{(ix)}.$$

Proof (i) $\implies$ (ii). Assume (i) and let $e := s(\psi + \varphi)$, $\psi_0 := \psi|_{eMe}$ and $\varphi_0 := \varphi|_{eMe}$. From the proof of [49, Proposition 6.7], there exists a commutative von Neumann subalgebra $\mathcal{A}_0$ of eMe and a normal conditional expectation $E_0 : eMe \to \mathcal{A}_0$ such that $\psi\_0 = \psi_0 \circ E_0$ and $\varphi_0 = \varphi_0 \circ E_0$. Hence (ii) holds with $\mathcal{A} = \mathcal{A}_0 + \mathbb{C}(1 - e)$ and $E : M \to \mathcal{A}$ given by

$$E(x) := E_0(exe) + \rho((1 - e)x(1 - e))(1 - e), \qquad x \in M,$$

where ρ is a normal state on $(1 - e)M(1 - e)$.

(ii) $\implies$ (iii) is clear with the injection $\alpha : \mathcal{A} \to M$ and $\beta = E$, and (iii) $\implies$ (iv) is obvious. (iv) $\implies$ (vii) follows since

$$\begin{aligned} S_f(\psi\|\varphi) &\le \liminf_k S_f(\psi \circ \alpha_k \circ \beta_k \| \varphi \circ \alpha_k \circ \beta_k) \\ &\le \liminf_k S_f(\psi \circ \alpha_k \| \varphi \circ \alpha_k) \\ &\le S_f^{\mathrm{meas}}(\psi\|\varphi) \le S_f(\psi\|\varphi) \end{aligned}$$

thanks to Theorem 2.7 (i), (iv) and (5.7).

(iii) $\implies$ (v) and (v) $\implies$ (vii) follow from Theorem 2.7 (iv) and Proposition 5.2. (ii) $\implies$ (vi) is seen as

$$\begin{aligned} S_f(\psi\|\varphi) &= S_f(\psi|_{\mathcal{A}} \| \varphi|_{\mathcal{A}}) \\ &= \sup\{S_f(\psi|_{\mathcal{A}_0} \| \varphi|_{\mathcal{A}_0}) : \mathcal{A}_0 \text{ a finite-dimensional subalgebra of } \mathcal{A}\} \\ &\le S_f^{\mathrm{pr}}(\psi\|\varphi), \end{aligned}$$

where the second equality is due to Theorem 2.7 (vii). (vi) $\implies$ (vii) is obvious from (5.7). (iii) $\implies$ (viii) is seen since (iii) implies that

$$\widehat{S}_f(\psi\|\varphi) = \widehat{S}_f(\psi\circ\alpha\circ\beta\|\varphi\circ\alpha\circ\beta) \leq S_f(\psi\circ\alpha\|\varphi\circ\alpha) \leq S_f^{\mathrm{meas}}(\psi\|\varphi),$$

where the first inequality above is due to Theorem 4.4 (i) and Proposition 4.20. Finally, (viii) $\Longrightarrow$ (ix) is obvious. $\square$

The main aim of the present section is to prove the following:

Theorem 7.3 *Assume that* supp μ_f, *the support of the representing measure of* f *(see* (2.5)*), has a limit point in* $(0,\infty)$. *If* $s(\psi) = s(\varphi)$ *and* $S_f(\psi\|\varphi) < +\infty$, *then the above conditions* (i)–(ix) *are all equivalent.*

Once we have shown the implications in Lemma 7.2, it only remains to prove (vii) $\Longrightarrow$ (i) and (ix) $\Longrightarrow$ (v). We prepare one more lemma.

Lemma 7.4 *Let* $\alpha_k : A_k \to B(\mathcal{H})$ *be a sequence of unital Schwarz maps where* A_k*'s are unital* C^**-algebras. Let* $u_k, a_k \in A_k$ *be given such that* u_k*'s are unitaries and* $\sup_k \|a_k\| < +\infty$. *If* $\alpha_k(u_k) \to u$ *and* $\alpha_k(a_k) \to a$ *in the strong* topology for some* $u, a \in B(\mathcal{H})$ *where* u *is a unitary. Then* $\alpha_k(a_k u_k) \to au$ *in the weak operator topology.*

Remark 7.5 The lemma is in [95, Lemma 3.2] and [88, Lemma 9.8] without proof, where the convergence $\alpha_k(a_k u_k) \to au$ in the strong operator topology is claimed. In fact, we don't see this stronger convergence while the weaker convergence as above will be enough for the proof of Theorem 7.3.

Proof (Lemma 7.4) For every $\lambda \in \mathbb{R}$ we have

$$\lambda\{\alpha_k(a_k)\alpha(u_k) + \alpha_k(u_k)^*\alpha_k(a_k)^*\}$$
$$= \alpha_k(a_k + \lambda u_k^*)\alpha_k(a_k^* + \lambda u_k) - \alpha_k(a_k)\alpha_k(a_k)^* - \lambda^2\alpha_k(u_k)^*\alpha_k(u_k)$$
$$\leq \alpha_k((a_k + \lambda u_k^*)(a_k^* + \lambda u_k)) - \alpha_k(a_k)\alpha_k(a_k)^* - \lambda^2\alpha_k(u_k)^*\alpha(u_k)$$
$$= \alpha_k(a_k a_k^*) - \alpha_k(a_k)\alpha_k(a_k)^* + \lambda\{\alpha_k(a_k u_k) + \alpha_k(u_k^* a_k^*)\}$$
$$+ \lambda^2\{\alpha_k(u_k^* u_k) - \alpha_k(u_k)^*\alpha(u_k)\}. \tag{7.13}$$

For any vector $\xi \in \mathcal{H}$ the above implies that for every $\lambda > 0$,

$$\langle\xi, \{\alpha_k(a_k)\alpha(u_k) + \alpha_k(u_k)^*\alpha_k(a_k)^*\}\xi\rangle$$
$$\leq \frac{1}{\lambda}\langle\xi, \{\alpha_k(a_k a_k^*) - \alpha_k(a_k)\alpha_k(a_k)^*\}\xi\rangle + \langle\xi, \{\alpha_k(a_k u_k) + \alpha_k(u_k^* a_k^*)\}\xi\rangle$$
$$+ \lambda\langle\xi, \{1 - \alpha_k(u_k)^*\alpha_k(u_k)\}\xi\rangle. \tag{7.14}$$

Since $\alpha_k(u_k)^*\alpha_k(u_k) \to u^*u = 1$ strongly, one can choose $\lambda_k > 0$, $k \in \mathbb{N}$, such that

$$\lambda_k \to +\infty, \qquad \lambda_k\langle\xi, \{1 - \alpha_k(u_k)^*\alpha(u_k)\}\xi\rangle \longrightarrow 0.$$

Then using (7.14) to $\lambda = \lambda_k$ and letting $k \to \infty$ one has

$$\langle\xi, (au + u^*a^*)\xi\rangle \leq \liminf_k\langle\xi, \{\alpha_k(a_k u_k) + \alpha_k(u_k^* a_k^*)\}\xi\rangle. \tag{7.15}$$

Similarly, for every $\lambda < 0$,

$$\langle \xi, \{\alpha_k(a_k)\alpha(u_k) + \alpha_k(u_k^*)\alpha_k(a_k^*)\}\xi\rangle$$
$$\geq \frac{1}{\lambda}\langle \xi, \{\alpha_k(a_k a_k^*) - \alpha_k(a_k)\alpha_k(a_k)^*\}\xi\rangle + \langle \xi, \{\alpha_k(a_k u_k) + \alpha_k(u_k^* a_k^*)\}\xi\rangle$$
$$+ \lambda\langle \xi, \{1 - \alpha_k(u_k)^*\alpha(u_k)\}\xi\rangle.$$

Taking $\lambda_k < 0$, $k \in \mathbb{N}$, with $\lambda_k \to -\infty$ and $\lambda_k\langle \xi, \{1 - \alpha_k(u_k)^*\alpha(u_k)\}\xi\rangle \to 0$, one has

$$\langle \xi, (au + u^*a^*)\xi\rangle \geq \limsup_k \langle \xi, \{\alpha_k(a_k u_k) + \alpha_k(u_k^* a_k^*)\}\xi\rangle. \qquad (7.16)$$

From (7.15) and (7.16) it follows that

$$\alpha_k(a_k u_k) + \alpha_k(u_k^* a_k^*) \longrightarrow au + u^*a^* \quad \text{weakly.} \qquad (7.17)$$

Next, replace u_k with iu_k in (7.13) to obtain

$$\lambda i\{\alpha_k(a_k)\alpha(u_k) - \alpha_k(u_k)^*\alpha_k(a_k)^*\}$$
$$\leq \alpha_k(a_k a_k^*) - \alpha_k(a_k)\alpha_k(a_k)^* + \lambda i\{\alpha_k(a_k u_k) - \alpha_k(u_k^* a_k^*)\}$$
$$+ \lambda^2\{\alpha_k(u_k^* u_k) - \alpha_k(u_k)^*\alpha(u_k)\}.$$

From this, the same argument as above gives

$$\alpha_k(a_k u_k) - \alpha_k(u_k^* a_k^*) \longrightarrow au - u^*a^* \quad \text{weakly.} \qquad (7.18)$$

Combining (7.17) and (7.18) yields the asserted conclusion.   □

Proof (Theorem 7.3) In view of Lemma 7.2 the proof will be completed when we show the following two implications.

(ix) $\Longrightarrow$ (v). By Theorem 4.17 there exist a unital normal positive map $\beta : M \to \mathcal{A}$ with a commutative von Neumann algebra $\mathcal{A}$ and $p, q \in \mathcal{A}_*^+$ such that $\psi = p \circ \beta$, $\varphi = q \circ \beta$ and $S_f(p\|q) = \widehat{S}_f(\psi\|\varphi)$. From (ix) this implies that $S_f(p\|q) = S_f(\psi\|\varphi) < +\infty$. Hence, from Theorem 6.19 it follows that β is reversible for $\{p, q\}$, so there exists a unital normal positive map $\alpha : \mathcal{A} \to M$ such that $p = \psi \circ \alpha$ and $q = \varphi \circ \alpha$. Since

$$S_f(\psi \circ \alpha\|\varphi \circ \alpha) = S_f(p\|q) = S_f(\psi\|\varphi),$$

condition (v) holds. (Note that the proof of this part works when $s(\psi) \leq s(\varphi)$.)

(vii) $\Longrightarrow$ (i). Condition (vii) means that that there exists a sequence of unital normal positive maps $\alpha_k : \mathcal{A}_k \to M$ with commutative $\mathcal{A}_k$ such that (7.1) holds. Put $e := s(\psi) = s(\varphi)$. For each k, by the normality of α_k there is a projection $e_k \in \mathcal{A}_k$ such that $\alpha_k(1_{\mathcal{A}_k} - e_k) = 0$ and α_k is faithful on $e_k\mathcal{A}_k e_k$. Define $\hat{\alpha}_k := e\alpha(\cdot)e|_{e_k\mathcal{A}_k e_k} : e_k\mathcal{A}_k e_k \to eMe$, which is obviously unital (i.e., $\hat{\alpha}_k(e_k) = e$), normal and faithful. Note that

$$S_f(\psi\|\varphi) = S_f(e\psi e\|e\varphi e).$$

Moreover,

$$(\psi \circ \alpha_k)(a) = (e\psi e) \circ \hat{\alpha}_k(e_k a e_k), \qquad a \in \mathcal{A}_k,$$

and similarly for φ, so that

$$S_f(\psi \circ \alpha_k \| \varphi \circ \alpha_k) = S_f((e\psi e) \circ \hat{\alpha}_k \| (e\varphi e) \circ \hat{\alpha}_k).$$

Hence, replacing ψ, φ, α_k by $e\psi e, e\varphi e, \hat{\alpha}_k$, we may assume that $s(\psi) = s(\varphi) = 1$ and all α_k are faithful. Then by Lemma 7.1 we have

$$\alpha_k([D\psi \circ \alpha_k : D\varphi \circ \alpha_k]_t) \longrightarrow [D\psi : D\varphi]_t \ \text{ strongly}, \quad t \in \mathbb{R}.$$

By changing the roles of ψ, φ with replacing f by its transpose $\widetilde{f}$ (see Proposition 2.3 (5)) and noting that supp $\mu_{\widetilde{f}}$ has a limit point in $(0, \infty)$ again, we also have

$$\alpha_k([D\varphi \circ \alpha_k : D\psi \circ \alpha_k]_t) \longrightarrow [D\varphi : D\psi]_t \ \text{ strongly}, \quad t \in \mathbb{R}.$$

Therefore, letting $u_{k,t} := [D\psi \circ \alpha_k : D\varphi \circ \alpha_k]_t$ and $u_t := [D\psi : D\varphi]_t$, by Proposition A.51 (1) we have $\alpha_k(u_{k,t}) \to u_t$ strongly* for all $t \in \mathbb{R}$. By Lemma 7.4 we have

$$\alpha_k(u_{k,s} u_{k,t}) \longrightarrow u_s u_t \ \text{ weakly}, \quad s, t \in \mathbb{R}.$$

Since $\mathcal{A}_k$ is commutative and so $\{u_{k,t}\}_{t\in\mathbb{R}}$ is a one-parameter unitary group by Proposition A.55, we have $u_s u_t = u_t u_s$ for all $s, t \in \mathbb{R}$, which means by Proposition A.55 again that ψ and φ are commuting. □

Theorem 7.6 *Assume that* supp μ_f *has a limit point in* $(0, \infty)$. *Let* $\psi, \varphi \in M_*^+$, *and assume that* $s(\psi) \leq s(\varphi)$ *and* $S_f(\psi\|\varphi) < +\infty$. *Let* (i)′–(ix)′ *denote the conditions corresponding to* (i)–(ix), *respectively, where* ψ *is replaced with* $\psi + \varphi$; *for example,* (vii)′ $S_f^{\mathrm{meas}}(\psi + \varphi\|\varphi) = S_f(\psi + \varphi\|\varphi)$. *Then the conditions* (i)–(v), (viii), (ix) *and all* (i)′–(ix)′ *are equivalent.*

Proof Each condition of (i)–(iv) is unchanged when ψ is replaced with $\psi + \varphi$. Note that $s(\psi + \varphi) = s(\varphi)$ from $s(\psi) \leq s(\varphi)$ and $S_f(\psi + \varphi\|\varphi) < +\infty$ by Proposition 2.9. Hence we can apply Theorem 7.3 to $\psi + \varphi$ and φ to see that the conditions (i)–(iv) and (v)′–(ix)′ are all equivalent. Furthermore, we have (iii) $\Longrightarrow$ (v) and (iii) $\Longrightarrow$ (viii) $\Longrightarrow$ (ix) by Lemma 7.2, (ix) $\Longrightarrow$ (v) by the above proof of Theorem 7.3 (this part has been done with $s(\psi) \leq s(\varphi)$), and (v) $\Longrightarrow$ (iii) from Theorem 6.19. Hence (v), (viii) and (ix) are also equivalent to (iii), so the result follows. □

Problem 7.7 Conditions (vi) and (vii) are missing in Theorem 7.6. But it is unknown to us how to prove the implication (vii) $\Longrightarrow$ (i) under $s(\psi) \leq s(\varphi)$. Here, note that although $[D\psi : D\varphi]_t$ is defined even when $s(\psi) \leq s(\varphi) = 1$ (in this case, $[D\psi : D\varphi]_0 = s(\psi)$), Lemma 7.1 does not hold in general. Indeed, one can easily construct an example of ψ, φ and von Neumann subalgebras $M_k \subset M$ such that $\psi|_{M_k}$ and $\varphi|_{M_k}$ are faithful and $S_f(\psi \circ \alpha_k \| \varphi \circ \alpha_k) \to S_f(\alpha\|\varphi) < +\infty$, for example,

in the case where M is commutative and $f(x) = x^\alpha$, $1 < \alpha < 2$. But in this case, $\alpha_k([D\psi_k : D\varphi_k]_0) = \alpha_k(1) = 1 \neq s(\psi) = [D\psi : D\varphi]_0$, so the strong convergence (7.2) is impossible. This phenomenon is similar to the convergence of positive self-adjoint operators. Let A_k, A be positive self-adjoint operators on a Hilbert space such that $A_k \to A$ in the sense that $(\lambda I + A_k)^{-1} \to (\lambda I + A)^{-1}$ strongly for all $\lambda > 0$ (see Proposition B.1). When all A_k are non-singular but A is singular, note that A_k^{it} cannot converge to $s(A)A^{it}$ for $t \in \mathbb{R}$; in fact, $A_k^0 = I \not\to s(A)$ for $t = 0$. Therefore, we need to find another route to prove (vii) $\Longrightarrow$ (i) for the case $s(\psi) \leq s(\varphi)$.

Problem 7.8 A significant problem is the attainability question for S_f^{meas}, that is, whether there exists a unital normal positive map $\alpha : \mathcal{A} \to M$ with commutative $\mathcal{A}$ such that $S_f(\psi \circ \alpha \| \varphi \circ \alpha) = S_f^{\mathrm{meas}}(\psi \| \varphi)$. Theorem 7.3 tells that the attainability holds for f satisfying the assumption in the theorem if $s(\psi) = s(\varphi)$ and $S_f^{\mathrm{meas}}(\psi \| \varphi) = S_f(\psi \| \varphi) < +\infty$. When $M = B(\mathcal{H})$ with $\dim \mathcal{H} < \infty$, this is true for arbitrary ψ, φ [52, Proposition 4.17]. Suppose the problem holds true in general for $\psi, \varphi \in M_*^+$ with $s(\psi) \leq s(\varphi)$, then conditions (v) and (vii) are equivalent and so Problem 7.7 becomes meaningless.

Chapter 8
Preservation of Maximal f-Divergences

8.1 Maximal f-Divergences and Operator Connections

In this chapter we will characterize the preservation of $\widehat{S}_f$ under a unital normal positive map γ, i.e., the equality case in the monotonicity inequality $\widehat{S}_f(\psi \circ \gamma \| \varphi \circ \varphi) \leq \widehat{S}_f(\psi \| \varphi)$. To do this, we work in the standard form

$$(M, L^2(M), J = {}^*, L^2(M)_+),$$

where $L^p(M)$ is Haagerup's L^p-space (see Appendix A.6). Recall that M_* is identified with $L^1(M)_+$ by the linear bijection $\varphi \leftrightarrow h_\varphi$ and the tr -functional on $L^1(M)$ is given by $\mathrm{tr}(h_\varphi) = \varphi(1)$. To present the main result, we introduce the notion of operator connections of $\psi, \varphi \in M_*^+$, which is a kind of extension of Kubo-Ando's operator connections [73]. This is an interesting topic in its own, so we develop the theory in some detail in Appendix D separately.

Definition 8.1 Let $\psi, \varphi \in M_*^+$ and assume that $(\psi, \varphi) \in (M_*^+ \times M_*^+)_{\leq}$, see Sec. 4.1. By Lemma A.21 we have a unique $A \in s(\varphi)Ms(\varphi)$ such that $h_\psi^{1/2} = Ah_\varphi^{1/2}$ (note that $h_\varphi^{1/2} \in L^2(M)_+$ is the vector representative of φ). Let $T_{\psi/\varphi} := A^*A$. For every bounded Borel function ϕ on $[0, \infty)$, since $h_\varphi^{1/2}\phi(T_{\psi/\varphi})h_\varphi^{1/2} \in L^1(M)_+$, we define the M_*-valued *operator perspective* $P_\phi(\psi, \varphi)$ by

$$h_{P_\phi(\psi,\varphi)} = h_\varphi^{1/2}\phi(T_{\psi/\varphi})h_\varphi^{1/2}. \tag{8.1}$$

Definition 8.2 Let σ be an operator connection in Kubo-Ando's sense [73] with the representing function k_σ, a non-negative operator monotone function on $[0, \infty)$. We define the M_*-valued *operator connection* $\varphi\sigma\psi$ as

$$\varphi\sigma\psi = P_{k_\sigma}(\psi, \varphi), \quad \text{i.e.,} \quad h_{\varphi\sigma\psi} = h_\varphi^{1/2}k_\sigma(T_{\psi/\varphi})h_\varphi^{1/2}. \tag{8.2}$$

For general $\psi, \varphi \in M_*^+$, choosing an $\omega \in M_*^+$ with $\omega \sim \varphi + \psi$, we define

$$\varphi\sigma\psi := \lim_{\varepsilon\searrow 0}(\varphi+\varepsilon\omega)\sigma(\psi+\varepsilon\omega) \in M_*^+ \quad \text{in the norm.} \tag{8.3}$$

From Lemma D.4 and Theorem D.6 of Appendix D, the limit in (8.3) as $\varepsilon\searrow 0$ exists independently of choice of $\omega \sim \varphi+\psi$, and it coincides with the definition (8.2) when $(\psi,\varphi) \in (M_*^+ \times M_*^+)_{\leq}$. In fact, a more intrinsic definition of $\varphi\sigma\psi$ is provided in Appendix D.

Now let $\gamma : N \to M$ be a unital normal positive linear map between von Neumann algebras, and let $\gamma_* : L^1(M) \to L^1(N)$ be the predual map of γ via $M_* \cong L^1(M)$ and $N_* \cong L^1(N)$, i.e., $\gamma_*(h_\varphi) = h_{\varphi\circ\gamma}$ for $\varphi \in M_*$ and the corresponding $h_\varphi \in L^1(M)$. For every $\varphi \in M_*^+$ let $e_0 \in N$ be the support projection of $\varphi\circ\gamma$, and $\gamma_\varphi^* : M \to e_0 N e_0$ be the (extended) Petz' recovery map of γ with respect to φ, given in Proposition 6.6. The next lemma is the description of γ_φ^* in terms of Haagerup's L^1-elements h_φ and $h_{\varphi\circ\gamma}$. Note that (8.4) has a strong resemblance to (6.2) in the finite-dimensional case.

Lemma 8.3 *For every $\varphi \in M_*^+$ let γ_φ^* be as stated above. Then for every $x \in M$, $\gamma_\varphi^*(x) \in e_0 N e_0$ is determined by*

$$h_{\varphi\circ\gamma}^{1/2}\gamma_\varphi^*(x)h_{\varphi\circ\gamma}^{1/2} = \gamma_*(h_\varphi^{1/2}xh_\varphi^{1/2}). \tag{8.4}$$

Proof Write $\beta = \gamma_\varphi^*$. For every $x \in M$ the defining condition of $\beta(x)$ is equation (6.1) for all $y \in e_0 N e_0$, and it can be rewritten in terms of of $h_\varphi, h_{\varphi\circ\gamma}$ as

$$\operatorname{tr}\beta(x)h_{\varphi\circ\gamma}^{1/2}yh_{\varphi\circ\gamma}^{1/2} = \operatorname{tr} xh_\varphi^{1/2}\gamma(y)h_\varphi^{1/2}, \quad y \in e_0 N e_0,$$

that is,

$$\operatorname{tr} h_{\varphi\circ\gamma}^{1/2}\beta(x)h_{\varphi\circ\gamma}^{1/2}y = \operatorname{tr}\gamma_*(h_\varphi^{1/2}xh_\varphi^{1/2})y, \qquad y \in e_0 N e_0.$$

Since $e_0 h_{\varphi\circ\gamma}^{1/2}e_0 = h_{\varphi\circ\gamma}^{1/2}$ and $e_0\gamma_*(h_\varphi^{1/2}xh_\varphi^{1/2})e_0 = \gamma_*(h_\varphi^{1/2}xh_\varphi^{1/2})$ as easily verified, the above holds for all $y \in N$, which means that $h_{\varphi\circ\gamma}^{1/2}\beta(x)h_{\varphi\circ\gamma}^{1/2} = \gamma_*(h_\varphi^{1/2}xh_\varphi^{1/2})$. $\square$

In this section we present the next theorem on the preservation of the maximal f-divergence, i.e., the equality case in the monotonicity inequality of $\widehat{S}_f(\psi\|\varphi)$, which is the extension of [52, Theorem 3.34] to the von Neumann algebra setting. It is remarkable that we treat general $\psi,\varphi \in M_*^+$ with no condition on their support projections, differently from the reversibility theorems in Chaps. 6 and 7, so that even in the finite-dimensional case, the next theorem improves [52, Theorem 3.34].

Theorem 8.4 *Let γ be as above. Let $\psi,\varphi \in M_*^+$ and set $\omega := \psi + \varphi$. As in Definition 8.1 let $T_{\psi/\omega} = A^*A$ with $A \in s(\omega)Ms(\omega)$ satisfying $h_\psi^{1/2} = Ah_\omega^{1/2}$, and $T_{\psi\circ\gamma/\omega\circ\gamma} = A_0^*A_0$ with $A_0 \in s(\omega\circ\gamma)Ns(\omega\circ\gamma)$ satisfying $h_{\psi\circ\gamma}^{1/2} = A_0 h_{\omega\circ\gamma}^{1/2}$. Then the following conditions are equivalent:*

(i) *$\widehat{S}_f(\psi\circ\gamma\|\varphi\circ\gamma) = \widehat{S}_f(\psi\|\varphi)$ for any operator convex function f on $(0,\infty)$;*

(ii) *$\widehat{S}_f(\psi\circ\gamma\|\varphi\circ\gamma) = \widehat{S}_f(\psi\|\varphi) < +\infty$ for some nonlinear operator convex function f on $(0,\infty)$ with $\lim_{t\to 0^+}tf(t) = 0$ (in particular, $f(0^+) < +\infty$) and $\lim_{t\to\infty}f(t)/t^2 = 0$ (in particular, $f'(\infty) < +\infty$);*

(iii) $\widehat{S}_f(\psi \circ \gamma \| \omega \circ \gamma) = \widehat{S}_f(\psi \| \omega)$ *for any operator convex function* f *on* $[0, \infty)$*;*

(iv) $\widehat{S}_f(\psi \circ \gamma \| \omega \circ \gamma) = \widehat{S}_f(\psi \| \omega)$ *for some nonlinear operator convex function* f *on* $[0, \infty)$*;*

(v) $S_{t^2}(\psi \circ \gamma \| \omega \circ \gamma) = S_{t^2}(\psi \| \omega)$ *(equivalently* $D_2(\psi \circ \gamma \| \omega \circ \gamma) = D_2(\psi \| \omega)$ *as far as* $\psi \ne 0$*);*

(vi) $(\varphi \circ \gamma)\sigma(\psi \circ \gamma) = (\varphi \sigma \psi) \circ \gamma$ *for any operator connection* σ*;*

(vii) $(\varphi \circ \gamma)\sigma(\psi \circ \gamma) = (\varphi \sigma \psi) \circ \gamma$ *for some nonlinear operator connection* σ*;*

(viii) $(\omega \circ \gamma)\sigma(\psi \circ \gamma) = (\omega \sigma \psi) \circ \gamma$ *for any operator connection* σ*;*

(ix) $(\omega \circ \gamma)\sigma(\psi \circ \gamma) = (\omega \sigma \psi) \circ \gamma$ *for some nonlinear operator connection* σ*;*

(x) $P_\phi(\psi \circ \gamma, \omega \circ \gamma) = P_\phi(\psi, \omega) \circ \gamma$ *for any bounded Borel function* ϕ *on* $[0, \infty)$*;*

(xi) $h_{\psi \circ \gamma}^{1/2} A_0 A_0^* h_{\psi \circ \gamma}^{1/2} = \gamma_*(h_\psi^{1/2} A A^* h_\psi^{1/2})$*;*

(xii) $\gamma_\omega^*((T_{\psi/\omega})^2) = (\gamma_\omega^*(T_{\psi/\omega}))^2$*.*

Moreover, if γ *is 2-positive and hence so is* γ_ω^**, then the above conditions are also equivalent to*

(xiii) $T_{\psi/\omega} \in \mathcal{M}_{\gamma_\omega^*}$ *(the multiplicative domain of* γ_ω^**).*

For example, the function $f(t) = t \log t$ on $(0, \infty)$ satisfies $f(0^+) < +\infty$ and $\lim_{t \to \infty} f(t)/t^2 = 0$. so that a typical realization of condition (ii) is that $D_{\mathrm{BS}}(\psi \circ \gamma \| \varphi \circ \gamma) = D_{\mathrm{BS}}(\psi \| \varphi) < +\infty$, see Example 4.21.

In the rest let $\psi_0 := \psi \circ \gamma$, $\varphi_0 := \varphi \circ \gamma$, $\omega_0 := \omega \circ \gamma = \psi_0 + \varphi_0$, $e := s(\omega) \in M$ and $e_0 := s(\omega_0) \in N$. Before proving the theorem we give two simple lemmas.

Lemma 8.5 *We have* $\gamma_\omega^*(T_{\psi/\omega}) = T_{\psi_0/\omega_0}$*.*

Proof By (8.4) we have

$$h_{\omega_0}^{1/2} \gamma_\omega^*(T_{\psi/\omega}) h_{\omega_0}^{1/2} = \gamma_*(h_\omega^{1/2} T_{\psi/\omega} h_\omega^{1/2}) = \gamma_*(h_\psi^{1/2} h_\psi^{1/2})$$
$$= \gamma_*(h_\psi) = h_{\psi_0} = h_{\omega_0}^{1/2} T_{\psi_0/\omega_0} h_{\omega_0}^{1/2}.$$

Hence, for every $y, z \in N$,

$$\langle h_{\omega_0}^{1/2} y, (\gamma_\omega^*(T_{\psi/\omega}) - T_{\psi_0/\omega_0}) h_{\omega_0}^{1/2} z \rangle = 0.$$

Since $\overline{h_{\omega_0}^{1/2} N} = e_0 L^2(\mathcal{H}_0)$ and $\gamma_\omega^*(T_{\psi/\omega}), T_{\psi_0/\omega_0} \in e_0 N e_0$, the assertion follows. $\square$

Lemma 8.6 *For every operator convex function* f *on* $[0, \infty)$*,*

$$P_f(\psi \circ \gamma, \omega \circ \gamma) \le P_f(\psi, \omega) \circ \gamma.$$

Proof By the definition in (8.1), Lemma 8.5, and by (8.4) applied to ω, we have

$$h_{P_f(\psi_0, \omega_0)} = h_{\varphi_0}^{1/2} f(T_{\psi_0/\omega_0}) h_{\omega_0}^{1/2} = h_{\omega_0}^{1/2} f(\gamma_\omega^*(T_{\psi/\omega})) h_{\omega_0}^{1/2}$$
$$\le h_{\omega_0}^{1/2} \gamma_\omega^*(f(T_{\psi/\omega})) h_{\omega_0}^{1/2} = \gamma_*(h_\omega^{1/2} f(T_{\psi/\omega}) h_\omega^{1/2})$$
$$= \gamma_*(h_{P_f(\psi, \omega)}) = h_{P_f(\psi, \omega) \circ \gamma},$$

where the inequality above is due to the Jensen inequality for operator convex functions (see [22, Theorem 2.1]). Hence we have the asserted inequality.   □

8.2 Proof of the Theorem

We divide the proof of Theorem 8.4 into two parts. The integral expression of an operator convex function on $(0, 1)$ will be useful in Part 2.

Proof (Part 1) We prove that (iii)–(v) and (viii)–(xii) are equivalent, and they are equivalent to (xiii) when γ is 2-positive.

In view of (4.4), it is obvious that (iii) $\Longrightarrow$ (v) $\Longrightarrow$ (iv). That (x) $\Longrightarrow$ (viii) $\Longrightarrow$ (ix) is trivial. By Proposition D.11 of Appendix D we have (x) $\Longrightarrow$ (iii) and (ix) $\Longrightarrow$ (iv).

(xi) $\Longleftrightarrow$ (v). Note that

$$
\begin{aligned}
h_{\psi_0}^{1/2} A_0 A_0^* h_{\psi_0}^{1/2} &= h_{\omega_0}^{1/2} (A_0^* A_0)^2 h_{\omega_0}^{1/2} \\
&= h_{P_{t^2}(\psi_0,\omega_0)} \le h_{P_{t^2}(\psi,\omega)\circ\gamma} \quad \text{(by Lemma 8.6)} \\
&= \gamma_*(h_\omega^{1/2}(A^*A)^2 h_\omega^{1/2}) = \gamma_*(h_\psi^{1/2} AA^* h_\psi^{1/2}).
\end{aligned}
$$

Hence (xi) is equivalent to

$$
\operatorname{tr} h_{\psi_0}^{1/2} A_0 A_0^* h_{\psi_0}^{1/2} = \operatorname{tr} h_\psi^{1/2} AA^* h_\psi^{1/2},
$$

or equivalently, $\operatorname{tr} h_{\omega_0}^{1/2}(A_0^* A_0)^2 h_{\omega_0}^{1/2} = \operatorname{tr} h_\omega^{1/2}(A^*A)^2 h_\omega^{1/2}$, which is (v) by (4.4).

(xi) $\Longleftrightarrow$ (xii). It follows from (8.4) that

$$
h_{\omega_0}^{1/2} \gamma_\omega^*((A^*A)^2) h_{\omega_0}^{1/2} = \gamma_*(h_\omega^{1/2}(A^*A)^2 h_\omega^{1/2}) = \gamma_*(h_\psi^{1/2} AA^* h_\psi^{1/2}).
$$

It also follows from Lemma 8.5 that

$$
h_{\omega_0}^{1/2}(\gamma_\omega^*(A^*A))^2 h_{\omega_0}^{1/2} = h_{\omega_0}^{1/2}(A_0^*A_0)^2 h_{\omega_0}^{1/2} = h_{\psi_0}^{1/2} A_0 A_0^* h_{\psi_0}^{1/2}.
$$

Hence (xi) is equivalent to $h_{\omega_0}^{1/2} \gamma_\omega^*((A^*A)^2) h_{\omega_0}^{1/2} = h_{\omega_0}^{1/2}(\gamma_\omega^*(A^*A))^2 h_{\omega_0}^{1/2}$, which implies (xii) as in the proof of Lemma 8.5. The converse is obvious.

(xii) $\Longrightarrow$ (x). Let $\widehat{\gamma_\omega^*}$ be the restriction of γ_ω^* to the commutative von Neumann subalgebra of M generated by $T_{\psi/\omega}$ and 1. Then condition (xii) implies that $T_{\psi/\omega}$ is in the multiplicative domain $\mathcal{M}_{\widehat{\gamma_\omega^*}}$ of $\widehat{\gamma_\omega^*}$ (see Definition 6.8). Then, since the restriction of $\widehat{\gamma_\omega^*}$ to $\mathcal{M}_{\widehat{\gamma_\omega^*}}$ is a $*$-homomorphism by Proposition 6.9, we have, for any Borel function ϕ on $[0, \infty)$,

$$
\phi(\gamma_\omega^*(T_{\psi/\omega})) = \phi(\widehat{\gamma_\omega^*}(T_{\psi/\omega})) = \widehat{\gamma_\omega^*}(\phi(T_{\psi/\omega})) = \gamma_\omega^*(\phi(T_{\psi/\omega})),
$$

which implies by Lemma 8.5 that $\phi(T_{\psi_0/\omega_0}) = \gamma_\omega^*(\phi(T_{\psi/\omega}))$. Multiplying $h_{\omega_0}^{1/2}$ from both sides gives

$$h_{\omega_0}^{1/2}\phi(T_{\psi_0/\omega_0})h_{\omega_0}^{1/2} = h_{\varphi_0}^{1/2}\gamma_\omega^*(\phi(T_{\psi/\omega}))h_{\omega_0}^{1/2} = \gamma_*(h_\omega^{1/2}\phi(T_{\psi/\omega})h_\omega^{1/2})$$

by Lemma 8.3. The above equality is (x).

(iv) $\implies$ (v). Recall [53, Theorem 8.1] that an operator convex function f on $[0,\infty)$ has the integral expression

$$f(t) = f(0^+) + at + bt^2 + \int_{(0,\infty)}\left(\frac{t}{1+s} - \frac{t}{t+s}\right)d\mu(s), \qquad t \in (0,\infty),$$

where $a \in \mathbb{R}$, $b \geq 0$ and μ is a positive measure on $(0,\infty)$ with $\int_{(0,\infty)}(1+s)^{-2}\,d\mu(s) < +\infty$. With $f_s(t) := -t/(t+s)$ $(t \in [0,\infty))$ for $s > 0$, one has

$$\widehat{S}_f(\psi\|\omega) = f(0^+)\omega(1) + a\psi(1) + bS_{t^2}(\psi\|\omega)$$
$$+ \int_{(0,\infty)}\left(\frac{\psi(1)}{1+s} + \widehat{S}_{f_s}(\psi\|\omega)\right)d\mu(s), \qquad (8.5)$$

$$\widehat{S}_f(\psi_0\|\omega_0) = f(0^+)\omega_0(1) + a\psi_0(1) + bS_{t^2}(\psi_0\|\omega_0)$$
$$+ \int_{(0,\infty)}\left(\frac{\psi_0(1)}{1+s} + \widehat{S}_{f_s}(\psi_0\|\omega_0)\right)d\mu(s)$$
$$= f(0^+)\omega(1) + a\psi(1) + bS_{t^2}(\psi_0\|\omega_0)$$
$$+ \int_{(0,\infty)}\left(\frac{\psi(1)}{1+s} + \widehat{S}_{f_s}(\psi_0\|\omega_0)\right)d\mu(s). \qquad (8.6)$$

By comparing (8.5) and (8.6) in view of the monotonicity in Theorem 4.4(1) one has

$$S_{t^2}(\psi_0\|\omega_0) = \widehat{S}_{t^2}(\psi\|\omega) \quad \text{if } b > 0,$$
$$\widehat{S}_{f_s}(\psi_0\|\omega_0) = \widehat{S}_{f_s}(\psi\|\omega) \quad \text{for } \mu\text{-a.e. } s > 0.$$

Since f is nonlinear, $b > 0$ or $\mu \neq 0$. So it suffices to show that (v) holds if $\widehat{S}_{f_s}(\psi_0\|\omega_0) = \widehat{S}_{f_s}(\psi\|\omega)$ for some $s \in (0,\infty)$. Since $f_s(A^*A) = -1 + s(A^*A + s1)^{-1}$ and similarly for $f_s(A_0^*A_0)$, the assumption means that for some $s \in (0,\infty)$,

$$\operatorname{tr} h_{\omega_0}^{1/2}(A_0^*A_0 + e_0)^{-1}h_{\omega_0}^{1/2} = \operatorname{tr} h_\omega^{1/2}(A^*A + se)^{-1}h_\omega^{1/2}, \qquad (8.7)$$

where $(A^*A + se)^{-1}$ is the inverse in eMe and $(A_0^*A_0 + se_0)^{-1}$ is in e_0Ne_0. We have $h_\omega^{1/2} = Bh_{\psi+s\omega}^{1/2}$ with an invertible $B \in eMe$ and $h_{\omega_0}^{1/2} = B_0h_{\psi_0+s\omega_0}^{1/2}$ with an invertible $B_0 \in e_0Ne_0$. Note that

$$h_{\psi+s\omega}^{1/2} = B^{-1}h_\omega^{1/2} = h_\omega^{1/2}B^{*-1}, \quad h_{\psi_0+s\omega_0}^{1/2} = B_0^{-1}h_{\omega_0}^{1/2} = h_{\omega_0}^{1/2}B_0^{*-1}. \qquad (8.8)$$

Since

$$h_\omega^{1/2}(A^*A + se)h_\omega^{1/2} = h_\psi + sh_\omega = h_{\psi+s\omega} = h_\omega^{1/2}(BB^*)^{-1}h_\omega^{1/2},$$

we have $A^*A + se = (BB^*)^{-1}$. Similarly, $A_0^*A_0 + se_0 = (B_0B_0^*)^{-1}$. Hence (8.7) means that

$$\operatorname{tr} h_{\omega_0}^{1/2} B_0 B_0^* h_{\omega_0}^{1/2} = \operatorname{tr} h_\omega^{1/2} BB^* h_\omega^{1/2},$$

which implies that (v) holds for $\omega, \psi + s\omega$ in place of ψ, ω (see the proof of (xi) $\Longleftrightarrow$ (v) above). Here note that the above proofs of (v) $\Longleftrightarrow$ (xi) $\Longleftrightarrow$ (xii) $\Longrightarrow$ (x) $\Longrightarrow$ (iii) have been done in the setting of arbitrary $(\psi, \omega) \in (M_*^+ \times M_*^+)_\leq$. So we find that (iii) holds for $\omega, \psi + s\omega$ in place of ψ, ω. Applying this to $f(t) = (t + \varepsilon)^{-1}$ for any $\varepsilon > 0$ gives

$$\operatorname{tr} h_{\psi_0 + s\omega_0}^{1/2} (B_0^* B_0 + \varepsilon e_0)^{-1} h_{\psi_0 + s\omega_0}^{1/2} = \operatorname{tr} h_{\psi + s\omega}^{1/2} (B^* B + \varepsilon e)^{-1} h_{\psi + s\omega}^{1/2}.$$

Letting $\varepsilon \searrow 0$ gives

$$\operatorname{tr} h_{\psi_0 + s\omega_0}^{1/2} (B_0^* B_0)^{-1} h_{\psi_0 + s\omega_0}^{1/2} = \operatorname{tr} h_{\psi + s\omega}^{1/2} (B^* B)^{-1} h_{\psi + s\omega}^{1/2}.$$

Therefore, from (8.8) we have

$$\operatorname{tr} h_{\omega_0}^{1/2} (B_0 B_0^*)^{-2} h_{\omega_0}^{1/2} = \operatorname{tr} h_\omega^{1/2} (BB^*)^{-2} h_\omega^{1/2},$$

i.e.,

$$\operatorname{tr} h_{\omega_0}^{1/2} (A_0^* A_0 + se_0)^2 h_{\omega_0}^{1/2} = \operatorname{tr} h_\omega^{1/2} (A^* A + se)^2 h_\omega^{1/2},$$

which means by (4.4) that

$$S_{t^2}(\psi_0 \| \omega_0) + 2s\psi_0(1) + s^2\omega_0(1) = S_{t^2}(\psi \| \omega) + 2s\psi(1) + s^2\omega(1)$$

so that (v) holds.

Finally, assume that γ is 2-positive. Then so is γ_ω^*, as mentioned just before Example (6.7), and (xii) $\Longleftrightarrow$ (xiii) is obvious by Definition 6.8. $\square$

Proof (Part 2) We prove that (i), (ii), (vi) and (vii) are equivalent to the conditions treated in Part 1.

(x) $\Longrightarrow$ (i). Let f be any operator convex function on $(0, \infty)$. Set $g(t) := (1 - t)f(\frac{t}{1-t})$ for $t \in [0, 1]$, where $g(0) := f(0^+)$ and $g(1) := f'(\infty)$ are in $(-\infty, +\infty]$. For each $n \in \mathbb{N}$ let $g_n(t) := g(t) \wedge n$, $t \in [0, 1]$. Then by (x) we have

$$h_{\omega_0}^{1/2} g_n(T_{\psi_0/\omega_0}) h_{\omega_0}^{1/2} = \gamma_*(h_\omega^{1/2} g_n(T_{\psi/\omega}) h_\omega^{1/2}).$$

Taking tr of both sides above yields that

$$\int_0^1 g_n(t) \, d\|E_{\psi_0/\omega_0}(t) h_{\omega_0}^{1/2}\|^2 = \int_0^1 g_n(t) \, d\|E_{\psi/\omega}(t) h_\omega^{1/2}\|^2,$$

where $T_{\psi/\omega} = \int_0^1 t \, dE_{\psi/\omega}(t)$ and $T_{\psi_0/\omega_0} = \int_0^1 t \, dE_{\psi_0/\omega_0}(t)$ are the spectral decompositions. By the monotone convergence theorem letting $n \to \infty$ gives

$$\int_0^1 g(t) \, d\|E_{\psi_0/\omega_0}(t)h_{\omega_0}^{1/2}\|^2 = \int_0^1 g(t) \, d\|E_{\psi/\omega}(t)h_\omega^{1/2}\|^2,$$

which means that $\widehat{S}_f(\psi_0\|\varphi_0) = \widehat{S}_f(\psi\|\varphi)$ by Theorem 4.8.

(i) $\Longrightarrow$ (vi). Let σ be any operator connection with the representing operator monotone function $k_\sigma \geq 0$ on $[0, \infty)$. Then by (i) for $f = -k_\sigma$ and by Proposition D.11 of Appendix D we have

$$(\varphi_0\sigma\psi_0)(1_N) = -\widehat{S}_{-k_\sigma}(\psi_0\|\varphi_0) = -\widehat{S}_{-k_\sigma}(\psi\|\varphi)$$
$$= (\varphi\sigma\psi)(1_M) = ((\varphi\sigma\psi) \circ \gamma)(1_N).$$

Since $\varphi_0\sigma\psi_0 \geq (\varphi\sigma\psi) \circ \gamma$ by Proposition D.10, (vi) follows.

(vi) $\Longrightarrow$ (vii) is trivial.

(vii) $\Longrightarrow$ (ii). Assume that (vii) holds for a nonlinear operator connection σ with the representation function of k_σ; then $f := -k_\sigma$ is a nonlinear operator convex function on $[0, \infty)$ with $f'(\infty) < +\infty$. Similarly to the proof of (i) $\Longrightarrow$ (vi) we have

$$-\widehat{S}_f(\psi_0\|\varphi_0) = (\varphi_0\sigma\psi_0)(1_N) = (\varphi\sigma\psi)(1_M) = -\widehat{S}_f(\psi\|\varphi).$$

(ii) $\Longrightarrow$ (v). Assume that (ii) holds with f as stated, and let $g(t) := (1-t)f\left(\frac{t}{1-t}\right)$ for $t \in [0, 1]$, where $g(0) := f(0^+)$ and $g(1) := f'(\infty) \in (-\infty, +\infty]$. Then by Theorem 4.8 we have

$$\int_0^1 g(t) \, d\|E_{\psi_0/\omega_0}(t)h_{\omega_0}^{1/2}\|^2 = \int_0^1 g(t) \, d\|E_{\psi/\omega}(t)h_\omega^{1/2}\|^2 \tag{8.9}$$

with the spectral measures $E_{\varphi/\omega}$ of $T_{\varphi/\omega}$ and E_{φ_0/ω_0} of T_{φ_0/ω_0}. Since

$$\lim_{t\to 0^+} tg(t) = \lim_{t\to 0^+} tf(t) = 0, \qquad \lim_{t\to 1^-} (1-t)g(t) = \lim_{t\to\infty} \frac{f(t)}{t^2} = 0,$$

it follows from Corollary C.3 that g has the integral expression

$$g(t) = \alpha + \beta t + \int_{(-\infty, 1)} \phi_s(t) \, d\nu(s), \qquad t \in [0, 1], \tag{8.10}$$

where $\alpha, \beta \in \mathbb{R}$ and ν is a positive measure ν on $(-\infty, 1)$ with $\int_{(-\infty, 1)} (2-s)^{-1} \, d\nu(s) < +\infty$, and we set

$$\phi_s(t) := \frac{(2t-1)^2}{1 - st}, \qquad t \in [0, 1], \ s \in (-\infty, 1).$$

By Fubini's theorem the LHS of (8.9) is

$$\alpha\omega_0(1) + \beta\psi_0(1) + \int_{(-\infty,1)}\left[\int_0^1 \phi_s(t)\, d\|E_{\psi_0/\omega_0}(t)h_{\omega_0}^{1/2}\|^2\right]d\nu(s)$$

$$= \alpha\omega(1) + \beta\psi(1) + \int_{(-\infty,1)} \operatorname{tr} h_{\omega_0}^{1/2}\phi_s(T_{\psi_0/\omega_0})h_{\omega_0}^{1/2}\, d\nu(s)$$

and the RHS of (8.9) is

$$\alpha\omega(1) + \beta\psi(1) + \int_{(-\infty,1)} \operatorname{tr} h_{\omega}^{1/2}\phi_s(T_{\psi/\omega})h_{\omega}^{1/2}\, d\nu(s).$$

Therefore, (8.9) implies that

$$\int_{(-\infty,1)} \operatorname{tr} h_{\omega_0}^{1/2}\phi_s(T_{\psi_0/\omega_0})h_{\omega_0}^{1/2}\, d\nu(s) = \int_{(-\infty,1)} \operatorname{tr} h_{\omega}^{1/2}\phi_s(T_{\psi/\omega})h_{\omega}^{1/2}\, d\nu(s). \quad (8.11)$$

For every $s \in (-\infty, 1)$, since ϕ_s is operator convex on $[0, 1]$, note by Lemma 8.5 that

$$\begin{aligned} h_{\omega_0}^{1/2}\phi_s(T_{\psi_0/\omega_0})h_{\omega_0}^{1/2} &= h_{\omega_0}^{1/2}\phi_s(\gamma_{\omega}^*(T_{\psi/\omega}))h_{\omega_0}^{1/2} \\ &\leq h_{\omega_0}^{1/2}\gamma_{\omega}^*(\phi_s(T_{\psi/\omega}))h_{\omega_0}^{1/2} = \gamma_*(h_{\omega}^{1/2}\phi_s(T_{\psi/\omega})h_{\omega}^{1/2}), \end{aligned}$$

where we have used [22, Theorem 2.1] for the inequality above and the last equality is due to (8.4). Therefore,

$$\operatorname{tr} h_{\omega_0}^{1/2}\phi_s(T_{\psi_0/\omega_0})h_{\omega_0}^{1/2}) \leq \operatorname{tr} h_{\omega}^{1/2}\phi_s(T_{\psi/\omega})h_{\omega}^{1/2}), \qquad s \in (-\infty, 1). \quad (8.12)$$

Since f is nonlinear and so is g, note that $\nu((-\infty, 1)) > 0$ in expression (8.10). Hence by (8.11) and (8.12) we find that there exists an $s \in (-\infty, 1)$ such that

$$\operatorname{tr} h_{\omega_0}^{1/2}\phi_s(T_{\psi_0/\omega_0})h_{\omega_0}^{1/2} = \operatorname{tr} h_{\omega}^{1/2}\phi_s(T_{\psi/\omega})h_{\omega}^{1/2}. \quad (8.13)$$

When $s = 0$, since $\phi_0(t) = (2t-1)^2$, equality (8.13) implies that $\operatorname{tr} h_{\omega_0}^{1/2}T_{\psi_0/\omega_0}^2 h_{\omega_0}^{1/2} = \operatorname{tr} h_{\omega}^{1/2}T_{\psi/\omega}^2 h_{\omega}^{1/2}$ so that $\widehat{S}_{t^2}(\psi_0\|\omega_0) = \widehat{S}_{t^2}(\psi\|\omega)$, i.e., (v) holds.

Next, when $s \neq 0$, note that

$$\phi_s(t) = \frac{(1 - \frac{2}{s} + \frac{2}{s}(1-st))^2}{1-st} = \left(1 - \frac{2}{s}\right)^2 \frac{1}{1-st} - \frac{4(1-s)}{s^2} - \frac{4}{s}t.$$

Hence, (8.13) implies that

$$\operatorname{tr} h_{\omega_0}^{1/2}(e_0 - sA_0^*A_0)^{-1}h_{\omega_0}^{1/2}) = \operatorname{tr}(h_{\omega}^{1/2}(e - sA^*A)^{-1}h_{\omega}^{1/2}). \quad (8.14)$$

Now, since $\omega \sim (1-s)\psi + \varphi$, by Lemma A.21 choose invertible $B \in eMe$ and $B_0 \in e_0 N e_0$ such that

$$h_{\omega}^{1/2} = B h_{(1-s)\psi+\varphi}^{1/2}, \qquad h_{\omega_0}^{1/2} = B_0 h_{(1-s)\psi_0+\varphi_0}^{1/2}.$$

We then have

$$h_\omega^{1/2}(e - sA^*A)h_\omega^{1/2} = h_\omega - sh_\psi = h_{(1-s)\psi+\varphi} = h_\omega^{1/2}(BB^*)^{-1}h_\omega^{1/2},$$

since $h_{(1-s)\psi+\varphi}^{1/2} = B^{-1}h_\omega^{1/2} = h_\omega^{1/2}B^{*-1}$. Therefore, we have $e - sA^*A = (BB^*)^{-1}$ so that $(e - sA^*A)^{-1} = BB^*$. Similarly, $e_0 - sA_0^*A_0 = (B_0B_0^*)^{-1}$ so that $(e_0 - sA_0^*A_0)^{-1} = B_0B_0^*$. Thus, (8.14) means that

$$\operatorname{tr} h_{\omega_0}^{1/2} B_0 B_0^* h_{\omega_0}^{1/2} = \operatorname{tr} h_\omega^{1/2} BB^* h_\omega^{1/2},$$

which implies from the proof of (xi) $\Longleftrightarrow$ (v) in Part 1 that (v) holds for ω, $(1-s)\psi+\varphi$ in place of ψ, ω. Furthermore, similarly to an argument in the proof of (iv) $\Longrightarrow$ (v) in Part 1 based on (v) $\Longrightarrow$ (iii), we find that

$$\operatorname{tr} h_{\omega_0}^{1/2}(B_0B_0^*)^{-2}h_{\omega_0}^{1/2} = \operatorname{tr} h_\omega^{1/2}(BB^*)^{-2}h_\omega^{1/2}.$$

Since $(BB^*)^{-2} = (e - sA^*A)^2$ and $(B_0B_0^*)^{-2} = (e_0 - sA_0^*A_0)^2$, we have

$$\omega_0(1) - 2s\psi_0(1) + s^2 S_{t^2}(\psi_0\|\omega_0) = \omega(1) - 2s\psi(1) + s^2 S_{t^2}(\psi\|\omega),$$

so that (v) holds. $\qquad\square$

Remark 8.7 (1) When $(\psi, \varphi) \in (M_*^+ \times M_*^+)_\le$, we can replace ω with φ in conditions (v) and (x)–(xiii) of Theorem 8.4. Indeed, all the proofs in Part 1 are valid for ψ, φ themselves instead of replacing φ with $\omega = \psi + \varphi$.

(2) As remarked in Remark 6.21, even in the finite-dimensional case, the preservation of S_{t^2} ($= \widehat{S}_{t^2}$) under a CP map γ does not imply the reversibility of γ. Therefore, by Theorem 8.4 (and (1) above) we see that a CP map γ is not necessarily reversible for $(\psi, \varphi) \in (M_*^+ \times M_*^+)_\le$ even if $\widehat{S}_f(\psi \circ \gamma\|\varphi \circ \gamma) = \widehat{S}_f(\psi\|\varphi)$ holds for all operator convex functions f on $(0, \infty)$. Thus, the preservation of $\widehat{S}_f$ is strictly weaker than that of S_f.

(3) Assume that the domain N of γ is abelian (i.e., γ is a quantum-classical channel). By Theorem 4.18 and Proposition 4.20 condition (iii) implies that $S_f(\psi \circ \gamma\|(\psi + \varphi) \circ \gamma) = S_f(\psi\|\psi + \varphi)$ for any operator convex function f on $[0, \infty)$, which implies by Theorem 6.20 that γ is reversible for $\{\psi, \psi + \varphi\}$. Hence from Theorem 7.6 it follows that $\psi, \psi + \varphi$ commute and so ψ, φ commute. Hence, in this case, the conditions of Theorem 8.4 are also equivalent to the commutativity of ψ, φ. Here note that ψ, φ are arbitrary without the assumption $s(\psi) \le s(\varphi)$, while it is assumed in Theorem 7.6.

Problem 8.8 It is desirable to remove the assumption $\lim_{t\to 0^+} tf(t) = 0$ and $\lim_{t\to\infty} f(t)/t^2 = 0$ on f in condition (ii) of Theorem 8.4. To do this, we need to take care of the case where $f(t) = t^2$ and t^{-1} only (see the proof of (ii) $\Longrightarrow$ (v) in Part 2). Since t^{-1} is the transpose of t^2, the problem is whether $S_{t^2}(\psi \circ \gamma\|\varphi \circ \gamma) = S_{t^2}(\psi\|\varphi) < +\infty$ implies (v) or not. In view of Remark 8.7 (1), this holds true when $(\psi, \varphi) \in (M_*^+ \times M_*^+)_\le$. Since $S_{t^2}(\psi\|\varphi) < +\infty$ implies $s(\psi) \le s(\varphi)$ by Proposition 4.5 (4), the problem holds true, in particular, in the finite-dimensional case, since $s(\psi) \le s(\varphi)$ implies $(\psi, \varphi) \in (M_*^+ \times M_*^+)_\le$ in this case.

Appendix A
Preliminaries on von Neumann algebras

A.1 Introduction of von Neumann algebras

The first section is a very brief introduction of von Neumann algebras. The set $B(\mathcal{H})$ of all bounded linear operators on a Hilbert space $\mathcal{H}$ with the inner product $\langle \cdot, \cdot \rangle$ is a vector space with the operator sum $a + b$ and the scalar multiplication λa ($a, b \in B(\mathcal{H})$, $\lambda \in \mathbb{C}$) and is a Banach space with the operator norm

$$\|a\| := \sup\{\|a\xi\| : \xi \in H, \|\xi\| \leq 1\}.$$

Moreover, $B(\mathcal{H})$ becomes a Banach $*$-algebra with the operator product ab and the adjoint operation $a \mapsto a^*$. A subspace of $B(\mathcal{H})$ is called an algebra if it is closed under the product, and a $*$-subalgebra if it is further invariant under the $*$-operation. In general, *operator algebras* mean $*$-subalgebras of $B(\mathcal{H})$.

The (operator) norm topology, the strong operator topology and the weak operator topology are defined on $B(\mathcal{H})$, which are weaker in this order. Since $B(\mathcal{H})$ is the dual Banach space of the Banach space $C_1(\mathcal{H})$ of trace-class operators with the trace-norm, the weak topology $\sigma(B(\mathcal{H}), C_1(\mathcal{H}))$ is also defined on $B(\mathcal{H})$, which is called the σ-*weak topology* and particularly important in study of von Neumann algebras.

We write $B(\mathcal{H})_{\mathrm{sa}}$ for the set of all self-adjoint (i.e., $a = a^*$) $a \in B(\mathcal{H})$. The order $a \leq b$ for $a, b \in B(\mathcal{H})_{\mathrm{sa}}$ means that $\langle \xi, a\xi \rangle \leq \langle \xi, b\xi \rangle$ for all $\xi \in \mathcal{H}$, which is a partial order on $B(\mathcal{H})$. When a net $\{a_\alpha\}$ in $B(\mathcal{H})_{\mathrm{sa}}$ is increasing and bounded above, it has the supremum $a \in B(\mathcal{H})_{\mathrm{sa}}$ and $a_\alpha \to a$ strongly; in this case, we write $a_\alpha \nearrow a$.

A $*$-subalgebra of $B(\mathcal{H})$ is called a *von Neumann algebra* (also W^*-*algebra*) if it contains the identity operator 1 and closed in the weak topology. For $S \subset B(\mathcal{H})$ define the *commutant* S' of S by

$$S' := \{a \in B(H) : ab = ba \text{ for all } b \in S\},$$

and also $S'' := (S')'$. For a $*$-subalgebra M of $B(\mathcal{H})$ with $1 \in M$, the *double commutation theorem* or *von Neumann's density theorem* says that the following three conditions are equivalent:

(a) M is a von Neumann algebra (i.e., weakly closed);
(b) M is strongly closed;
(c) $M'' = M$.

From this, the polar decompositions and the spectral decompositions of operators in a von Neumann algebra M are takin inside itself and M contains plenty of projections. Starting from this theorem (1929), J. von Neumann developed basics, including the classification as we will explain shortly, of the von Neumann algebra theory in a series of joint papers with F. J. Murray.

C^*-*algebras* are another major subject of operator algebras, which are faithfully represented as norm-closed subalgebras of $B(\mathcal{H})$. Sakai [101] gave the abstract characterization of von Neumann algebras in such a way that a C^*-algebra is isomorphic to a von Neumann algebra if and only if it is the dual Banach space of some Banach space. In this case, the predual space is unique in a strong sense. The term W^*-algebra is often used to stress this abstract (or Hilbert space-free) situation. Although von Neumann algebras are special C^*-algebras, both categories of operator algebras are quite different theoretically and methodologically.

Let M_* be the set of all σ-weakly continuous linear functionals on a von Neumann algebra M, which is a Banach space as a closed subspace of the dual Banach space M^*. Then the dual Banach space of M_* is isometric to M, so M_* is identified with the *predual* of M. A positive linear functional φ on M is in M_* if and only if φ is *normal* (i.e., $a_\alpha \nearrow a \implies \varphi(a_\alpha) \nearrow \varphi(a)$). Let M, N be von Neumann algebras. A $*$-homomorphism $\pi : M \to N$ is normal (i.e., $a_\alpha \nearrow a$ in M implies $\pi(a_\alpha) \nearrow \pi(a)$) if and only if it is continuous with respect to the σ-weak topologies on M, N. A $*$-homomorphism $\pi : M \to B(\mathcal{K})$ ($\mathcal{K}$ is a Hilbert space) is called a $*$-*representation* (or simply representation). The range $\pi(M)$ of a normal representation π is a von Neumann algebra and its kernel is represented as $M(1-e)$ for some central projection e (i.e., a projection in the *center* $Z(M) := M \cap M'$), so π induces a $*$-isomorphism between Me and $\pi(M)$. Note that a faithful representation is normal automatically.

The notion of the *Murray-von Neumann equivalence* on the set $\mathrm{Proj}(M)$ of all projections in a von Neumann algebra M is essential: $e, f \in \mathrm{Proj}(M)$ is said to be equivalent ($e \sim f$) if there is a $v \in M$ such that $v^*v = e$ and $vv^* = f$. A projection $e \in \mathrm{Proj}(M)$ is called an abelian projection if eMe is an abelian algebra, and a finite projection if for $f \in \mathcal{P}(M)$, $e \sim f \leq e \implies f = e$.

The von Neumann algebra M is said to be *finite* if 1 is a finite projection, and *semifinite* if, for every central projection $e \neq 0$ (i.e., a projection in the center of M), there is a finite projection $f \in \mathrm{Proj}(M)$ such that $0 \neq f \leq e$. If M has no finite central projection ($\neq 0$), then M is said to be *properly infinite*. If M has no finite projection ($\neq 0$), then M is said to be *purely infinite* or *of type III*. A von Neumann algebra M is properly infinite if and only if M is isomorphic to the tensor product $M \otimes B(\mathcal{H})$, and M is semifinite if and only if M has a faithful normal and semifinite trace.

A von Neumann algebra M is called a *factor* if the center is trivial (i.e., $Z(M) = \mathbb{C}1$). The factors are classified into one of the following types:

- type I_n if M is isomorphic to the $n \times n$ matrix algebra $\mathbb{M}_n$ for some $n \in \mathbb{N}$, or equivalently, M is finite and has a finite abelian projection $\neq 0$,
- type I_∞ if M is isomorphic to $B(\mathcal{H})$ ($\dim \mathcal{H} = \infty$), or equivalently, M is properly infinite and has an abelian projection $\neq 0$,
- type II_1 if M is finite and has no abelian projection $\neq 0$,
- type II_∞ if M is semifinite and properly infinite, and has no abelian projection $\neq 0$,
- type III.

A finite factor has a faithful finite normal trace (unique up to positive constants). A factor of type II_∞ is represented as (a factor of type II_1) $\otimes B(H)$. Corresponding to the types of factors, the quotient set $\mathrm{Proj}(M)/\sim$ is identified with one of the following: $\{0, 1, \ldots, n\}$ for type I_n, $\{0, 1, \ldots, \infty\}$ for type I_∞, $[0, 1]$ for type II_1, $[0, \infty]$ for type II_∞, and $\{0, \infty\}$ for type III. Von Neumann algebras of type I (i.e., a direct sum of factors of type I) is said to be *discrete* or *atomic*. Type III factors were further classified by Connes [24] as follows: type III_1, type III_λ ($0 < \lambda < 1$), and type III_0.

Assume that a von Neumann algebra M is on a separable Hilbert space, or equivalently, M has the separable predual M_*. If M is generated by an increasing sequence of finite-dimensional $*$-subalgebras, then M is said to be *hyperfinite* or *AFD* (*approximately finite dimensional*). The uniqueness of type II_1 AFD factors is an old result of Murray-von Neumann. A von Neumann algebra $M \subset B(\mathcal{H})$ is said to be *injective* if there exists a norm one projection from $B(\mathcal{H})$ onto M. In [26] Connes proved that a von Neumann algebra is AFD if and only if it is injective, and in the same time, he proved that injective factors of types II_1, II_∞, III_λ ($0 < \lambda < 1$) are unique for each type. Later in 1985, Haagerup [42] proved the uniqueness of injective factors of type III_1, thus completing the classification of AFD (= injective) factors. In other words, the conjugacy class of the *flow of weights* due to Connes and Takesaki [27] is a complete invariant for injective factors of type III.

We end the overview with a list [66, 89, 101, 104, 108, 109] of standard textbooks on von Neumann algebras.

A.2 Tomita-Takesaki Modular Theory

Let M be a von Neumann algebra. Throughout the rest of the chapter, we assume[1] that there is a faithful $\omega \in M_*^+$, or equivalently, M is *σ-finite*, i.e., mutually orthogonal projections in M are at most countable. By making the *GNS cyclic representation* $\{\pi_\omega, \mathcal{H}_\omega, \Omega\}$ of M with respect to ω, we have a $*$-isomorphism $\pi_\omega : M \to B(\mathcal{H}_\omega)$ with a cyclic and separating vector $\Omega \in \mathcal{H}_\omega$ for $\pi_\omega(M)$ such that

---

[1] When M is not σ-finite, the construction of the modular theory is essentially similar to the σ-finite case (but technically more complicated) with a faithful normal semifinite weight on M based on the left Hilbert algebra theory.

$$\omega(x) = \langle \Omega, \pi_\omega(x)\Omega \rangle, \qquad x \in M.$$

Thus, by identifying M with $\pi_\omega(M)$, we may assume that M itself is a von Neumann algebra on $\mathcal{H}$ with a cyclic and separating vector Ω for M. Here, Ω is *cyclic* for M if $\overline{M\Omega} = \mathcal{H}$, and Ω is *separating* for M if $x \in M$, $x\Omega = 0 \implies x = 0$, equivalently $\overline{M'\Omega} = \mathcal{H}$.

We begin to define two conjugate-linear operators S_0 and F_0 with the dense domains $\mathcal{D}(S_0) = M\Omega$ and $\mathcal{D}(F_0) = M'\Omega$ by

$$S_0 x\Omega := x^*\Omega, \qquad x \in M,$$
$$F_0 x'\Omega := x'^*\Omega, \qquad x' \in M'.$$

For any $x \in M$ and $x' \in M'$ note that

$$\langle x'\Omega, S_0 x\Omega \rangle = \langle x'\Omega, x^*\Omega \rangle = \langle x\Omega, x'^*\Omega \rangle = \langle x\Omega, F_0 x'\Omega \rangle,$$

which implies that S_0 and F_0 are closable, $\overline{F_0} \subset S_0^*$ and $\overline{S_0} \subset F_0^*$. So we set $S := \overline{S_0}$ and $F := S^* = S_0^*$, and take the polar decomposition of S as

$$S = J\Delta^{1/2}, \qquad \Delta := S^*S = FS.$$

Since the ranges of S and S^* are dense, it follows that J is a conjugate-linear unitary and Δ is a strictly positive self-adjoint operator. We then have the following:

(1) $J = J^*$ and $J^2 = 1$.
(2) $\Delta = FS$ and $\Delta^{-1} = SF$.
(3) $S = J\Delta^{1/2} = \Delta^{-1/2}J$ and $F = J\Delta^{-1/2} = \Delta^{1/2}J$.
(4) $\Delta^{-1} = J\Delta J$ and $J\Delta^{it} = \Delta^{-it}J$ for all $t \in \mathbb{R}$.
(5) $J\Omega = \Omega$ and $\Delta\Omega = \Omega$.

The next theorem is *Tomita's fundamental theorem.*

Theorem A.1 (Tomita) *With J and Δ given above, we have*

$$JMJ = M', \tag{A.1}$$
$$\Delta^{it} M\Delta^{-it} = M, \qquad t \in \mathbb{R}. \tag{A.2}$$

Definition A.2 The operator Δ is called the *modular operator* with respect to Ω (or ω), and J is called the *modular conjugation* with respect to Ω (or ω). The one-parameter automorphism group $\sigma_t = \sigma_t^\omega$ of M defined by $\sigma_t^\omega(x) := \Delta^{it} x\Delta^{-it}$ ($x \in M, t \in \mathbb{R}$) is called the *modular automorphism group* with respect to Ω (or ω).

Proposition A.3 *If $x \in M \cap M'$ (the center of M), then*

$$JxJ = x^*, \qquad \sigma_t(x) = \Delta^{it} x\Delta^{-it} = x, \qquad t \in \mathbb{R}.$$

Example A.4 Consider the simple case where $\omega = \tau$ is a faithful normal finite trace of M, so M is a finite von Neumann algebra of type II$_1$. Since $\|x\Omega\|^2 = \tau(x^*x) =$

$\tau(xx^*) = \|x^*\Omega\|^2$ for all $x \in M$, S is a conjugate-linear unitary, which means that $S = J$ and $\Delta = 1$. Hence (A.2) trivially holds. For every $x, y, y_1 \in M$ note that

$$JxJyy_1\Omega = Jxy_1^*y^*\Omega = yy_1x^*\Omega = yJxy_1^*\Omega = yJxJy_1\Omega$$

so that $JxJy = yJxJ$. Hence $JMJ \subset M'$. Moreover, for every $x \in M$ and $x' \in M'$ note that

$$\langle x\Omega, Jx'\Omega \rangle = \langle x'\Omega, Jx\Omega \rangle = \langle x'\Omega, x^*\Omega \rangle = \langle x\Omega, x'^*\Omega \rangle$$

so that $Jx'\Omega = x'^*\Omega$. Hence, similarly to the above, $JM'J \subset M$, so $M' \subset JMJ$. Therefore, (A.1) holds. In this way, Tomita's theorem in this case is quite easy.

Let α_t ($t \in \mathbb{R}$) be a one-parameter weakly continuous (hence automatically strongly* continuous) automorphism group of M. For $\beta \in \mathbb{R}$ with $\beta \neq 0$, if $\beta < 0$,

$$D_\beta := \{z \in \mathbb{C} : 0 < \operatorname{Im} z < -\beta\}, \qquad \overline{D}_\beta := \{z \in \mathbb{C} : 0 \leq \operatorname{Im} z \leq -\beta\},$$

and if $\beta > 0$, $D_\beta := \{z \in \mathbb{C} : -\beta < \operatorname{Im} z < 0\}$ and $\overline{D}_\beta$ is similar.

Definition A.5 A functional $\varphi \in M_*^+$ is said to satisfy the *KMS (Kubo-Martin-Schwinger) condition* with respect to α_t at β, or the (α_t, β)-KMS condition, if for every $x, y \in M$ there is a bounded continuous function $f_{x,y}(z)$ on $\overline{D}_\beta$, that is analytic in D_β, such that

$$f_{x,y}(t) = \varphi(\alpha_t(x)y), \qquad f_{x,y}(t - i\beta) = \varphi(y\alpha_t(x)), \qquad t \in \mathbb{R}.$$

This condition proposed by Haag, Hugenholtz and Winnink serves as a mathematical formulation of equilibrium states in the quantum statistical mechanics, which is also defined and more useful in C^*-algebraic dynamical systems, see [21]. To illustrate this, given a Hamiltonian $H \in B(\mathcal{H})_{\mathrm{sa}}$ in a finite-dimensional $\mathcal{H}$, consider the Gibbs state $\varphi(x) = \operatorname{Tr} e^{-\beta H} x / \operatorname{Tr} e^{-\beta H}$ and the corresponding dynamics $\alpha_t(x) = e^{itH} x e^{-itH}$, $t \in \mathbb{R}$. For any $x, y \in B(\mathcal{H})$ the entire function $f_{x,y}(z) := \operatorname{Tr}(e^{-\beta H} e^{izH} x e^{-izH} y) / \operatorname{Tr} e^{-\beta H}$ satisfies

$$f_{x,y}(t) = \operatorname{Tr}(e^{-\beta H} e^{itH} x e^{-itH} y) / \operatorname{Tr} e^{-\beta H} = \varphi(\alpha_t(x)y),$$

$$f_{x,y}(t - i\beta) = \operatorname{Tr}(e^{itH} x e^{-itH} e^{-\beta H} y) / \operatorname{Tr} e^{-\beta H} = \varphi(y\alpha_t(x)), \qquad t \in \mathbb{R},$$

so that φ satisfies the (α_t, β)-KMS condition. Moreover, the Gibbs state φ is a unique state satisfying the (α_t, β)-KMS condition.

The following proposition is another justification for the KMS condition to describe equilibrium states.

Proposition A.6 *If $\varphi \in M_*^+$ satisfies the (α_t, β)-KMS condition, then $\varphi \circ \alpha_t = \varphi$ for all $t \in \mathbb{R}$.*

The following is Takesaki's theorem in [105], an important ingredient of the Tomita-Takesaki theory in addition to Tomita's theorem.

Theorem A.7 (Takesaki) *In the same situation as in Theorem A.1, the ω satisfies the $(\sigma_t^\omega, -1)$-KMS condition. Furthermore, σ_t^ω is uniquely determined as a weakly continuous one-parameter automorphism group of M for which ω satisfies the KMS condition at $\beta = -1$.*

Definition A.8 The *centralizer* of ω is defined as

$$M_\omega := \{x \in M : \omega(xy) = \omega(yx),\ y \in M\}.$$

It is obvious that M_ω is a subalgebra of M including the center $M \cap M'$.

Proposition A.9 *The centralizer M_ω coincides with the fixed-point algebra of σ_t^ω, i.e.,*

$$M_\omega = \{x \in M : \sigma_t^\omega(x) = x,\ t \in \mathbb{R}\}.$$

Proposition A.9 contains the second assertion of Proposition A.3. Also, it follows that ω is a trace if and only if $\sigma_t = \mathrm{id}$ for all $t \in \mathbb{R}$.

The notion of conditional expectations in the von Neumann algebra setting was first introduced by Umegaki [113]. If M_0 is a von Neumann subalgebra of M and $E : M \to M_0$ is a norm one projection onto M_0, then we call E a *conditional expectation* onto M_0. Tomiyama [112] showed that even when M is a C^*-algebra and M_0 is a C^*-subalgebra of M, a norm one projection $E : M \to M_0$ satisfies (1) E is positive, (2) $E(y_1 x y_2) = y_1 E(x) y_2$ for all $x \in M$ and $y_1, y_2 \in N$, and (3) $E(x)^* E(x) \leq E(x^* x)$ for all $x \in M$. In fact, it is well-known that E is completely positive, see the first paragraph of Sec. 6.1.

Takesaki's theorem [106] characterizes von Neumann subalgebras onto which the normal conditional expectation exists for a given faithful $\omega \in M_*^+$.

Theorem A.10 ([106]) *For a von Neumann subalgebra M_0 of M and a faithful $\omega \in M_*^+$, there exists a (unique faithful normal) conditional expectation $E : M \to M_0$ satisfying $\omega = \omega \circ E$ on M if and only if M_0 is globally invariant under σ_t^ω, i.e., $\sigma_t^\omega(M_0) = M_0$ for all $t \in \mathbb{R}$.*

Furthermore, in this case, E is unique and $\sigma_t^{\omega|_{M_0}} = \sigma_t^\omega|_{M_0}$ for any $t \in \mathbb{R}$.

A.3 Standard Forms

The theory of the standard form of von Neumann algebras was developed, independently, by Araki [7] and Connes [25] in the case of σ-finite von Neumann algebras, and by Haagerup [37, 38] in the general case.

Let M be a σ-finite von Neumann algebra, thus represented on a Hilbert space $\mathcal{H}$ with a cyclic and separating vector Ω, for which we have the modular operator Δ and the modular conjugation J, as in Sec. A.2. Let $j : M \to M'$ be the conjugate-linear $*$-isomorphism defined by $j(x) := JxJ$, $x \in M$.

Definition A.11 The *natural positive cone* $\mathcal{P} = \mathcal{P}^\natural$ in $\mathcal{H}$ associated with (M, Ω) is defined by

$$\mathcal{P} := \overline{\{xj(x)\Omega : x \in M\}} = \overline{\{xJx\Omega : x \in M\}}. \tag{A.3}$$

Theorem A.12 *We have:*

(1) $\mathcal{P} = \overline{\Delta^{1/4} M_+ \Omega} = \overline{\Delta^{-1/4} M'_+ \Omega}$. *In particular, $\mathcal{P}$ is a closed cone.*
(2) $J\xi = \xi$ *for all $\xi \in \mathcal{P}$.*
(3) $\Delta^{it} \mathcal{P} = \mathcal{P}$ *for all $t \in \mathbb{R}$.*
(4) $xj(x)\mathcal{P} \subset \mathcal{P}$ *for all $x \in M$.*
(5) $\mathcal{P}$ *is self-dual, i.e., $\mathcal{P} = \{\eta \in \mathcal{H} : \langle \xi, \eta \rangle \geq 0, \ \xi \in \mathcal{P}\}$.*

We thus conclude that any (σ-finite) von Neumann algebra is faithfully represented on a Hilbert space $\mathcal{H}$ with a conjugate-linear involution J and a self-dual cone $\mathcal{P}$ such that

(a) $JMJ = M'$ (Theorem A.1),
(b) $JxJ = x^*$ for all $x \in M \cap M'$ (Proposition A.3),
(c) $J\xi = \xi$ for all $\xi \in \mathcal{P}$,
(d) $xj(x)\mathcal{P} \subset \mathcal{P}$ for all $x \in M$, where $j(x) := JxJ$.

Definition A.13 A quadruple $(M, \mathcal{H}, J, \mathcal{P})$ satisfying the above conditions (i)–(iv) is called a *standard form* of a von Neumann algebra M. This is the abstract (or axiomatic) definition, and we have shown the existence of a standard form for any σ-finite von Neumann algebra.

Example A.14 (1) Let $(X, \mathcal{X}, \mu)$ be a σ-finite (or more generally, localizable) measure space. The commutative von Neumann algebra $M = L^\infty(X, \mu)$ is faithfully represented on the Hilbert space $L^2(X, \mu)$ as multiplication operators $\pi(f)\xi := f\xi$ for $f \in L^\infty(X, \mu)$ and $\xi \in L^2(X, \mu)$. The standard form of M is

$$(L^\infty(X, \mu), L^2(X, \mu), J\xi = \overline{\xi}, L^2(X, \mu)_+),$$

where $L^2(X, \mu)_+$ is the cone of non-negative functions $\xi \in L^2(X, \mu)$.

(2) Let $M = B(\mathcal{H})$, a factor of type I. Let $C_2(\mathcal{H})$ be the space of Hilbert-Schmidt operators (i.e., $a \in B(\mathcal{H})$ with $\operatorname{Tr} a^*a < +\infty$), which is a Hilbert space with the Hilbert-Schmidt inner product $\langle a, b \rangle := \operatorname{Tr} a^*b$ for $a, b \in C_2(\mathcal{H})$. Then $M = B(\mathcal{H})$ is faithfully represented on $C_2(\mathcal{H})$ as left multiplication operators $\pi(x)a := xa$ for $x \in B(\mathcal{H})$ and $a \in C_2(\mathcal{H})$. The standard form of M is

$$(B(\mathcal{H}), C_2(\mathcal{H}), J = \ ^*, C_2(\mathcal{H})_+),$$

where $J = \ ^*$ is the adjoint operation and $C_2(\mathcal{H})_+$ is the cone of positive $a \in C_2(\mathcal{H})$. In this case, note that for any $x \in B(\mathcal{H})$, JxJ is the right multiplication of x^* on $C_2(\mathcal{H})$ and $xj(x)C_2(\mathcal{H})_+ = xC_2(\mathcal{H})_+x^* \subset C_2(\mathcal{H})_+$.

The following gives geometric properties of the cone $\mathcal{P}$.

Proposition A.15 *Let $(M, \mathcal{H}, J, \mathcal{P})$ be a standard form. Then:*

(1) $\mathcal{P}$ *is a pointed cone, i.e., $\mathcal{P} \cap (-\mathcal{P}) = \{0\}$.*
(2) *If $\xi \in \mathcal{H}$ and $J\xi = \xi$, then ξ has a unique decomposition $\xi = \xi_1 - \xi_2$ with $\xi_1, \xi_2 \in \mathcal{P}$ and $\xi_1 \perp \xi_2$.*
(3) $\mathcal{H}$ *is linearly spanned by $\mathcal{P}$.*

The next proposition is the description of the standard form of a reduced von Neumann algebra eMe, where e is a projection in M.

Proposition A.16 *Let $(M, \mathcal{H}, J, \mathcal{P})$ be a standard form. Let $e \in M$ be a projection and set $q := ej(e)$. Then:*

(1) $exe \mapsto qxq$ $(x \in M)$ *is a $*$-isomorphism of eMe onto qMq. In particular, $e \neq 0$ $\iff q \neq 0$.*
(2) $qJ = Jq$ *and $(qMq, q\mathcal{H}, qJq, q\mathcal{P})$ is a standard form of $qMq \cong eMe$.*

The next proposition says the universality of $(J, \mathcal{P})$ for the choice of a cyclic and separating vector ξ in $\mathcal{P}$.

Proposition A.17 *Let $(M, \mathcal{H}, J, \mathcal{P})$ be a standard form. If $\xi \in \mathcal{P}$ is cyclic and separating for M, then*

$$J_\xi = J, \qquad \mathcal{P}_\xi = \mathcal{P},$$

where J_ξ is the modular conjugation and $\mathcal{P}_\xi$ is the natural positive cone associated with (M, ξ).

For each $\xi \in \mathcal{H}$ we denote by ω_ξ $(\in M_*^+)$ the vector functional $x \mapsto \langle \xi, x\xi \rangle$ on M. The following is the most important property of the standard form, establishing the relation between $\mathcal{P}$ and M_*^+.

Theorem A.18 *Let $(M, \mathcal{H}, J, \mathcal{P})$ be a standard form. For every $\varphi \in M_*^+$ there exists a $\xi \in \mathcal{P}$ such that $\varphi(x) = \langle \xi, x\xi \rangle$ for all $x \in M$. Furthermore, the following estimates hold:*

$$\|\xi - \eta\|^2 \leq \|\omega_\xi - \omega_\eta\| \leq \|\xi - \eta\| \, \|\xi + \eta\|, \qquad \xi, \eta \in \mathcal{P}.$$

Consequently, the map $\xi \mapsto \omega_\xi$ is a homeomorphism from $\mathcal{P}$ onto M_*^+ when $\mathcal{P}$ and M_*^+ are equipped with the norm topology.

The vector $\xi \in \mathcal{P}$ such that $\omega_\xi = \varphi$ is called the *vector representative* of $\varphi \in M_*^+$.

Another essential property of the standard form is the universality (uniqueness) in the following strict sense.

Theorem A.19 *Let $(M, \mathcal{H}, J, \mathcal{P})$ and $(M_1, \mathcal{H}_1, J_1, \mathcal{P}_1)$ be standard forms of von Neumann algebras of M and M_1, respectively. If $\gamma : M \to M_1$ is a $*$-isomorphism, then there exists a unique unitary $U : \mathcal{H} \to \mathcal{H}_1$ such that*

(1) $\gamma(x) = UxU^*$ *for all $x \in M$,*
(2) $J_1 = UJU^*$,
(3) $\mathcal{P}_1 = U\mathcal{P}$.

Example A.20 Let $(M, \mathcal{H}, J, \mathcal{P})$ be a standard form. We here describe the standard form of $M^{(2)} := M \otimes \mathbb{M}_2$, the tensor product of M with the 2×2 matrix algebra $\mathbb{M}_2$. To do this, choose a cyclic and separating vector $\Omega \in \mathcal{P}$ and consider a faithful $\omega^{(2)} \in (M^{(2)})_*^+$ defined by $\omega^{(2)}\left(\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}\right); = \omega_\Omega(x_{11}) + \omega_\Omega(x_{22})$. Then the GNS cyclic representation $\{\pi^{(2)}, \mathcal{H}^{(2)}, \Omega^{(2)}\}$ of $M^{(2)}$ with respect to $\omega^{(2)}$ is given as follows:

$$\mathcal{H}^{(2)} = \mathcal{H} \oplus \mathcal{H} \oplus \mathcal{H} \oplus \mathcal{H} = \left\{ \begin{bmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{bmatrix} : \xi_{ij} \in \mathcal{H} \right\}, \tag{A.4}$$

$\Omega^{(2)} = \begin{bmatrix} \Omega & 0 \\ 0 & \Omega \end{bmatrix}$, and $\pi^{(2)}\left(\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}\right)$ acts like 2×2 matrix product as $\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}\begin{bmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{bmatrix}$, whose 4×4 representation is

$$\begin{bmatrix} x_{11} & 0 & x_{12} & 0 \\ 0 & x_{11} & 0 & x_{12} \\ x_{21} & 0 & x_{22} & 0 \\ 0 & x_{22} & 0 & x_{22} \end{bmatrix} \begin{bmatrix} \xi_{11} \\ \xi_{12} \\ \xi_{21} \\ \xi_{22} \end{bmatrix} \quad \text{for} \quad \begin{bmatrix} \xi_{11} \\ \xi_{12} \\ \xi_{21} \\ \xi_{22} \end{bmatrix} \in \mathcal{H}^{(2)}. \tag{A.5}$$

Write $S^{(2)}$ and $\Delta^{(2)}$ for $(M^{(2)}, \Omega^{(2)})$ as well as S and Δ for (M, Ω) (see Sec. A.2). Since

$$S^{(2)}\left(\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \Omega^{(2)}\right) = \begin{bmatrix} x_{11}^* \Omega & x_{21}^* \Omega \\ x_{12}^* \Omega & x_{22}^* \Omega \end{bmatrix},$$

one can write

$$S^{(2)} = \begin{bmatrix} S & 0 & 0 & 0 \\ 0 & 0 & S & 0 \\ 0 & S & 0 & 0 \\ 0 & 0 & 0 & S \end{bmatrix} \quad \text{and} \quad \Delta^{(2)} = \begin{bmatrix} \Delta & 0 & 0 & 0 \\ 0 & \Delta & 0 & 0 \\ 0 & 0 & \Delta & 0 \\ 0 & 0 & 0 & \Delta \end{bmatrix}. \tag{A.6}$$

From this with $S = J\Delta^{1/2}$ one has the polar decomposition $S^{(2)} = J^{(2)}(\Delta^{(2)})^{1/2}$ where

$$J^{(2)} = \begin{bmatrix} J & 0 & 0 & 0 \\ 0 & 0 & J & 0 \\ 0 & J & 0 & 0 \\ 0 & 0 & 0 & J \end{bmatrix}. \tag{A.7}$$

Therefore, the standard form of $M^{(2)}$ is given as $(M^{(2)}, \mathcal{H}^{(2)}, J^{(2)}, \mathcal{P}^{(2)})$ with identifications (A.4), (A.5) and (A.7). Moreover, by (A.3), $\mathcal{P}^{(2)} = \overline{\{x J^{(2)} x \Omega^{(2)} : x \in M^{(2)}\}}$. In particular, restricting to $x = \begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix}$ one has $\mathcal{P}^{(2)} \supset \left\{ \begin{bmatrix} \xi & 0 \\ 0 & \eta \end{bmatrix} : \xi, \eta \in \mathcal{P} \right\}$.

We end the section with the next simple lemma, which is often used in the main body of this monograph. So we give a proof for convenience of the reader.

Lemma A.21 *Assume that $\psi \leq \alpha\varphi$ for some $\alpha > 0$. Let $\Phi, \Psi \in \mathcal{P}$ be the vector representatives of φ, ψ, respectively. Then there exists a unique $A \in s(\varphi)Ms(\varphi)$ such that $\Psi = A\Phi$. The A satisfies $\|A\| \leq \alpha^{1/2}$. Moreover, if $\beta\varphi \leq \psi \leq \alpha\varphi$ for some $\alpha, \beta > 0$, then the above A satisfies $\beta s(\varphi) \leq A^*A \leq \alpha s(\varphi)$.*

Proof Since

$$\|x'\Psi\|^2 = \|Jx'J\Psi\|^2 = \psi((Jx'J)^*(Jx'J))$$
$$\leq \alpha\varphi((Jx'J)^*(Jx'J)) = \alpha\|x'\Phi\|^2, \quad x' \in M',$$

there is a unique $A \in B(\mathcal{H})$ such that $As(\varphi) = A$ and $A(x'\Phi) = x'\Psi$ for all $x' \in M$. Moreover, $\|A\| \leq \alpha^{1/2}$, and for every unitary u' in $M's(\varphi)$,

$$u'^*Au'(x'\Phi) = u'^*u'x'\Psi = x'\Psi = A(x'\Phi), \quad x' \in M,$$

which implies that $u'^*Au' = A$, so $A \in (M's(\varphi))' = s(\varphi)Ms(\varphi)$.

Next, assume that $\beta\varphi \leq \psi$ in addition to $\psi \leq \alpha\varphi$. Then $s(\psi) = s(\varphi)$, and there is a unique $B \in s(\varphi)Ms(\varphi)$ such that $\Phi = B\Psi$. It is easy to see that $AB = BA = s(\varphi)$, hence $B = A^{-1}$ in $s(\varphi)Ms(\varphi)$. Since $BB^* \leq \beta^{-1}s(\varphi)$, we have $\beta s(\varphi) \leq A^*A \leq \alpha s(\varphi)$. $\square$

A.4 Relative Modular Operators

The notion of relative modular operators was introduced by Araki [9] to extend the relative entropy to the general von Neumann algebra setting. Let $(M, \mathcal{H}, J, \mathcal{P})$ be a standard form. Let $\psi, \varphi \in M_*^+$, whose vector representatives are $\Psi, \Phi \in \mathcal{P}$, respectively. Define the operators $S_{\psi,\varphi}^0$ and $F_{\psi,\varphi}^0$ by

$$S_{\psi,\varphi}^0(x\Phi + \eta) := s_M(\varphi)x^*\Psi, \qquad x \in M, \ \eta \in (1 - s_{M'}(\varphi))\mathcal{H},$$
$$F_{\psi,\varphi}^0(x'\Phi + \zeta) := s_{M'}(\varphi)x'^*\Psi, \qquad x' \in M', \ \zeta \in (1 - s_M(\varphi))\mathcal{H}.$$

where $s_M(\varphi)$ is the orthogonal projection onto $\overline{M'\Phi}$ (the M-support) and $s_{M'}(\varphi)$ is that onto $\overline{M\Phi}$ (the M'-support). Note that $Js_M(\varphi)J = s_{M'}(\varphi)$. In fact, since $J\Phi = \Phi$ and $JM'J = M$ (Theorem A.1),

$$Js_M(\varphi)J\mathcal{H} = Js_M(\varphi)\mathcal{H} = \overline{JM'\Phi} = \overline{JM'J\Phi} = \overline{M\Phi} = s_{M'}(\varphi)\mathcal{H}.$$

Note that $S_{\psi,\varphi}^0$ and $F_{\psi,\varphi}^0$ are well defined. In fact, assume that $x_1\Phi + \eta_1 = x_2\Phi + \eta_2$ for $x_i \in M$ and $\eta_i \in (1 - s_{M'}(\varphi))\mathcal{H}$. Then $(x_1 - x_2)\Phi = \eta_2 - \eta_1 = 0$, so $\eta_1 = \eta_2$ and $x_1 s_M(\varphi) = x_2 s_M(\varphi)$, so that $s_M(\varphi)x_1^*\Psi = s_M(\varphi)x_2^*\Psi$. Hence $S_{\psi,\varphi}^0$ is well defined and similarly for $F_{\psi,\varphi}^0$. For every $x \in M$, $\eta \in (1 - s_{M'}(\varphi))\mathcal{H}$ and $x' \in M'$, $\zeta \in (1 - s_M(\varphi))\mathcal{H}$, one has

$$\langle S^0_{\psi,\varphi}(x\Phi + \eta), x'\Phi + \zeta\rangle = \langle s_M(\varphi)x^*\Psi, x'\Phi + \zeta\rangle$$
$$= \langle x^*\Psi, x'\Phi\rangle = \langle x'^*\Psi, x\Phi\rangle$$
$$= \langle s_{M'}(\varphi)x'^*\Psi, x\Phi + \eta\rangle$$
$$= \langle F^0_{\psi,\varphi}(x'\Phi + \zeta), x\Phi + \eta\rangle.$$

Since $S^0_{\psi,\varphi}$ and $F^0_{\psi,\varphi}$ have dense domains, they have the closures $S_{\psi,\varphi}$ and $F_{\psi,\varphi}$ satisfying

$$S^*_{\psi,\varphi} = F_{\psi,\varphi}, \qquad F^*_{\psi,\varphi} = S_{\psi,\varphi}. \tag{A.8}$$

Definition A.22 For every $\psi, \varphi \in M^+_*$ define $\Delta_{\psi,\varphi} := S^*_{\psi,\varphi}S_{\psi,\varphi}$, called the *relative modular operator* with respect to ψ, φ.

The notion of the relative modular operator plays a crucial role in this monograph, so we give the next proposition with proof.

Proposition A.23

(1) *The support projection of $\Delta_{\psi,\varphi} := S^*_{\psi,\varphi}S_{\psi,\varphi}$ is $s_M(\psi)s_{M'}(\varphi)$.*

(2) $S_{\psi,\varphi} = J\Delta^{1/2}_{\psi,\varphi}$ *(the polar decomposition).*

(3) $\Delta^{-1}_{\varphi,\psi} = J\Delta_{\psi,\varphi}J.$

Proof (1) By (A.8) the support projection of $S^*_{\psi,\varphi}S_{\psi,\varphi}$ is the orthogonal projection onto the closure of the range of $F_{\psi,\varphi}$, which is $\overline{s_{M'}(\psi)M'\Phi} = s_{M'}(\psi)s_M(\varphi)\mathcal{H}$. Hence the assertion follows.

(2) First, prove the case $\psi = \varphi$. We write $S_\varphi := S_{\varphi,\varphi}$ and $\Delta_\varphi := \Delta_{\varphi,\varphi}$. Let $e := s_M(\varphi)$, $e' := s_{M'}(\varphi)$ and $q := ee' = eJeJ$. By Proposition A.16, $(qMq, q\mathcal{H}, qJq, q\mathcal{P})$ is a standard form of $qMq \cong eMe$. Define $\overline{\varphi} \in (qMq)^+_*$ by $\overline{\varphi}(qxq) := \varphi(exe)$ for $x \in M$, whose vector representative is $\Phi = q\Phi$. Note that Φ is cyclic and separating for qMq on $q\mathcal{H}$. Let $\Delta_{\overline{\varphi}}$ and $J_{\overline{\varphi}}$ be the modular operator and the modular conjugation with respect to $\overline{\varphi}$. Since $S_\varphi((1-e)x\Phi) = ex^*(1-e)\Phi = 0$ for all $x \in M$, $S_\varphi|_{(1-e)e'\mathcal{H}} = 0$ as well as $S_\varphi|_{(1-e')\mathcal{H}} = 0$. Since $(1-e')+(1-e)e' = 1-q$, $S_\varphi|_{(1-q)\mathcal{H}} = 0$. Moreover, for every $x \in M$, note that $ex\Phi = exq\Phi = qxq\Phi$ and $S_\varphi(ex\Phi) = ex^*\Phi = (qxq)^*\Phi$, so $S_\varphi|_{q\mathcal{H}}$ coincides with $S_{\overline{\varphi}}$. Therefore,

$$S_\varphi = S_{\overline{\varphi}} \oplus 0, \quad \text{so} \quad \Delta_\varphi = \Delta_{\overline{\varphi}} \oplus 0$$

on the decomposition $\mathcal{H} = q\mathcal{H} \oplus (1-q)\mathcal{H}$. Since $J_{\overline{\varphi}} = qJ$ by Proposition A.17, one has

$$S_\varphi = J_{\overline{\varphi}}\Delta^{1/2}_{\overline{\varphi}} \oplus 0 = (Jq \oplus J(1-q))(\Delta^{1/2}_{\overline{\varphi}} \oplus 0) = J\Delta^{1/2}_\varphi.$$

Next, prove the case of general ψ, φ. Let $M^{(2)} := M \otimes \mathbb{M}_2$, whose standard form $(M^{(2)}, \mathcal{H}^{(2)}, J^{(2)}, \mathcal{P}^{(2)})$ was described in Example A.20. Define $\theta \in (M^{(2)})^+_*$ by
$$\theta\left(\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}\right) := \varphi(x_{11})+\psi(x_{22}), \text{ whose vector representative in } \mathcal{P}^{(2)} \text{ is } \Theta = \begin{bmatrix} \Phi & 0 \\ 0 & \Psi \end{bmatrix}.$$
It is clear that $s_{M^{(2)}}(\theta) = \begin{bmatrix} s_M(\varphi) & 0 \\ 0 & s_M(\psi) \end{bmatrix}$ or in the 4×4 form,

$$s_{M^{(2)}}(\theta) = \begin{bmatrix} s_M(\varphi) & 0 & 0 & 0 \\ 0 & s_M(\varphi) & 0 & 0 \\ 0 & 0 & s_M(\psi) & 0 \\ 0 & 0 & 0 & s_M(\psi) \end{bmatrix}, \tag{A.9}$$

and also by (A.7),

$$s_{(M^{(2)})'}(\theta) = J^{(2)} s_{M^{(2)}}(\theta) J^{(2)} = \begin{bmatrix} s_{M'}(\varphi) & 0 & 0 & 0 \\ 0 & s_{M'}(\psi) & 0 & 0 \\ 0 & 0 & s_{M'}(\varphi) & 0 \\ 0 & 0 & 0 & s_{M'}(\psi) \end{bmatrix}, \tag{A.10}$$

Furthermore, S_θ is defined as

$$S_\theta\left(\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \Theta + \begin{bmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \end{bmatrix}\right) = s_{M^{(2)}}(\theta) \begin{bmatrix} x_{11}^* & x_{21}^* \\ x_{12}^* & x_{22}^* \end{bmatrix} \Theta \tag{A.11}$$

for $\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \in M^{(2)}$ and $\begin{bmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \end{bmatrix} \in (1 - s_{(M^{(2)})'}(\theta))\mathcal{H}^{(2)}$. Rewrite (A.11) in the 4×4 form and combine with (A.9) and (A.10). Then extending (A.6) we find that

$$S_\theta = \begin{bmatrix} S_\varphi & 0 & 0 & 0 \\ 0 & 0 & S_{\psi,\varphi} & 0 \\ 0 & S_{\varphi,\psi} & 0 & 0 \\ 0 & 0 & 0 & S_\psi \end{bmatrix} \tag{A.12}$$

and so

$$\Delta_\theta = S_\theta^* S_\theta = \begin{bmatrix} S_\varphi^* S_\varphi & 0 & 0 & 0 \\ 0 & S_{\varphi,\psi}^* S_{\varphi,\psi} & 0 & 0 \\ 0 & 0 & S_{\psi,\varphi}^* S_{\psi,\varphi} & 0 \\ 0 & 0 & 0 & S_\psi^* S_\psi \end{bmatrix} = \begin{bmatrix} \Delta_\varphi & 0 & 0 & 0 \\ 0 & \Delta_{\varphi,\psi} & 0 & 0 \\ 0 & 0 & \Delta_{\psi,\varphi} & 0 \\ 0 & 0 & 0 & \Delta_\psi \end{bmatrix}. \tag{A.13}$$

From the first case (applied to $\theta \in (M^{(2)})_*^+$) and (A.7) it follows that

$$S_\theta = J^{(2)} \Delta_\theta^{1/2} = \begin{bmatrix} J\Delta_\varphi^{1/2} & 0 & 0 & 0 \\ 0 & 0 & J\Delta_{\psi,\varphi}^{1/2} & 0 \\ 0 & J\Delta_{\varphi,\psi}^{1/2} & 0 & 0 \\ 0 & 0 & 0 & J\Delta_\psi^{1/2} \end{bmatrix}. \tag{A.14}$$

Comparing the (2,3)-entries of (A.12) and (A.14) implies that $S_{\psi,\varphi} = J\Delta_{\psi,\varphi}^{1/2}$.

(3) For the case $\psi = \varphi$, from the above proof of (2) and the property (4) of Sect. A.2, we have

$$\Delta_\varphi^{-1} = \Delta_{\overline{\varphi}}^{-1} \oplus 0 = (J_{\overline{\varphi}} \Delta_{\overline{\varphi}} J_{\overline{\varphi}}) \oplus 0 = J\Delta_\varphi J.$$

For general ψ, φ, apply the above case to θ in the proof of (2); then by (A.13) and (A.7) we have

$$
\begin{bmatrix}
\Delta_\varphi^{-1} & 0 & 0 & 0 \\
0 & \Delta_{\varphi,\psi}^{-1} & 0 & 0 \\
0 & 0 & \Delta_{\psi,\varphi}^{-1} & 0 \\
0 & 0 & 0 & \Delta_\psi^{-1}
\end{bmatrix}
=
\begin{bmatrix}
J\Delta_\varphi J & 0 & 0 & 0 \\
0 & J\Delta_{\psi,\varphi} J & 0 & 0 \\
0 & 0 & J\Delta_{\varphi,\psi} J & 0 \\
0 & 0 & 0 & J\Delta_\psi J
\end{bmatrix},
$$

implying that $\Delta_{\psi,\varphi}^{-1} = J\Delta_{\varphi,\psi} J$. $\square$

Example A.24 (1) Let $M = L^\infty(X, \mu)$ as in Example A.14 (1). For every $\psi, \varphi \in L^1(X, \mu)_+ \cong M_*^+$, it is easy to verify that $\Delta_{\psi,\varphi}$ is the multiplication of $1_{\{\varphi>0\}}(\psi/\varphi)$, which is the Radon-Nikodym derivative of $\psi \, d\mu$ with respect to $\varphi \, d\mu$ (restricted on the support of φ) in the classical sense.

(2) Let $M = B(\mathcal{H})$ as in Example A.14 (2). For every $\psi, \varphi \in B(\mathcal{H})_*^+$ we have the density operators (positive trace-class operators) D_ψ, D_φ such that $\varphi(X) = \operatorname{Tr} D_\varphi X$ for $X \in B(\mathcal{H})$ and similarly for D_ψ. Let $D_\psi = \sum_{a>0} aP_a$ and $D_\varphi = \sum_{b>0} bQ_b$ be the spectral decompositions of D_ψ, D_φ, where P_a and Q_b are finite-dimensional orthogonal projections. Then the relative modular operators $\Delta_{\psi,\varphi}$ on $C_2(\mathcal{H})$ is given as

$$
\Delta_{\psi,\varphi} = L_{D_\psi} R_{D_\varphi^{-1}} = \sum_{a>0, \, b>0} ab^{-1} L_{P_a} R_{Q_b}, \tag{A.15}
$$

where $L_{[-]}$ and $R_{[-]}$ denote the left and the right multiplications and D_φ^{-1} is the generalized inverse of D_φ.

A.5 τ-measurable operators

The non-commutative integration theory was created by Segal [102], where measurable operators affiliated with a von Neumann algebra were discussed. Later in [86], Nelson introduced the notion of τ-measurable operators in a stricter connection with a given trace τ. This section is a brief survey of the theory of τ-measurable operators. Assume that M is a semifinite von Neumann algebra on a Hilbert space $\mathcal{H}$ and τ is a faithful normal semifinite trace on M.

Let $a : \mathcal{D}(a) \to \mathcal{H}$ be a linear operator, where $\mathcal{D}(a)$ is a linear subspace of $\mathcal{H}$. We say that a is *affiliated with M*, denoted by $a \, \eta M$, if $x'a \subset ax'$ for all $x' \in M'$, or equivalently, if $u'au'^* = a$ for all unitaries $u' \in M'$. The following facts are easy to verify:

(a) If a, b are linear operators affiliated with M, then $a + b$ with $\mathcal{D}(a + b) = \mathcal{D}(a) \cap \mathcal{D}(b)$ and ab with $\mathcal{D}(ab) = \{\xi \in \mathcal{D}(b) : b\xi \in D(a)\}$ are affiliated with M.
(b) If a is densely defined and $a \, \eta M$, then $a^* \, \eta M$.

(c) If a is closable and $a \, \eta \, M$, then $\overline{a} \, \eta \, M$.
(d) Assume that a is densely defined and closed, so we have the polar decomposition $a = w|a|$ and the spectral decomposition $|a| = \int_0^\infty \lambda \, e_\lambda$. Then $a \, \eta \, M$ if and only if $w, e_\lambda \in M$ for all $\lambda \geq 0$.

Definition A.25 Let a be a densely-defined closed operator such that $a \, \eta \, M$. We say that a is τ-*measurable* if, for any $\delta > 0$, there exists an $e \in \mathrm{Proj}(M)$ such that $e\mathcal{H} \subset \mathcal{D}(a)$ and $\tau(e^\perp) \leq \delta$. We denote by $\widetilde{M}$ the set of such τ-measurable operators.

Proposition A.26 *Let a be a densely-defined closed operator affiliated with M with $a = w|a|$ and $|a| = \int_0^\infty \lambda \, de_\lambda$ as above. Then the following conditions are equivalent:*

 (i) $a \in \widetilde{M}$;
 (ii) $|a| \in \widetilde{M}$;
(iii) $\tau(e_\lambda^\perp) \to 0$ *as* $\lambda \to \infty$;
(iv) $\tau(e_\lambda^\perp) < +\infty$ *for some* $\lambda > 0$.

For each $\varepsilon, \delta > 0$ define

$$\mathcal{N}(\varepsilon, \delta) := \{a \in \widetilde{M} : \|ae\| \leq \varepsilon \text{ and } \tau(e^\perp) \leq \delta \text{ for some } e \in \mathrm{Proj}(M)\}.$$

Then the following assertions hold:

(1) If $a \in \widetilde{M}$, then $a^* \in \widetilde{M}$. Moreover, $a \in \mathcal{N}(\varepsilon, \delta) \iff a^* \in \mathcal{N}(\varepsilon, \delta)$.
(2) If $a, b \in \widetilde{M}$, then $a + b$ and ab are densely defined and closable, and $\overline{a + b}, \overline{ab} \in \widetilde{M}$. Moreover, if $a \in \mathcal{N}(\varepsilon_1, \delta_1)$ and $b \in \mathcal{N}(\varepsilon_2, \delta_2)$, then $\overline{a + b} \in \mathcal{N}(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$ and $\overline{ab} \in \mathcal{N}(\varepsilon_1 \varepsilon_2, \delta_1 + \delta_2)$.

In view of the above assertions, for every $a, b \in \widetilde{M}$ we may use the convention that $a + b$ and ab mean the strong sum $\overline{a + b}$ and the strong product $\overline{ab}$, respectively. A big advantage of τ-measurable operators is that we can freely take adjoint, sum and product in $\widetilde{M}$. The following is the main result on τ-measurable operators.

Theorem A.27 $\widetilde{M}$ *is a complete metrizable Hausdorff topological $*$-algebra with the strong sum, the strong product, and $\{\mathcal{N}(\varepsilon, \delta) : \varepsilon, \delta > 0\}$ as a neighborhood basis of 0. Moreover, M is dense in $\widetilde{M}$.*

Let $\widetilde{M}_+$ be the set of positive self-adjoint $a \in \widetilde{M}$ (denoted by $a \geq 0$). Note that $\widetilde{M}_+$ is a closed convex cone in $\widetilde{M}$ and $\widetilde{M}$ is an ordered topological space with the order $a \geq b$ defined by $a - b \geq 0$ for self-adjoint $a, b \in \widetilde{M}$. The trace τ on M_+ extends to all $a \in \widetilde{M}_+$ as $\tau(a) := \int_0^\infty \lambda \, d\tau(e_\lambda)$, where $a = \int_0^\infty \lambda \, de_\lambda$ is the spectral decomposition.

The topology on $\widetilde{M}$ given in Theorem A.27 is called the *measure topology*. This topology is not necessarily locally convex. For instance, if M is a finite non-atomic von Neumann algebra with a faithful normal finite trace τ, then a non-empty open convex set in $\widetilde{M}$ is only the whole $\widetilde{M}$ and there is no non-zero continuous functional on $\widetilde{M}$.

Example A.28 (1) When $M = B(\mathcal{H})$ and τ is the usual trace Tr, $\widetilde{M} = B(\mathcal{H})$ and the measure topology is the operator norm topology.

(2) Let M be finite with a faithful normal finite trace τ. Then $\widetilde{M}$ is the set of all densely-defined closed operators $x \eta M$.

(3) Let $(X, \mathcal{X}, \mu)$ be a localizable measure space, where $(X, \mathcal{X}, \mu)$ is *localizable* if for every $A \in \mathcal{X}$ there is a $B \in \mathcal{X}$ such that $B \subset A$ and $\mu(B) < +\infty$. For an abelian von Neumann algebra $\mathcal{A} = L^\infty(X, \mu) = L^1(X, \mu)^*$ with $\tau(f) := \int_X f \, d\mu$ for $f \in L^\infty(X, \mu)_+$, $\widetilde{\mathcal{A}}$ is the space of measurable functions f on X such that there is an $A \in \mathcal{X}$ such that $\mu(A) < +\infty$ and f is bounded on $X \setminus A$, where $f = g$ in $\widetilde{\mathcal{A}} \iff f(x) = g(x)$ μ-a.e.

We end the section with the following two definitions.

Definition A.29 For $a \in \widetilde{M}$ and $t > 0$ the (tth) *generalized s-number* $\mu_t(a)$ is defined to be

$$\mu_t(a) := \inf\{s \geq 0 : \tau(e_{(s,\infty)}(|a|)) \leq t\},$$

where $e_{(s,\infty)}(|a|)$ denotes the spectral projection of $|a|$ corresponding to the interval (s, ∞). Note that $\mu_t(a) < +\infty$ for all $t > 0$ since $\tau(e_{(s,\infty)}(|a|)) \to 0$ as $s \to \infty$ by Proposition A.26.

Definition A.30 For each $a \in \widetilde{M}$ define

$$\|a\|_p := \tau(|a|^p)^{1/p} \in [0, +\infty], \qquad 0 < p < \infty.$$

Furthermore, $\|a\|_\infty := \|a\| \in [0, +\infty]$. The *non-commutative L^p-space* on (M, τ) is defined as

$$L^p(M, \tau) := \{a \in \widetilde{M} : \|a\|_p < +\infty\}, \qquad 0 < p \leq \infty.$$

(In particular, $L^\infty(M, \tau) = M$.)

The trace τ uniquely extends to a positive linear functional on $L^1(M, \tau)$ and it satisfies $|\tau(a)| \leq \|a\|_1$, $a \in L^1(M, \tau)$. Assume that $1 \leq P < \infty$ and $1/p + 1/q = 1$. If $a \in L^p(M, \tau)$ and $b \in L^q(M, \tau)$, then $ab \in L^1(M, \tau)$ and

$$|\tau(ab)| \leq \|ab\|_1 \leq \|a\|_p \|b\|_q \quad \textit{(Hölder's inequality)}.$$

Also, $L^p(M, \tau)$ is a Banach space with respect to the norm $\|\cdot\|_p$, and its dual Banach space is $L^q(M, \tau)$ under the duality pairing

$$(a, b) \in L^p(M, \tau) \times L^q(M, \tau) \longmapsto \tau(ab) \in \mathbb{C}.$$

In particular, we have $M_* = L^1(M, \tau)$ under the correspondence $\psi \in M_* \leftrightarrow a \in L^1(M, \tau)$ given by $\psi(x) = \tau(ax)$, $x \in M$, and $\psi \in M_*^+ \iff a \in L^1(M, \tau)_+$ $(= L^1(M, \tau) \cap \widetilde{M}+)$. This a is often called the *Radon-Nikodym derivative* of ψ with respect to τ and denoted by $d\psi/d\tau$.

More details on the generalized s-numbers and the non-commutative $L^p(M, \tau)$ are found in [33] that is the best literature on those. More general Haagerup's L^p-spaces will be explained in the next section.

A.6 Haagerup's L^p-Spaces

We begin with the crossed product of M by the modular automorphism group, based on which Haagerup's L^p-spaces are constructed. Let $\sigma = \sigma_t^\omega$ ($t \in \mathbb{R}$) be the modular automorphism group with respect to a faithful $\omega \in M_*^+$ (see Definition A.2). (We may take a faithful normal semifinite weight ω when M is not σ-finite.) Define a faithful normal representation π_σ of M and a one-parameter strongly continuous unitary representation λ on $L^2(\mathbb{R}, \mathcal{H}) \cong \mathcal{H} \otimes L^2(\mathbb{R})$ by

$$(\pi_\sigma(x)\xi)(s) := \sigma_{s^{-1}}(x)\xi(s), \quad (\lambda(t)\xi)(s) := \xi(s - t), \quad s \in \mathbb{R}, \ \xi \in L^2(\mathbb{R}, \mathcal{H}),$$

satisfying the covariance property $\pi_\sigma(\sigma_t(x)) = \lambda(t)\pi_\sigma(x)\lambda(t)^*$ for all $x \in M$, $t \in \mathbb{R}$, Then the *crossed product*

$$N := M \rtimes_\sigma \mathbb{R} = \{\pi_\sigma(M) \cup \lambda(\mathbb{R})\}''$$

is the von Neumann algebra generated by $\pi_\sigma(x)$ ($x \in M$) and $\lambda(t)$ ($t \in \mathbb{R}$). The crossed product construction is indeed performed for a continuous (locally compact abelian) group action on M, for which Takesaki's duality theorem [107] was established, where the dual action plays a role. In the case of σ^ω the *dual action* $\theta = \widehat{\sigma^\omega}$ is determined by $\theta_s(\pi_\sigma(x)) = \pi_\sigma(x)$ and $\theta_s(\lambda(t)) = e^{ist}\lambda(t)$ for $x \in M$, $s, t \in \mathbb{R}$. Then $N^\theta = \pi_\sigma(M)$ holds, where N^θ is the fixed-point algebra of θ. We may assume that $M \subset N$ by identifying $x \in M$ with $\pi_\sigma(x)$.

An important fact is that N is a *semifinite* von Neumann algebra with an faithful normal semifinite trace τ (called the canonical trace on N) satisfying the trace scaling property

$$\tau \circ \theta_s = e^{-s}\tau, \qquad s \in \mathbb{R},$$

which is shown by using the *dual weight* $\widetilde{\omega}$ on N though we omit its details here (see [39, 40] for theory of dual weights). Thus, let $\widetilde{N}$ be the space of τ-measurable operators affiliated with $N = M \rtimes_\sigma \mathbb{R}$. It is easy to see that θ_s ($s \in \mathbb{R}$) on N uniquely extends to $\widetilde{N}$ as a one-parameter group of homeomorphic $*$-isomorphisms with respect to the measure topology. It is also worth noting that the triplet

$$(N := M \rtimes_{\sigma^\omega} \mathbb{R}, \ \theta := \widehat{\sigma^\omega}, \ \tau)$$

is canonical in the sense that for another faithful $\omega_1 \in M_*^+$ there is a unitary U on $L^2(\mathbb{R}, \mathcal{H})$ such that

$$UNU^* = N_1, \quad \text{Ad}(U) \circ \theta_s = \theta_{1s} \circ \text{Ad}(U) \ (s \in \mathbb{R}), \quad \tau = \tau_1 \circ \text{Ad}(U),$$

where (N_1, θ_1, τ_1) is the triplet associated with ω_1. When M is a factor of type III, the *flow of weights* of M mentioned at the end of Sec. A.1 is defined by $(X, F_t^M) := (Z(N), \theta_t|_{Z(N)})$, which is an ergodic flow and can be used for the type III$_\lambda$ ($0 \leq \lambda \leq 1$) classification, see [27].

Definition A.31 ([41]) For $0 < p \leq \infty$ *Haagerup's L^p-space $L^p(M)$ is defined by*

$$L^p(M) := \{x \in \widetilde{N} : \theta_s(x) = e^{-s/p}x, \ s \in \mathbb{R}\}.$$

Let $L^p(M)_+ = L^p(M) \cap \widetilde{N}_+$ where $\widetilde{N}_+$ is the positive part of $\widetilde{N}$.

Clearly, $L^p(M)$'s are (closed) linear subspaces of $\widetilde{N}$, which are M-bimodules and invariant under $a \mapsto a^*$ and $a \mapsto |a|$. Moreover, they are linearly spanned by their positive part $L^p(M)_+$. It follows that every $a \in L^\infty(M)$ is bounded, so we have $L^\infty(M) = M$. The following is the starting point of the theory of Haagerup's L^p-spaces developed in [110].

Lemma A.32 *There is a bijection $\psi \in M_*^+ \mapsto h_\psi \in L^1(M)_+$ satisfying, for every $\psi, \varphi \in M_*^+$ and $x \in M$,*

$$h_{\psi+\varphi} = h_\psi + h_\varphi, \qquad h_{x\psi x^*} = xh_\psi x^*,$$

where $(x\psi x^*)(y) := \psi(x^*yx)$, $y \in M$.

Although we omit the details here, note that $\psi \in M_*^+ \mapsto h_\psi \in L^1(M)_+$ is determined in such a way that the dual weight $\widetilde{\psi}$ on N is equal to $\tau(h_\psi \cdot)$, and that $s(\psi) = s(h_\psi)$ for the support projections. By linearly extending the above bijection we have the next theorem.

Theorem A.33 *There is a linear bijection $\psi \in M_* \mapsto h_\psi \in L^1(M)$ satisfying, for every $\psi \in M_*$ and $x, y \in M$,*

$$h_{x\psi y^*} = xh_\psi y^*, \qquad h_{\psi^*} = h_\psi^*.$$

Moreover, if $\psi = u|\psi|$ is the polar decomposition of $\psi \in M_*$ (see [108, Sec. III.4]), then $h_\psi = uh_{|\psi|}$ is the polar decomposition of h_ψ. Hence

$$|h_\psi| = h_{|\psi|}$$

and the partial isometry part of h_ψ is $u \in M$.

Due to the linear bijection in Theorem A.33, define a linear functional tr on $L^1(M)$ by

$$\mathrm{tr}\,(h_\psi) := \psi(1), \qquad \psi \in M_*.$$

Then

$$\mathrm{tr}\,(|h_\psi|) = \mathrm{tr}\,(h_{|\psi|}) = |\psi|(1) = \|\psi\|, \qquad \psi \in M_*. \tag{A.16}$$

Lemma A.34 *Let $a \in \widetilde{N}$ with the polar decomposition $a = u|a|$. Then for every $p \in [1, \infty)$,*

$$a \in L^p(M) \iff u \in M \text{ and } |a|^p \in L^1(M).$$

Definition A.35 In view of Lemma A.34, for every $a \in L^p(M)$ define $\|a\|_p \in [0, +\infty)$ by

$$\|a\|_p := \operatorname{tr}(|a|^p)^{1/p} \quad \text{if } 0 < p < \infty,$$
$$\|a\|_\infty := \|a\| \qquad \text{if } p = \infty.$$

In the case $p = 1$, by (A.16), $\|a\|_1 := \operatorname{tr}(|a|)$ for $a \in L^1(M)$ is the norm on $L^1(M)$ copied from the norm on M_* by the linear bijection. In this way,

$$(L^1(M), \|\cdot\|_1)$$

becomes a Banach space identified with M_*.

Example A.36 Assume that M is semifinite with a faithful normal semifinite trace τ_0. Then (N, θ_s, τ) is identified with

$$\left(M \otimes L^\infty(\mathbb{R}), \; \operatorname{id} \otimes (f \mapsto f(\cdot + s)), \; \tau_0 \otimes \int_{\mathbb{R}} \cdot \, e^t \, dt \right).$$

In this case, for each $\psi \in M_*^+$, $h_\psi = (d\psi/d\tau) \otimes e^{-t}$ holds, where $d\psi/d\tau$ is the Radon-Nikodym derivative of ψ with respect to τ (see Sec. A.5). Hence, for each $p \in (0, \infty]$,

$$L^p(M) = L^p(M, \tau_0) \otimes e^{-t/p}$$

and $\|a \otimes e^{-t/p}\|_{L^p(M)} = \|a\|_{L^p(M, \tau_0)}$ for all $a \in L^p(M, \tau_0)$. With neglecting the superfluous tensor factor $e^{-t/p}$ we may identify $L^p(M)$ with $L^p(M, \tau_0)$.

In the rest of the section we present properties of Haagerup's L^p-spaces. See [110] for their proofs and more details.

Proposition A.37 *Let $1 \le p, q \le \infty$ with $1/p + 1/q = 1$. Let $a \in L^p(M)$ and $b \in L^q(M)$. Then $ab, ba \in L^1(M)$ and*

$$\operatorname{tr}(ab) = \operatorname{tr}(ba).$$

Theorem A.38 (Hölder's inequality) *Let $1 \le p, q \le \infty$ with $1/p + 1/q = 1$. Then for every $a \in L^p(M)$ and $b \in L^q(M)$,*

$$|\operatorname{tr}(ab)| \le \|ab\|_1 \le \|a\|_p \|b\|_q.$$

Moreover, for every $p, q, r \in (0, \infty]$ with $1/p + 1/q = 1/r$, if $a \in L^p(M)$ and $b \in L^q(M)$, then $ab \in L^r(M)$ and

$$\|ab\|_r \le \|a\|_p \|b\|_q.$$

The reference [110] contains only the first assertion of the above theorem, while the second assertion and other inequalities for Haagerup's $\| \cdot \|_p$ are found in [33].

Proposition A.39 *Let* $1 \le p, q \le \infty$ *with* $1/p + 1/q = 1$. *Then for every* $a \in L^p(M)$,

$$\|a\|_p = \sup\{|\mathrm{tr}\,(ab)| : b \in L^q(M),\ \|b\|_q \le 1\}. \tag{A.17}$$

Theorem A.40 (1) *For every* $p \in [1, \infty]$, $(L^p(M), \| \cdot \|_p)$ *is a Banach space.*
(2) *In particular,* $L^2(M)$ *is a Hilbert space with respect to the inner product* $\langle a, b \rangle :=$ $\mathrm{tr}\,(a^*b)$ $(= \mathrm{tr}\,(ba^*))$ *for* $a, b \in L^2(M)$.
(3) *For any* $p \in [1, \infty)$, *the norm topology on* $L^p(M)$ *coincides with the relative topology induced from* $\widetilde{N}$ *with the measure topology. More precisely, the uniform structure on* $L^p(M)$ *by* $\| \cdot \|_p$ *coincides with that induced from* $\widetilde{N}$.

Theorem A.41 *Let* $1 \le p < \infty$ *and* $1/p + 1/q = 1$. *Then the dual Banach space of* $L^p(M)$ *is* $L^q(M)$ *under the duality pairing*

$$(a, b) \in L^p(M) \times L^q(M) \longmapsto \mathrm{tr}\,(ab) \in \mathbb{C}.$$

Proposition A.42 *Let* $1 \le p, q \le \infty$ *with* $1/p + 1/q = 1$. *Let* $a \in L^q(M)$. *Then*

$$a \ge 0 \iff \mathrm{tr}\,(ab) \ge 0 \ \text{for all } b \in L^p(M)_+.$$

Theorem A.43 *For each* $x \in M$ *we define the left action* $\lambda(x)$ *and the right action* $\rho(x)$ *on the Hilbert space* $L^2(M)$ *by*

$$\lambda(x)a := xa, \qquad \rho(x)a := ax, \qquad a \in L^2(M),$$

and the involution J *on* $L^2(M)$ *by* $Ja := a^*$. *Then:*

(1) λ *(resp.,* ρ*) is a normal faithful representation (resp., anti-representation) of* M *on* $L^2(M)$.
(2) *The von Neumann algebras* $\lambda(M)$ *and* $\rho(M)$ *are the commutants of each other and*

$$\rho(M) = J\lambda(M)J.$$

(3) $(\lambda(M), L^2(M), J, L^2(M)_+)$ *is a standard form of* M.

Remark A.44 For any projection $e \in M$, Haagerup's L^p-space $L^p(eMe)$ is identified with $eL^p(M)e$. Furthermore, since $ej(e)L^2(M) = eL^2(M)e$, we see from Proposition A.16 that

$$(eMe, eL^2(M)e, J = {}^*, eL^2(M)_+e)$$

is a standard form of eMe, where eMe acts on $eL^2(M)e$ as the left multiplication.

Finally in this section we give a lemma, which will be useful in the next section.

Lemma A.45 ([69]) *For every* $\varphi \in M_*^+$ *let* $s(\varphi)$ *be the* M-*support of* φ *and* σ_t^φ *is the modular automorphism group with respect to* $\varphi|_{s(\varphi)Ms(\varphi)}$. *Then*

$$\sigma_t^\varphi(x) = h_\varphi^{it} x h_\varphi^{-it}, \qquad x \in s(\varphi)Ms(\varphi),\ t \in \mathbb{R}.$$

Proof First, assume that φ is faithful. Then $\alpha_t(x) := h_\varphi^{it} x h_\varphi^{-it}$ ($x \in M,\ t \in \mathbb{R}$) defines a strongly continuous one-parameter automorphism group. Let $x, y \in M$ and assume that x is entire α-analytic with the analytic extension $\alpha_z(x)$ ($z \in \mathbb{C}$), see, e.g., [20, Sec. 2.5.3]. By analytic continuation one can see that $h_\varphi^{is} \alpha_z(x) = \alpha_{z+s}(x) h_\varphi^{is}$ for all $s \in \mathbb{R}$ and $z \in \mathbb{C}$, which implies further that $h_\varphi \alpha_z(x) = \alpha_{z-i}(x) h_\varphi$. Since

$$\varphi(\alpha_t(x)y) = \operatorname{tr}(h_\varphi \alpha_t(x)y) = \operatorname{tr}(\alpha_{t-i}(x) h_\varphi y) = \varphi(y \alpha_{t-i}(x)), \qquad t \in \mathbb{R},$$

it follows that φ satisfies the $(\sigma_t^\varphi, -1)$-KMS condition. Hence Theorem A.7 implies that $\sigma_t^\varphi = \alpha_t$. For general $\varphi \in M_*^+$ let $e := s(\varphi)$. Since $h_\varphi \in eL^1(M)e$ corresponds to $\varphi|_{eMe}$ (see Remark A.44), the result follows from the above case. $\qquad\square$

The above lemma says that $\Delta_\varphi^{it}(x h_\varphi^{1/2}) = h_\varphi^{it}(x h_\varphi^{1/2}) h_\varphi^{-it}$ for all $x \in M$ and $t \in \mathbb{R}$. This shows that $\Delta_\varphi^{it}\xi = h_\varphi^{it}\xi h_\varphi^{-it}$ for all $\xi \in L^2(M)$ and $t \in \mathbb{R}$. Also, from the uniqueness of analytic continuation it follows that $\Delta_\varphi^{p/2}(x h_\varphi^{1/2}) = h_\varphi^{p/2} x h_\varphi^{(1-p)/2}$ for $0 \le p \le 1$. Furthermore, in a way similar to the proof of Proposition A.23 (2) we have for any $\psi, \varphi \in M_*^+$,

$$\Delta_{\psi,\varphi}^{it}\xi = h_\psi^{it}\xi h_\varphi^{-it}, \qquad \xi \in L^2(M),\ t \in \mathbb{R}, \tag{A.18}$$

$$\Delta_{\psi,\varphi}^{p/2} x h_\varphi^{1/2} = h_\psi^{p/2} x h_\varphi^{(1-p)/2}, \quad x \in M,\ 0 \le p \le 1, \tag{A.19}$$

with the convention that $h_\psi^0 = s(\psi)$, $h_\varphi^0 = s(\varphi)$ and $\Delta_{\psi,\varphi}^0 = s(\psi)Js(\varphi)J$.

A.7 Connes' Cocycle Derivatives

Consider the tensor product $M^{(2)} := M \otimes \mathbb{M}_2$ of M with the 2×2 matrix algebra $\mathbb{M}_2$. For each $\varphi, \psi \in M_*^+$ the *balanced functional* $\theta = \theta(\varphi, \psi)$ on $M^{(2)}$ is given by

$$\theta\left(\sum_{i,j=1}^2 x_{ij} \otimes e_{ij}\right) := \varphi(x_{11}) + \psi(x_{22}), \qquad x_{ij} \in M,$$

where e_{ij} ($i, j = 1, 2$) are the matrix units of $\mathbb{M}_2$. The functional θ was used in the proof of Proposition A.23 (3). The support projection of θ is $s(\theta) = s(\varphi) \otimes e_{11} + s(\psi) \otimes e_{22}$. Concerning the modular automorphism group σ_t^θ on $s(\theta)M^{(2)}s(\theta)$ the following hold:

(1) $s(\varphi) \otimes e_{11}, s(\psi) \otimes e_{22} \in (s(\theta)M^{(2)}s(\theta))^\theta$ (the centralizer of $\theta|_{s(\theta)M^{(2)}s(\theta)}$).
(2) $\sigma_t^\theta(x \otimes e_{11}) = \sigma_t^\varphi(x) \otimes e_{11}$ for all $t \in \mathbb{R}$ and $x \in s(\varphi)Ms(\varphi)$.
(3) $\sigma_t^\theta(x \otimes e_{22}) = \sigma_t^\psi(x) \otimes e_{11}$ for all $t \in \mathbb{R}$ and $x \in s(\psi)Ms(\psi)$.
(4) $\sigma_t^\theta(s(\psi)Ms(\varphi) \otimes e_{21}) \subset s(\psi)Ms(\varphi) \otimes e_{21}$.

Definition A.46 ([23]) Let $\varphi, \psi \in M_*^+$ and $\theta = \theta(\varphi, \psi)$. By the above (4) there exists a strongly* continuous map $t \in \mathbb{R} \mapsto u_t \in s(\psi)Ms(\varphi)$ such that

$$\sigma_t^\theta(s(\psi)s(\varphi) \otimes e_{21}) = u_t \otimes e_{21}, \qquad t \in \mathbb{R}.$$

The map $t \mapsto u_t$ is called *Connes' cocycle (Radon-Nikodym) derivative* of ψ with respect to φ, and denoted by $u_t = [D\psi : D\varphi]_t, t \in \mathbb{R}$.

The expression of $[D\psi : D\varphi]_t$ in terms of Haagerup's L^1-elements is quite convenient to derive properties of Connes' cocycle derivative $[D\psi : D\varphi]_t$. For this, we first examine Haagerup's L^p-spaces for $M^{(2)} = M \otimes \mathbb{M}_2$. Take the tensor product $\omega \otimes \mathrm{Tr}$ of a faithful $\omega \in M_*^+$ and the trace functional Tr on $\mathbb{M}_2$. Then $\sigma_t^{\omega \otimes \mathrm{Tr}} = \sigma_t^\omega \otimes \mathrm{id}_2$, where id_2 is the identity map on $\mathbb{M}_2$. From the construction of the crossed products $N := M \rtimes_{\sigma^\omega} \mathbb{R}$ and $N^{(2)} := M^{(2)} \rtimes_{\sigma^\omega \otimes \mathrm{id}_2} \mathbb{R}$ (see the first paragraph of Sec. A.6), it is easy to see that

(a) $N^{(2)} = N \otimes \mathbb{M}_2$,
(b) the canonical trace on $N^{(2)}$ is $\tau \otimes \mathrm{Tr}$, where τ is the canonical trace on N,
(c) the dual action on $N^{(2)}$ is $\theta_s \otimes \mathrm{id}_2$ ($s \in \mathbb{R}$), where θ_s is the dual action on N.

Based on these facts we see that $\widetilde{N}^{(2)} = \widetilde{N} \otimes \mathbb{M}_2$ for the spaces $\widetilde{N}$ and $\widetilde{N}^{(2)}$ of τ-measurable and $\tau \otimes \mathrm{Tr}$-measurable operators affiliated with N and $N^{(2)}$, respectively. Therefore, for $0 < p \le \infty$, Haagerup's L^p-space

$$L^p(M^{(2)}) := \{a \in \widetilde{N}^{(2)} = \widetilde{N} \otimes \mathbb{M}_2 : (\theta_s \otimes \mathrm{id}_2)(a) = e^{-s/p}a, \ s \in \mathbb{R}\}$$

is given as

$$L^p(M^{(2)}) = L^p(M) \otimes \mathbb{M}_2 = \left\{a = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} : a_{ij} \in L^p(M), \ i, j = 1, 2\right\},$$

and its positive part is $(L^p(M) \otimes \mathbb{M}_2) \cap (\widetilde{N}^{(2)})_+$. Moreover, by closely looking the construction of the functional tr, we notice that

(d) the tr-functional on $L^1(M^{(2)}) = L^1(M) \otimes \mathbb{M}_2$ is $\mathrm{tr} \otimes \mathrm{Tr}$, where tr is the tr-functional on $L^1(M)$.

In this way, the standard form of $M^{(2)} = M \otimes \mathbb{M}_2$ in terms of Haagerup's L^2-space is canonically given as

$$(M^{(2)}, L^2(M) \otimes \mathbb{M}_2, J = {}^*, (L^2(M) \otimes \mathbb{M}_2)_+),$$

where $[x_{ij}]_{i,j=1}^2 \in M^{(2)}$ acts on $L^2(M) \otimes \mathbb{M}_2$ as the left multiplication as 2×2 matrices $\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{bmatrix}$ and $J = {}^*$ is the matrix $*$-operation $\begin{bmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{bmatrix}^* = \begin{bmatrix} \xi_{11}^* & \xi_{21}^* \\ \xi_{12}^* & \xi_{22}^* \end{bmatrix}$ for $[\xi_{ij}]_{i,j=1}^2 \in L^2(M) \otimes \mathbb{M}_2$.

Lemma A.47 *For every $\varphi, \psi \in M_*^+$,*

$$[D\psi : D\varphi]_t = h_\psi^{it} h_\varphi^{-it}, \qquad t \in \mathbb{R}, \tag{A.20}$$

where h_φ, h_ψ are the elements of Haagerup's L^1-space $L^1(M)$ corresponding to φ, ψ.

Proof In view of the above description of $L^p(M \otimes \mathbb{M}_2)$, in particular, the fact in (d), the element of $L^1(N)$ corresponding to $\theta(\varphi, \psi)$ is $h_\theta = \begin{bmatrix} h_\varphi & 0 \\ 0 & h_\psi \end{bmatrix}$. Hence it follows from Lemma A.45 that for $[x_{ij}] \in s(\theta)Ns(\theta)$,

$$\begin{aligned}
\sigma_t^\theta \left(\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \right) &= \begin{bmatrix} h_\varphi & 0 \\ 0 & h_\psi \end{bmatrix}^{it} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} h_\varphi & 0 \\ 0 & h_\psi \end{bmatrix}^{-it} \\
&= \begin{bmatrix} h_\varphi^{it} x_{11} h_\varphi^{-it} & h_\varphi^{it} x_{12} h_\psi^{-it} \\ h_\psi^{it} x_{21} h_\varphi^{-it} & h_\psi^{it} x_{22} h_\psi^{-it} \end{bmatrix}, \qquad t \in \mathbb{R}. \tag{A.21}
\end{aligned}$$

Therefore,

$$\begin{bmatrix} 0 & 0 \\ [D\psi : D\varphi]_t & 0 \end{bmatrix} = \sigma_t^\theta \left(\begin{bmatrix} 0 & 0 \\ s(\psi)s(\varphi) & 0 \end{bmatrix} \right) = \begin{bmatrix} 0 & 0 \\ h_\psi^{it} h_\varphi^{-it} & 0 \end{bmatrix}$$

so that (A.20) follows. $\square$

Remark A.48 Indeed, the assertions stated just before Definition A.46 are immediately seen from the expression in (A.21) and Lemma A.45.

Example A.49 Assume that M is a semifinite von Neumann algebra with a faithful normal semifinite trace τ. As explained in Example A.36, Haagerup's L^1-space $L^1(M)$ in this case is identified with the conventional L^1-space $L^1(M, \tau)$ with respect to τ. More precisely, for each $\psi \in M_*$, h_ψ in $L^1(M)$ and the Radon-Nikodym derivative $\frac{d\psi}{d\tau} \in L^1(M, \tau)$ are in the relation that $h_\psi = \frac{d\psi}{d\tau} \otimes e^{-t}$. Hence, for every $\varphi, \psi \in M_*^+$ we have

$$[D\psi : D\varphi]_t = \left(\frac{d\psi}{d\tau} \right)^{it} \left(\frac{d\varphi}{d\tau} \right)^{-it}.$$

In particular, when $B = B(\mathcal{H})$ with the usual trace Tr, $[D\psi : D\varphi]_t = D_\psi^{it} D_\varphi^{-it}$ where D_φ, D_ψ are the density (trace-class) operators representing $\varphi, \psi \in B(\mathcal{H})_*^+$.

In the rest of the section, we present important properties of Connes' cocycle derivative $[D\psi : D\varphi]_t$, which were first given by Connes [23, 24] for the case of the faithful normal semifinite weights. Their proofs and more details are found in [103, 109].

Theorem A.50 *Let $\varphi, \psi \in M_*^+$ and assume that $s(\psi) \leq s(\varphi)$. Then $u_t := [D\psi : D\varphi]_t$ satisfies the following properties:*

(i) *$u_t u_t^* = s(\psi) = u_0$ and $u_t^* u_t = \sigma_t^\varphi(s(\psi))$ for all $t \in \mathbb{R}$. In particular, u_t's are partial isometries with the final projection $s(\psi)$.*

(ii) *$u_{s+t} = u_s \sigma_s^\varphi(u_t)$ for all $s, t \in \mathbb{R}$ (the cocycle identity).*

(iii) $u_{-t} = \sigma_{-t}^{\varphi}(u_t^*)$ for all $t \in \mathbb{R}$.

(iv) $\sigma_t^{\psi}(x) = u_t \sigma_t^{\varphi}(x) u_t^*$ for all $t \in \mathbb{R}$ and $x \in s(\psi) M s(\psi)$.

(v) For every $x \in s(\varphi) M s(\psi)$ and $y \in s(\psi) M s(\varphi)$, there exists a continuous bounded function F on $0 \leq \operatorname{Im} z \leq 1$, that is analytic in $0 < \operatorname{Im} z < 1$, such that

$$F(t) = \psi(u_t \sigma_t^{\varphi}(y) x), \qquad F(t+i) = \varphi(x u_t \sigma_t^{\varphi}(y)), \qquad t \in \mathbb{R}.$$

Furthermore, u_t ($t \in \mathbb{R}$) is uniquely determined by a strongly* continuous map $t \in \mathbb{R} \mapsto u_t \in M$ satisfying the above (i), (ii), (iv) and (v). (Note that (iii) follows from (i) and (ii).)

Proposition A.51 Let $\varphi, \psi, \omega \in M_*^+$.

(1) $[D\psi : D\varphi]_t^* = [D\varphi : D\psi]_t$ for all $t \in \mathbb{R}$.

(2) If either $s(\psi) \leq s(\omega)$ or $s(\varphi) \leq s(\omega)$, then $[D\psi : D\omega]_t [D\omega : D\varphi]_t = [D\psi : D\varphi]_t$ for all $t \in \mathbb{R}$ (the chain rule).

(3) If $s(\psi) \leq s(\omega)$ and $s(\varphi) \leq s(\omega)$, then $[D\psi; D\omega]_t = [D\varphi : D\omega]_t$ for all $t \in \mathbb{R}$ if and only if $\psi = \varphi$.

(4) For every $\alpha \in \operatorname{Aut}(M)$, $[D(\psi \circ \alpha) : D(\varphi \circ \alpha)]_t = \alpha^{-1}([D\psi : D\varphi]_t)$ for all $t \in \mathbb{R}$.

Proposition A.52 Let M_0 be a von Neumann subalgebra of M and $E : M \to M_0$ be a faithful normal conditional expectation onto M_0. Then for every faithful $\varphi, \psi \in (M_0)_*^+$,

$$[D(\psi \circ E) : D(\varphi \circ E)]_t = [D\psi : D\varphi]_t, \qquad t \in \mathbb{R}.$$

The next proposition can be shown by using (A.13).

Proposition A.53 For every $\varphi, \psi \in M_*^+$,

$$[D\psi : D\varphi]_t J s(\varphi) J = \Delta_{\psi, \varphi}^{it} \Delta_{\varphi}^{-it}, \quad [D\psi : D\varphi]_t J s(\psi) J = \Delta_{\psi}^{it} \Delta_{\varphi, \psi}^{-it}, \quad t \in \mathbb{R}.$$
$$(A.22)$$

Remark A.54 Since $x \in s(\varphi) M s(\varphi) \mapsto x J s(\varphi) J \in s(\varphi) M s(\varphi) J s(\varphi) J$ is an isomorphism, from (A.22) we may write

$$[D\psi : D\varphi]_t = \Delta_{\psi, \varphi}^{it} \Delta_{\varphi}^{-it} \qquad \text{if } s(\psi) \leq s(\varphi),$$

$$[D\psi : D\varphi]_t = \Delta_{\psi}^{it} \Delta_{\varphi, \psi}^{-it} \qquad \text{if } s(\varphi) \leq s(\psi),$$

Proposition A.55 Let $\varphi, \psi \in M_*^+$ with $s(\psi) \leq s(\varphi)$. The following conditions are equivalent:

(i) $\psi \circ \sigma_t^{\varphi} = \psi$ for all $t \in \mathbb{R}$;

(ii) $[D\psi : D\varphi]_t \in (s(\varphi) M s(\varphi))^{\varphi}$ (the centralizer of $\varphi|_{s(\varphi) M s(\varphi)}$) for all $t \in \mathbb{R}$;

(iii) $[D\psi : D\varphi]_t \in (s(\psi) M s(\psi))^{\psi}$ for all $t \in \mathbb{R}$;

(iv) $t \in \mathbb{R} \mapsto [D\psi : D\varphi]_t$ is a one-parameter group of unitaries in $s(\psi) M s(\psi)$;

(v) $h_{\psi}^{is} h_{\varphi}^{it} = h_{\varphi}^{it} h_{\psi}^{is}$ for all $s, t \in \mathbb{R}$;

(vi) $h_{\psi} h_{\varphi} = h_{\varphi} h_{\psi}$ as elements of $\widetilde{N}$ (the τ-measurable operators affiliated with N).

Definition A.56 We say that ψ *commutes* with φ if the equivalent conditions of Proposition A.55 hold. Condition (i) is often used to define the commutativity for normal functionals (also normal semifinite weights), but (v) and (vi) are quite natural definitions of commutativity that are available for any $\varphi, \psi \in M_*^+$ without $s(\psi) \leq s(\varphi)$, as stated in Lemma 4.19.

The following is *Connes' inverse theorem* in [24, Theorem 1.2.4].

Theorem A.57 ([24]) *Let φ is a faithful normal semifinite weight on M. Let $t \in \mathbb{R} \mapsto u_t \in M$ is a strongly* continuous map satisfying*

$$u_{s+t} = u_s \sigma_s^\varphi(u_t), \qquad t \in \mathbb{R},$$
$$u_{-t} = \sigma_{-t}^\varphi(u_t^*), \qquad t \in \mathbb{R}. \tag{A.23}$$

Then there exists a unique normal semifinite weight ψ on M such that $u_t = [D\psi : D\varphi]_t$ for all $t \in \mathbb{R}$.

In [24] u_t's are assumed to be unitaries in M. In this case, (A.23) is redundant and ψ is faithful as well. The above version without u_t's being unitaries is taken from [103, Theorem 51]. Note that even when $\varphi \in M_*^+$, ψ in the theorem cannot be in M_*^+ in general. This fact suggests us that the von Neumann algebra theory cannot be self-completed when we stick to functionals in M_*, so the normal semifinite weight theory is unavoidable.

For use in Secs. 3.2 and 5.3 we state the following lemmas. The first lemma was also given in [48, Lemma A.1], which generalizes [23, Theorem 3] and [25, Lemma 3.13]. In fact, the lemma can be shown by (A.20) and [104, Sec. 9.24] together with a usual argument in analytic function theory. See also [72]. In the notation of [27, Definition 4.1] condition (ii) can be written as $\psi \leq \lambda^{1/\delta} \varphi \ (\delta/2)$.

Lemma A.58 *For every $\psi, \varphi \in M_*^+$ and $\delta > 0$, the following conditions are equivalent:*

(i) $h_\psi^\delta \leq \lambda h_\varphi^\delta$, *i.e.,* $\lambda h_\varphi^\delta - h_\psi^\delta \in L^\delta(M)_+$ *for some $\lambda > 0$;*
(ii) $s(\psi) \leq s(\varphi)$ *and $[D\psi : D\varphi]_t$ extends to a σ-weakly continuous (M-valued) function $[D\psi : D\varphi]_z$ on $-\delta/2 \leq \operatorname{Im} z \leq 0$ which is analytic in the interior.*

If the above conditions hold, then $\|[D\psi : D\varphi]_z\| \leq \lambda^{1/2}$ and

$$h_\psi^{p/2} = [D\psi : D\varphi]_{-ip/2} h_\varphi^{p/2}, \qquad 0 < p \leq \delta. \tag{A.24}$$

The $\delta = 1$ case of the lemma is Lemma A.21, and so A in Lemma A.21 is equal to $[D\psi : D\varphi]_{-i/2}$.

Kosaki [72] proved the following which contains Sakai's Radon-Nikodym theorem [101, 1.24.3] as the special case when $\delta = p = 1$.

Lemma A.59 *If $\psi, \varphi \in M_*^+$ and $h_\psi^\delta \leq \lambda h_\varphi^\delta$ for some $\delta, \lambda > 0$, then for every $p > 0$ there exists a unique $k_p \in M_+$ with $0 \leq k_p \leq \lambda^{p/2\delta}$ such that $s(k_p) \leq s(\varphi)$ and $h_\psi^p = k_p h_\varphi^p k_p$, where $s(k_p)$ is the support projection of k_p. Moreover, $h_\psi^p = k_p h_\varphi^p k_p$ implies that $(h_\varphi^{p/2} h_\psi^p h_\varphi^{p/2})^{1/2} = h_\varphi^{p/2} k_p h_\varphi^{p/2}$.*

The construction of k_p in [72] is as follows. By virtue of Furuta's inequality [36] (for τ-measurable operators) the above assumption implies that $(h_\varphi^{p/2} h_\psi^p h_\varphi^{p/2})^{1/2} \le \lambda^{p/2\delta} h_\varphi^p$ for any $p > 0$. Hence there exists a unique $b_p \in M$ with $\|b_p\| \le \lambda^{p/4\delta}$ such that $s(|b_p|) \le s(\varphi)$ and $(h_\varphi^{p/2} h_\psi^p h_\varphi^{p/2})^{1/4} = b_p h_\varphi^{p/2}$. In fact, $b_p = [D\chi(p) : D\varphi]_{-ip/2}$ where $\chi(p) \in M_*^+$ is given by $h_{\chi(p)} = (h_\varphi^{p/2} h_\psi^p h_\varphi^{p/2})^{1/2p}$ (see (A.24)). Let $k_p := b_p^* b_p$. Then $h_\varphi^{p/2} h_\psi^p h_\varphi^{p/2} = h_\varphi^{p/2} k_p h_\varphi^p k_p h_\varphi^{p/2}$ and so $h_\psi^p = k_p h_\varphi^p k_p$ and $(h_\varphi^{p/2} h_\psi^p h_\varphi^{p/2})^{1/2} = h_\varphi^{p/2} k_p h_\varphi^{p/2}$.

A.8 Kosaki's L^p-Spaces

In this section, in connection with Haagerup's L^p-spaces in Sec. A.6, we give a brief survey on Kosaki's L^p-spaces [69].

Let $\varphi_0 \in M_*^+$ be a faithful state and $h_0 := h_{\varphi_0} \in L^1(M)$. For each $\theta \in [0, 1]$, M is imbedded into $L^1(M)$ by $x \in M \mapsto h_0^\theta x h_0^{1-\theta} \in L^1(M)$. Define the norm $\|h_0^\theta x h_0^{1-\theta}\| := \|x\|$ on $h_0^\theta M h_0^{1-\theta} (\subset L^1(M))$, so that $h_0^\theta M h_0^{1-\theta} \cong M$. Then $(h_0^\theta M h_0^{1-\theta}, L^1(M))$ becomes a pair of compatible Banach spaces.

Definition A.60 Let $1 < p < \infty$ and $0 \le \theta \le 1$. *Kosaki's L^p-space* $L^p(M; \varphi_0)_\theta$ with respect to φ_0 is defined to be the complex interpolation space

$$C_{1/p}(h_0^\theta M h_0^{1-\theta}, L^1(M))$$

equipped with the complex interpolation norm $\| \cdot \|_{p,\theta} (= \| \cdot \|_{C_{1/p}})$, see [15] for general theory on the complex interpolation method. In particular, $L^p(M; \varphi_0)_\theta$ for $\theta = 0, 1$ and $1/2$ are respectively called the *left*, the *right* and the *symmetric L^p-spaces*. We may write the norm $\| \cdot \|_{p,\theta}$ as $\| \cdot \|_{p,\varphi_0,\theta}$ to specify its dependence on φ_0.

The following is a main theorem [69, Theorem 9.1].

Theorem A.61 *Let $1 \le p \le \infty$ and $1/p + 1/q = 1$. Then*

$$L^p(M; \varphi_0)_\theta = h_0^{\theta/q} L^p(M) h_0^{(1-\theta)/q} \ (\subset L^1(M)),$$

$$\|h_0^{\theta/q} a h_0^{(1-\theta)/q}\|_{p,\theta} = \|a\|_p, \qquad a \in L^p(M),$$

That is, $L^p(M) \cong L^p(M; \varphi_0)_\theta$ by the isometry $a \mapsto h_0^{\theta/q} a h_0^{(1-\theta)/q}$.

Note that a more general embedding $x \in M \mapsto h_0^\theta x k_0^{1-\theta} \in L^1(M)$ was treated in [69], where $k_0 := h_{\psi_0}$ for another faithful state $\psi_0 \in M_*^+$.

Furthermore, when $1 < p < p' < \infty$,

$$L^1(M) = L^1(M; \varphi_0)_\theta \supset L^p(M; \varphi_0)_\theta \supset L^{p'}(M; \varphi_0)_\theta \supset h_0^\theta M h_0^{1-\theta} \cong M,$$

$$\|h_0^\theta x h_0^{1-\theta}\|_1 \le \|h_0^\theta x h_0^{1-\theta}\|_{p,\theta} \le \|h_0^\theta x h_0^{1-\theta}\|_{p',\theta}$$

$$\le \|h_0^\theta x h_0^{1-\theta}\|_{\infty,\theta} = \|x\|, \qquad x \in M.$$

Note that $h_0^\theta M h_0^{1-\theta}$ is dense in $L^p(M;\varphi_0)_\theta$ for every $p \in [1,\infty)$. When $1 \le p < \infty$ and $1/p + 1/q = 1$, $L^q(M;\varphi_0)_{\theta'}$ ($\cong L^q(M)$) is isometric to the dual Banach space of $L^p(M;\varphi_0)_\theta$ ($\cong L^p(M)$) for any $\theta, \theta' \in [0,1]$. But when $\theta' = 1 - \theta$, the duality between $L^p(M;\varphi_0)_\theta$ and $L^q(M;\varphi_0)_{1-\theta}$ is explicitly written as

$$\langle h_0^{\theta/q} a h_0^{(1-\theta)/q}, h_0^{(1-\theta)/p} b h_0^{\theta/p} \rangle_{p,q} = \mathrm{tr}\,(ab), \quad a \in L^p(M), b \in L^q(M).$$

This duality is convenient in the sense that for every $x, y \in M$,

$$\langle h_0^\theta x h_0^{1-\theta}, h_0^{1-\theta} y h_0^\theta \rangle_{p,q} = \mathrm{tr}\,(h_0^\theta x h_0^{1-\theta} y)$$

independently of $p \in [1,\infty)$.

When $\varphi \in M_*^+$ is not necessarily faithful with the M-support $e := s(\varphi)$, Kosaki's L^p-space with respect to φ is defined as $L^p(eMe;\varphi|_{eMe})_\theta$ with restriction to eMe, where the imbedding of eMe into $L^1(eMe)$ is realized as $x \in eMe \mapsto h_\varphi^\theta x h_\varphi^{1-\theta} \in eL^1(M)e$ from Remark A.44. Hence by Theorem A.61 we have

$$L^p(M;\varphi)_\theta = h_\varphi^{\theta/q} e L^p(M) e h_\varphi^{(1-\theta)/q} = h_\varphi^{\theta/q} L^p(M) h_\varphi^{(1-\theta)/q} \quad (\subset eL^1(M)e)$$

for $1 \le p \le \infty$, $1/p + 1/q = 1$ and $0 \le \theta \le 1$. In particular, for the symmetric case $\theta = 1/2$, we simply write

$$L^p(M;\varphi) = L^p(M;\varphi)_{\theta=1/2} = h_\varphi^{1/2q} L^p(M) h_\varphi^{1/2q}.$$

Appendix B
Preliminaries on Positive Self-adjoint Operators

Let A be a positive self-adjoint operator on a Hilbert space $\mathcal{H}$ with the spectral decomposition

$$A = \int_0^\infty t \, dE_t$$

with the spectral resolution E_t, $t \geq 0$. It is well-known that $\xi \in \mathcal{H}$ belongs to $\mathcal{D}(A^{1/2})$, the domain of $A^{1/2}$, if and only if $\int_0^\infty t \, d\|E_t \xi\|^2 < +\infty$, and in this case,

$$\|A^{1/2}\xi\|^2 = \int_0^\infty t \, d\|E_t \xi\|^2.$$

Below we will sometimes use a slightly abuse notation $\|A^{1/2}\xi\|^2$ for all $\xi \in \mathcal{H}$ with the convention that

$$\|A^{1/2}\xi\|^2 := \int_0^\infty t \, d\|E_t \xi\|^2 = \begin{cases} \|A^{1/2}\xi\|^2 < +\infty & \text{if } \xi \in \mathcal{D}(A^{1/2}), \\ +\infty & \text{otherwise.} \end{cases}$$

The following fact is well-known (see [103, Sec. A.4]).

Proposition B.1 *Let A, B be positive self-adjoint operators on $\mathcal{H}$. Then the following conditions are equivalent:*

(i) *$\mathcal{D}(B^{1/2}) \subset \mathcal{D}(A^{1/2})$ and $\|A^{1/2}\xi\|^2 \leq \|B^{1/2}\xi\|^2$ for all $\xi \in \mathcal{D}(B^{1/2})$;*
(ii) *there exists a core $\mathcal{D}$ of $B^{1/2}$ such that $\mathcal{D} \subset \mathcal{D}(A^{1/2})$ and $\|A^{1/2}\xi\|^2 \leq \|B^{1/2}\xi\|^2$ for all $\xi \in \mathcal{D}$;*
(iii) *$(s1 + B)^{-1} \leq (s1 + A)^{-1}$ for some (equivalently, any) $s > 0$;*
(iv) *$A(1 + sA)^{-1} \leq B(1 + sB)^{-1}$ for some (equivalently, any) $s > 0$.*

For positive self-adjoint operators A, B we write $A \leq B$ if the equivalent conditions in Proposition B.1 hold. From Proposition B.1 we have the next result.

Proposition B.2 *For densely-defined closed operators A, B on $\mathcal{H}$ the following are equivalent:*

(i) $A^*A \leq B^*B$;

(ii) *there exists a core $\mathcal{D}$ of B such that $\mathcal{D} \subset \mathcal{D}(A)$ and $\|A\xi\| \leq \|B\xi\|$ for all $\xi \in \mathcal{D}$.*

Proposition B.3 *Let A, B be a positive self-adjoint operators with $s(A) = s(B)$, where $s(A)$ is the support projection of A. If $A \leq B$, then $A^{-1} \geq B^{-1}$, where A^{-1} is defined with restriction to $s(A)\mathcal{H}$ (i.e., in the sense of the generalized inverse).*

Proof Let $e := s(A) = s(B)$. By Proposition B.1 it is clear that $A \leq B$ if and only if $(e + A)^{-1} \geq (e + B)^{-1}$, which is equivalent to $(e + A^{-1})^{-1} \leq (e + B^{-1})^{-1}$ since $(e + A^{-1})^{-1} = e - (e + A)^{-1}$. Hence the assertion holds.   □

The following holds concerning the relative modular operator $\Delta_{\psi,\varphi}$ introduced in Sec. A.4.

Proposition B.4 *Let $\psi, \varphi, \psi_i, \varphi_i \in M_*^+$ $(i = 1, 2)$. If $\psi_1 \leq \psi_2$, then $\Delta_{\psi_1,\varphi} \leq \Delta_{\psi_2,\varphi}$. If $s(\varphi_1) = s(\varphi_2)$ and $\varphi_1 \geq \varphi_2$, then $\Delta_{\psi,\varphi_1} \leq \Delta_{\psi,\varphi_2}$.*

Proof With the vector representatives Ψ_1, Ψ_2, Φ (in the natural positive cone) of ψ_1, ψ_2, φ, respectively, if $\psi_1 \leq \psi_2$, then for every $x \in M$ and $\zeta \in (1 - s_{M'}(\varphi))\mathcal{H}$ one has

$$\|\Delta_{\psi_1,\varphi}^{1/2}(x\Phi + \zeta)\|^2 = \|S_{\psi_1,\varphi}(x\Phi + \zeta)\|^2 = \|x^*\Psi_1\|^2$$
$$= \psi_1(xx^*) \leq \psi_2(xx^*) = \|\Delta_{\psi_2,\varphi}^{1/2}(x\Phi + \zeta)\|^2.$$

Since $M\Phi + (1 - s_{M'}(\varphi))\mathcal{H}$ is a core of $\Delta_{\psi_i,\varphi}^{1/2}$, it follows from Proposition B.1 that $\Delta_{\psi_1,\varphi} \leq \Delta_{\psi_2,\varphi}$. Hence the first assertion holds. For the second, assume that $s(\varphi_1) = s(\varphi_2)$ and $\varphi_1 \geq \varphi_2$. Then $s(\Delta_{\varphi_1,\psi}) = s(\Delta_{\varphi_2,\psi})$ by Proposition A.23 (1). Hence from the first assertion and Proposition B.3 one has $\Delta_{\varphi_1,\psi}^{-1} \leq \Delta_{\varphi_2,\psi}^{-1}$. Since $\Delta_{\psi,\varphi_i} = J\Delta_{\varphi_i,\psi}^{-1}J$ by Proposition A.23 (3), the second holds.   □

Lemma B.5 *Let A, B be a positive self-adjoint operators on $\mathcal{H}$ such that $A \leq B$. Then, for every operator monotone function $f \geq 0$ on $[0, \infty)$ we have $f(A) \leq f(B)$, that is,*

$$\|f(A)^{1/2}\xi\| \leq \|f(B)^{1/2}\xi\|$$

for all $\xi \in \mathcal{H}$ (with the convention mentioned above Proposition B.1).

Proof We write $f(A) = \int_0^\infty f(t)\, dE_t$ with the spectral decomposition $A = \int_0^\infty t\, dE_t$. Note that $\|f(A)^{1/2}\xi\|^2 = \int_0^\infty f(t)\, d\|E_t\xi\|^2$ for every $\xi \in \mathcal{H}$. Now we recall that f has the integral expression

$$f(t) = a + bt + \int_{[0,\infty)} \frac{t}{t + s}\, d\mu(s), \qquad t \in [0, \infty),$$

where $a, b \geq 0$ and μ is a positive measure on $[0, \infty)$ with $\int_{[0,\infty)} (1+s)^{-1}\, d\mu(s) < +\infty$. By Fubini's theorem we have

$$\int_0^\infty f(t)\, d\|E_t\xi\|^2$$

$$= a\|\xi\|^2 + b\|A^{1/2}\xi\|^2 + \int_{[0,\infty)}\left[\int_0^\infty\left(1 - \frac{s}{t+s}\right)d\|E_t\xi\|^2\right]d\mu(s)$$

$$= a\|\xi\|^2 + b\|A^{1/2}\xi\|^2 + \int_{[0,\infty)}[\|\xi\|^2 - s\langle\xi, (s1 + A)^{-1}\xi\rangle]\, d\mu(s). \qquad \text{(B.1)}$$

The same argument as above can be applied to B as well. Condition (iii) of Proposition B.1 implies that for every $\xi \in \mathcal{H}$,

$$\int_{[0,\infty)}[\|\xi\|^2 - s\langle\xi, (sI + A)^{-1}\xi\rangle]\, d\mu(s) \le \int_{[0,\infty)}[\|\xi\|^2 - s\langle\xi, (sI + B)^{-1}\xi\rangle]\, d\mu(s)$$

as well as $\|A^{1/2}\xi\|^2 \le \|B^{1/2}\xi\|^2$. Therefore, we see that $\|f(A)^{1/2}\|^2 \le \|f(B)^{1/2}\|^2$ for all $\xi \in \mathcal{H}$. (In particular, this implies that $\mathcal{D}(f(B)^{1/2}) \subset \mathcal{D}(f(A)^{1/2})$.) $\qquad\square$

The following fact on convergence of positive self-adjoint operators is also well-known (see [99, Sec. VIII.7], [103, Sec. A.3]).

Proposition B.6 *Let A_n ($n \in \mathbb{N}$) and A be positive self-adjoint operators on $\mathcal{H}$. Then the following conditions are equivalent:*

(i) $(s1 + A_n)^{-1} \to (s1 + A)^{-1}$ *strongly for some (equivalently, any) $s > 0$;*
(ii) $(i + A_n)^{-1} \to (i + A)^{-1}$ *strongly;*
(iii) $A_n(1 + sA_n)^{-1} \to A(1 + sA)^{-1}$ *strongly for some (equivalently, any) $s > 0$.*

If the equivalent conditions in Proposition B.6 hold, then we say that A_n converges to A *in the strong resolvent sense* or just *strongly*.

Lemma B.7 *Let A_n ($n \in \mathbb{N}$) and A be positive self-adjoint operators on $\mathcal{H}$ such that $A_1 \le A_2 \le \cdots \le A$ and*

$$(1 + A_n^{1/2})^{-1} \longrightarrow (1 + A^{1/2})^{-1} \quad \text{strongly}. \qquad \text{(B.2)}$$

Then, for every $r \in (0, 1)$ and every $\xi \in \mathcal{H}$ we have

$$\lim_{n\to\infty}\|A_n^{r/2}\xi\|^2 = \|A^{r/2}\xi\|^2 \quad \text{increasingly}.$$

Proof Since $A_1 \le A_2 \le \cdots \le A$, we see by Lemma B.5 that for every $r \in (0, 1)$ and every $\xi \in \mathcal{H}$,

$$\|A_1^{r/2}\xi\| \le \|A_2^{r/2}\xi\| \le \cdots \le \|A^{r/2}\xi\|. \qquad \text{(B.3)}$$

Hence it suffices to prove that

$$\|A^{r/2}\xi\|^2 \le \liminf_{n\to\infty}\|A_n^{r/2}\xi\|^2. \qquad \text{(B.4)}$$

Below we prove the result in the cases $0 < r < 1/2$, $r = 1/2$, and $1/2 < r < 1$, separately. Let $dE_t^{(n)}$, $t \geq 0$, be the spectral resolution for A_n together with dE_t for A.

Case $0 < r < 1/2$. We have the well-known integral expression

$$t^{2r} = \int_0^\infty \frac{t}{t+s} \, d\mu_r(s), \qquad t > 0,$$

with a positive measure $d\mu_r(s) := \frac{\sin(2r\pi)}{\pi} s^{2r-1} \, ds$ on $(0, \infty)$. Hence

$$t^r = \int_0^\infty \frac{t^{1/2}}{t^{1/2} + \lambda} \, d\mu_r(\lambda), \qquad t > 0. \tag{B.5}$$

For every $\xi \in \mathcal{H}$, similarly to (B.1) we can write

$$\|A_n^{r/2}\xi\|^2 = \int_0^\infty t^r \, d\|E_t^{(n)}\xi\|^2 = \int_0^\infty [\|\xi\|^2 - s\langle\xi, (s1 + A_n^{1/2})^{-1}\xi\rangle] \, d\mu_r(s),$$

$$\|A^{r/2}\xi\|^2 = \int_0^\infty t^r \, d\|E_t\xi\|^2 = \int_0^\infty [\|\xi\|^2 - s\langle\xi, (s1 + A^{1/2})^{-1}\xi\rangle] \, d\mu_r(s).$$

Hence, (B.4) follows from (B.2) (with Proposition B.6) and Fatou's lemma.

Case $r = 1/2$. Let $\xi \in \mathcal{H}$ and $\varepsilon > 0$ be arbitrary. Since $t^r \to t^{1/2}$ uniformly on $[0, 1]$ as $r \nearrow 1/2$, choose a $\delta \in (0, 1/2)$ such that if $1/2 - \delta < r < 1/2$ then $t^{1/2} \leq t^r \leq t^{1/2} + \varepsilon$ for all $t \in [0, 1]$. Applying the monotone convergence theorem to $\int_1^\infty t^r \, d\|E_t\xi\|^2$ as $r \nearrow 1/2$, we have

$$\|A^{1/4}\xi\|^2 = \int_0^\infty t^{1/2} \, d\|E_t\xi\|^2 = \lim_{r \nearrow 1/2} \int_0^\infty t^r \, d\|E_t\xi\|^2 = \lim_{r \nearrow 1/2} \|A^{r/2}\xi\|^2$$

so that we can choose an $r \in (1/2 - \delta, 1/2)$ such that $\|A^{1/4}\xi\|^2 \leq \|A^{r/2}\xi\|^2 + \varepsilon$. Since the first case implies that $\|A_n^{r/2}\xi\|^2 \to \|A^{r/2}\xi\|^2$, we can choose an n such that $\|A^{r/2}\xi\|^2 \leq \|A_n^{r/2}\xi\|^2 + \varepsilon$. Therefore,

$$\|A^{1/4}\xi\|^2 \leq \|A_n^{r/2}\xi\|^2 + 2\varepsilon$$

$$\leq \int_{[0,1]} (t^{1/2} + \varepsilon) \, d\|E_t^{(n)}\xi\|^2 + \int_{(1,\infty)} t^{1/2} \, d\|E_t^{(n)}\xi\|^2 + 2\varepsilon$$

$$\leq \|A_n^{1/4}\xi\|^2 + 3\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, it follows from (B.3) that $\|A_n^{1/4}\xi\|^2 \to \|A^{1/4}\xi\|^2$. In the above we have assumed that $\|A^{1/4}\xi\|^2 < +\infty$, but the proof can easily be modified when $\|A^{1/4}\xi\|^2 = +\infty$.

Case $1/2 < r < 1$. By multiplying $t^{1/2}$ to the both sides of (B.5) for $r - \frac{1}{2}$ in place of r, we have

$$t^r = \int_0^\infty \frac{t}{t^{1/2} + s} \, d\mu_r(s), \qquad t > 0,$$

where we have replaced $\mu_{r-\frac{1}{2}}$ in (B.5) with μ_r for notational simplicity. Since

$$\frac{t}{t^{1/2} + s} = t^{1/2} - s + \frac{s^2}{t^{1/2} + s},$$

we can write similarly to (B.1), for every $\xi \in \mathcal{H}$,

$$\|A_n^{r/2}\xi\|^2 = \int_0^\infty [\|A_n^{1/4}\xi\|^2 - s\|\xi\|^2 + s^2 \langle \xi, (s1 + A_n^{1/2})^{-1}\xi\rangle]\, d\mu_r(s),$$

$$\|A^{r/2}\xi\|^2 = \int_0^\infty [\|A^{1/4}\xi\|^2 - s\|\xi\|^2 + s^2 \langle \xi, (s1 + A^{1/2})^{-1}\xi\rangle]\, d\mu_r(s)$$

Again by Fatou's lemma, (B.4) follows from (B.2) and the above case $r = 1/2$. $\square$

Appendix C
Operator Convex Functions on $(0, 1)$

As is well-known, e.g., [46, Theorem 2.7.6], a real function h on $(-1, 1)$ is operator convex if and only if it has the integral expression

$$h(t) = a + bt + \int_{[-1,1]} \frac{t^2}{1 - ut} \, d\lambda(u), \qquad t \in (-1, 1), \qquad \text{(C.1)}$$

where $a, b \in \mathbb{R}$ and a finite positive measure λ on $[-1, 1]$ are uniquely determined.

On the other hand, a real function f on $(0, \infty)$ is operator convex if and only if it has the integral expression (2.5) in Sect. 2.2, where $a, b \in \mathbb{R}$, $c \geq 0$ and a positive measure μ on $[0, \infty)$ with $\int_{[0,\infty)} (1 + s)^{-1} \, d\mu(s) < +\infty$ are uniquely determined.

In Theorem 4.8 the function $g(t) := (1 - t) f\left(\frac{t}{1-t}\right)$ on $(0, 1)$ induced from an operator convex function f on $(0, \infty)$ plays a role. In this appendix we prove the next theorem, which says that the correspondence $f \leftrightarrow g$ is bijective between the operator convex functions on $(0, \infty)$ and those on $(0, 1)$. The integral expression in (C.2) is used in the proof of Theorem 8.4.

Theorem C.1 *Let f be a real function on $(0, \infty)$ and let $g(t) := (1 - t) f\left(\frac{t}{1-t}\right)$ for $t \in (0, 1)$. Then the following conditions are equivalent:*

(a) *f is operator convex on $(0, \infty)$;*
(b) *g is operator convex on $(0, 1)$;*
(c) *There exist $\alpha, \beta \in \mathbb{R}$, $\gamma \geq 0$ and a positive measure ν on $(-\infty, 1]$ satisfying $\int_{(-\infty, 1]} (2 - w)^{-1} \, d\nu(w) < +\infty$ such that*

$$g(t) = \alpha + \beta t + \gamma \frac{(2t - 1)^2}{t} + \int_{(-\infty, 1]} \frac{(2t - 1)^2}{1 - wt} \, d\nu(w), \quad t \in (0, 1), \quad \text{(C.2)}$$

Moreover, in this case,

$$g(0^+) = f(0^+), \qquad g(1^-) = f'(\infty), \qquad \text{(C.3)}$$

and α, β, γ and the measure ν in (c) are uniquely determined by f (hence by g).

Proof (b) $\Longleftrightarrow$ (c). Obviously, g is operator convex on $(0, 1)$ if and only if so is $g\left(\frac{t+1}{2}\right)$ on $(-1, 1)$. So we may show that the expression (C.1) of $h(t) = g\left(\frac{t+1}{2}\right)$ is transformed into (C.2) of g and vice versa. From (C.1) we write for every $t \in (0, 1)$,

$$g(t) = a + b(2t - 1) + \int_{[-1,1]} \frac{(2t - 1)^2}{1 - u(2t - 1)}\, d\lambda(u)$$

$$= (a - b) + 2bt + \frac{\lambda(\{-1\})}{2} \frac{(2t - 1)^2}{t} + \int_{(-1,1]} \frac{(2t - 1)^2}{1 - \frac{2u}{1+u}t} \frac{d\lambda(u)}{1 + u}.$$

By transformation $w = \frac{2u}{1+u} : (-1, 1] \rightarrow (-\infty, 1]$, define a measure ν on $(-\infty, 1]$ by $d\nu(w) = (1 + u)^{-1}\, d\lambda(u)$. Since $\frac{1}{1+u} = \frac{2-w}{2}$, $\int_{(-1,1]} d\lambda(u) < +\infty$ means that $\int_{(-\infty,1]}(2-w)^{-1}\, d\nu(w) < +\infty$. Hence we have the expression (C.2) of g. The converse transformation from (C.2) into (C.1) can similarly be shown, so (b) $\Longleftrightarrow$ (c) follows.

(a) $\Longleftrightarrow$ (b). For this we may show that the expression (2.5) of f is transformed to (C.2) of g and vice versa. From (2.5) we write for every $t \in (0, 1)$,

$$g(t) = a(1 - t) + b(2t - 1) + c\frac{(2t - 1)^2}{1 - t} + \int_{[0,\infty)} \frac{(2t - 1)^2}{t + s(1 - t)}\, d\mu(s)$$

$$= (a - b) + (2b - a)t + c\frac{(2t - 1)^2}{1 - t} + \mu(\{0\})\frac{(2t - 1)^2}{t}$$

$$+ \int_{(0,\infty)} \frac{(2t - 1)^2}{1 - \frac{s-1}{s}t} \frac{d\mu(s)}{s}.$$

By transformation $w = \frac{s-1}{s} : (0, \infty) \rightarrow (-\infty, 1)$, define a measure ν on $(-\infty, 1)$ by $d\nu(w) = s^{-1}\, d\mu(s)$. Since $\frac{1}{2-w} = \frac{s}{1+s}$, we have $\int_{(0,\infty)}(2 - w)^{-1}\, d\nu(w) = \int_{(0,\infty)}(1 + s)^{-1}\, d\mu(s) < +\infty$. Hence, setting $\nu(\{1\}) := c$, we find that g has an expression in the form (C.2).

Conversely, from expression (C.2) of g we have for every $t \in (0, \infty)$,

$$f(t) = (t + 1)g\left(\frac{t}{t + 1}\right)$$

$$= \alpha(t + 1) + \beta t + \gamma\frac{(t - 1)^2}{t} + \int_{(-\infty,1]} \frac{(t - 1)^2}{(t + 1) - wt}\, d\nu(w)$$

$$= \alpha + (\alpha + \beta)t + \gamma\frac{(t - 1)^2}{t} + \nu(\{1\})(t - 1)^2 + \int_{(-\infty,1)} \frac{(t - 1)^2}{t + \frac{1}{1-w}} \frac{d\nu(w)}{1 - w}.$$

By transformation $s = \frac{1}{1-w} : (-\infty, 1) \rightarrow (0, \infty)$, define a measure μ on $(0, \infty)$ by $d\mu(s) = (1 - w)^{-1}\, d\nu(w)$. Since $\frac{1}{1+s} = \frac{1-w}{2-w}$, we have $\int_{(0,\infty)}(1 + s)^{-1}\, d\mu(s) = \int_{(-\infty,1)}(2 - w)^{-1}\, d\nu(w) < +\infty$. Hence, setting $\mu(\{0\}) := \gamma$, we find that f has an expression in the form (2.5). So (a) $\Longleftrightarrow$ (b) has been shown.

Finally, the equalities in (C.3) are immediate and the uniqueness in expression (C.2) follows from that in (C.1) (or in (2.5)). $\square$

Remark C.2 The above proof indeed presents a proof of the integral expression in (2.5) of an operator convex function f on $(0, \infty)$ based on the well-known integral expression (C.1) of an operator convex function on $(-1, 1)$.

From $\gamma = \mu(\{0\})$ and $\nu(\{1\}) = c$ given in the proof of Theorem C.1, one can easily find that

$$\gamma = \lim_{t \to 0^+} t f(t) = \lim_{t \to 0^+} t g(t), \quad \nu(\{1\}) = \lim_{t \to \infty} \frac{f(t)}{t^2} = \lim_{t \to 1^-} (1 - t) g(t).$$

Thus, we have the following:

Corollary C.3 *Let g be an operator convex function on* $(0, 1)$ *such that*

$$\lim_{t \to 0^+} t g(t) = \lim_{t \to 1^-} (1 - t) g(t) = 0.$$

Then g has the integral expression

$$g(t) = \alpha + \beta t + \int_{(-\infty, 1)} \frac{(2t - 1)^2}{1 - wt} \, d\nu(w), \qquad t \in (0, 1),$$

where $\alpha, \beta \in \mathbb{R}$ *and* ν *is a positive measure on* $(-\infty, 1)$ *with* $\int_{(-\infty, 1)} (2-w)^{-1} \, d\nu(w) < +\infty$.

Corollary C.4 *Let g be a continuous real function on* $[0, 1]$. *Then the following conditions are equivalent:*

 (i) *g is operator convex on* $[0, 1]$;
(ii) *there exist* $\alpha, \beta \in \mathbb{R}$ *and an operator monotone function k on* $[0, \infty)$ *such that*

$$g(t) = \alpha + \beta t + (1 - t) k \left(\frac{t}{1 - t} \right), \qquad t \in [0, 1);$$

(iii) *there exist* $\alpha, \beta \in \mathbb{R}$ *and a positive measure* μ *on* $(0, \infty)$ *satisfying* $\int_{(0, \infty)} (1 + s)^{-1} \, d\mu(s) < +\infty$ *such that*

$$g(t) = \alpha(1 - t) + \beta t - \int_{(0, \infty)} \frac{t(1 - t)}{t + s(1 - t)} \, d\mu(s), \qquad t \in [0, 1].$$

Proof By Theorem C.1, g is operator convex on $[0, 1]$ if and only if there exists an operator convex function f on $(0, \infty)$ with $f(0^+), f'(\infty) < +\infty$ such that $g(t) = (1 - t) f(\frac{t}{1-t})$, $t \in (0, 1)$. Hence the corollary is immediately seen from [53, Proposition 8.4], while we omit the details. □

Appendix D
Operator Connections of Normal Positive Functionals

Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. An *operator connection* σ in Kubo-Ando's sense [73] is a binary operation $(A, B) \in B(\mathcal{H})_+ \times B(\mathcal{H})_+ \mapsto A \sigma B \in B(\mathcal{H})_+$ satisfying the following properties:

(I) *Joint monotonicity:* $A_1 \leq A_2$ and $B_1 \leq B_2$ imply $A_1 \sigma B_1 \leq A_2 \sigma B_2$.
(II) *Transformer inequality:* $C(A \sigma B)C \leq (CAC)\sigma(CBC)$ for any $C \in B(\mathcal{H})_+$. This extends to $C^*(A \sigma B)C \leq (C^*AC)\sigma(C^*BC)$ for any $C \in B(\mathcal{H})$ and equality holds if $C \in B(\mathcal{H})$ is invertible.
(III) *Downward continuity:* $A_n \searrow A$ and $B_n \searrow B$ imply $A_n \sigma B_n \searrow A \sigma B$, where $A_n \searrow A$ means that $A_1 \geq A_2 \geq \cdots$ and $A_n \to A$ in the strong operator topology.

It is well-known [73, Theorem 3.3] that there is a one-to-one correspondence between the operator connections σ and the operator monotone functions $k \geq 0$ on $[0, \infty)$, which are determined each other by

$$k(t)I = I \sigma (tI), \qquad t > 0,$$

$$A \sigma B = A^{1/2} k(A^{-1/2} B A^{-1/2}) A^{1/2}$$

for every $A, B \in B(\mathcal{H})_+$ with A invertible. In this case, write $k = k_\sigma$ and call it the *representing function* of σ. The transpose $\widetilde{\sigma}$ of σ is the operator connection corresponding to the operator monotone function $k_{\widetilde{\sigma}}(t) := t k_\sigma(t^{-1})$, $t > 0$, the transpose of k_σ. We have $A \widetilde{\sigma} B = B \sigma A$ for $A, B \in B(\mathcal{H})_+$. The *parallel sum* $\varphi : \psi$ is the operator connection of A, B corresponding to the operator monotone function $(1 + t^{-1})^{-1} = t/(1 + t)$ on $[0, \infty)$.

Theory of operator connections has recently been extended in [51] to unbounded positive operators (or positive forms), including τ-measurable positive operators affiliated with a semifinite von Neumann algebra and positive elements of Haagerup's L^p-spaces over a general von Neumann algebra.

In this section we will discuss operator connections of functionals in M_*^+ via the identification with Haagerup's $L^1(M)$ under the correspondence $\psi \in M_* \leftrightarrow h_\psi \in L^1(M)$ (Theorem A.33).

We begin with the following definition:

Definition D.1 For each $\varphi, \psi \in M_*^+$ we have positive operators $T_{\varphi/\varphi+\psi}$ and $T_{\psi/\varphi+\psi}$ in $s(\varphi + \psi) M s(\varphi + \psi)$ (see Lemma A.21 and Definition 8.1). Let σ be an operator mean (in Kubo-Ando's sense). Note that $T_{\varphi/\varphi+\psi} \sigma T_{\psi/\varphi+\psi} \in M_+$ since

$$u'^* (T_{\varphi/\varphi+\psi} \sigma T_{\psi/\varphi+\psi}) u' = (u'^* T_{\varphi/\varphi+\psi} u') \sigma (u'^* T_{\psi/\varphi+\psi} u') = T_{\varphi/\varphi+\psi} \sigma T_{\psi/\varphi+\psi}$$

for any unitary $u' \in M'$ (the commutant of M). Hence we can define the *connection* $\varphi \sigma \psi \in M_*^+$ of φ, ψ by

$$h_{\varphi\sigma\psi} = h_{\varphi+\psi}^{1/2} (T_{\varphi/\varphi+\psi} \sigma T_{\psi/\varphi+\psi}) h_{\varphi+\psi}^{1/2}.$$

Proposition D.2 *For every $\varphi, \psi \in M_*^+$,*

$$\varphi \widetilde{\sigma} \psi = \psi \sigma \varphi.$$

Proof From

$$\begin{aligned}
h_{\varphi\widetilde{\sigma}\psi} &= h_{\varphi+\psi}^{1/2} (T_{\varphi/\varphi+\psi} \widetilde{\sigma} T_{\psi/\varphi+\psi}) h_{\varphi+\psi}^{1/2} \\
&= h_{\varphi+\psi}^{1/2} (T_{\psi/\varphi+\psi} \sigma T_{\varphi/\varphi+\psi}) h_{\varphi+\psi}^{1/2} = h_{\psi\sigma\varphi},
\end{aligned}$$

the assertion follows. $\square$

Before giving main properties of $\varphi \sigma \psi$, we provide the next lemma on the equality case of the inequality (II) above. The equality is well-known when C is invertible, but this extension is not found in the literature. First recall the integral expression of the representing function k_σ:

$$k_\sigma(t) = \alpha + \beta t + \int_{(0,\infty)} \frac{t(1+s)}{t+s} \, d\mu(s), \qquad t \in [0, \infty), \tag{D.1}$$

where $\alpha, \beta \geq 0$ and μ is a finite positive measure on $(0, \infty)$.

Lemma D.3 *Let $A, B \in B(\mathcal{H})_+$ and $C \in B(\mathcal{H})$ be such that $s(A + B) \leq s(C^*C)$. Then for any operator connection σ,*

$$C^*(A \sigma B)C = (C^*AC) \sigma (C^*BC).$$

Proof Recall the integral expression [73, Theorem 3.4] corresponding to (D.1):

$$A \sigma B = \alpha A + \beta B + \int_{(0,\infty)} \frac{1+s}{s} \{(sA) : B\} \, d\mu(s). \tag{D.2}$$

Hence it suffices to prove that

$$C^*(A : B)C = (C^*AC) : (C^*BC). \tag{D.3}$$

To show this, we may assume that $s(C^*C) = I$. Here, recall the well-known variational expression of $A : B$ (see, e.g., [45, Lemma 3.1.5]): for every $\xi \in \mathcal{H}$,

$$\langle \xi, (A : B)\xi \rangle = \inf\{\langle \eta, A\eta \rangle + \langle \zeta, B\zeta \rangle : \eta, \zeta \in \mathcal{H}, \ \eta + \zeta = \xi\}. \qquad (D.4)$$

This gives

$$\langle \xi, C^*(A : B)C\xi \rangle = \inf\{\langle \eta, A\eta \rangle + \langle \zeta, B\zeta \rangle : \eta, \zeta \in \mathcal{H}, \ \eta + \zeta = C\xi\}.$$

For any $\eta, \zeta \in \mathcal{H}$ with $\eta + \zeta = C\xi$ one can choose $\eta_n \in \mathcal{H}$ such that $\|C\eta_n - \eta\| \to 0$. Set $\zeta_n := \xi - \eta_n$; then $C\eta_n + C\zeta_n = C\xi$ and $\|C\zeta_n - \zeta\| \to 0$. Since

$$\langle \eta_n, C^*AC\eta_n \rangle + \langle \zeta_n, C^*BC\zeta_n \rangle \geq \langle \xi, ((C^*AC) : (C^*BC))\xi \rangle.$$

letting $n \to \infty$ gives $\langle \eta, A\eta \rangle + \langle \zeta, B\zeta \rangle \geq \langle \xi, ((C^*AC) : (C^*BC))\xi \rangle$, which implies that

$$\langle C\xi, (A : B)C\xi \rangle \geq \langle \xi, ((C^*AC) : (C^*BC))\xi \rangle.$$

Therefore, $C^*(A : B)C \geq (C^*AC) : (C^*BC)$, and (D.3) follows since the opposite inequality holds in general. $\qquad \square$

Lemma D.4 *Let $\varphi, \psi, \omega \in M_*^+$ be such that $\varphi + \psi \leq \alpha\omega$ for some $\alpha > 0$. Then*

$$h_{\varphi\sigma\psi} = h_\omega^{1/2}(T_{\varphi/\omega}\sigma T_{\psi/\omega})h_\omega^{1/2}.$$

In particular, if $\psi \leq \alpha\varphi$ for some $\alpha > 0$, then $h_{\varphi\sigma\psi} = h_\varphi^{1/2}k_\sigma(T_{\psi/\varphi})h_\varphi^{1/2}$, where k_σ is the representing function of σ.

Proof By Lemma A.21 there are $A, B \in s(\varphi + \psi)Ms(\varphi + \psi)$ and $C \in s(\omega)Ms(\omega)$ such that

$$h_\varphi^{1/2} = Ah_{\varphi+\psi}^{1/2}, \qquad h_\psi^{1/2} = Bh_{\varphi+\psi}^{1/2}, \qquad h_{\varphi+\psi}^{1/2} = Ch_\omega^{1/2},$$

and C is invertible in $s(\omega)Ms(\omega)$. Since $h_\varphi^{1/2} = ACh_\omega^{1/2}$ and $h_\psi^{1/2} = BCh_{\varphi+\psi}^{1/2}$, we have

$$T_{\varphi/\omega} = (AC)^*(AC) = C^*T_{\varphi/\varphi+\psi}C, \quad T_{\psi/\omega} = (BC)^*(BC) = C^*T_{\psi/\varphi+\psi}C.$$

Therefore, by Lemma D.3,

$$\begin{aligned}
h_\omega^{1/2}(T_{\varphi/\omega}\sigma T_{\psi/\omega})h_\omega^{1/2} &= h_\omega^{1/2}((C^*T_{\varphi/\varphi+\psi}C)\sigma(C^*T_{\psi/\varphi+\psi}C))h_\omega^{1/2} \\
&= h_\omega^{1/2}C^*(T_{\varphi/\varphi+\psi}\sigma T_{\psi/\varphi+\psi})Ch_\omega^{1/2} \\
&= h_{\varphi+\psi}^{1/2}(T_{\varphi/\varphi+\psi}\sigma T_{\psi/\varphi+\psi})h_{\varphi+\psi}^{1/2} = h_{\varphi\sigma\psi},
\end{aligned}$$

showing the first result. The latter result follows from the first by taking $\omega = \varphi$. $\quad \square$

Example D.5 Consider the case when $M = B(\mathcal{H})$ on an arbitrary Hilbert space $\mathcal{H}$, which is standardly represented by the left multiplication on the Hilbert-Schmidt class $C_2(\mathcal{H})$, a Hilbert space with the Hilbert-Schmidt inner product. The trace-class $C_1(\mathcal{H})$ is identified with $B(\mathcal{H})_*$ as usual. Let $A, B \in C_1(\mathcal{H})_+$, corresponding to $\varphi = \mathrm{Tr}(A\cdot), \psi = \mathrm{Tr}(B\cdot) \in M(\mathcal{H})_*^+$. We have $\varphi\sigma\psi$ in Definition D.1 as well as $A\sigma B$ in Kubo-Ando's sense. [73]. We write $A = (A + B)^{1/2}T_{A/A+B}(A + B)^{1/2}$

and $B = (A + B)^{1/2}T_{B/A+B}(A + B)^{1/2}$ with positive operators $T_{A/A+B}, T_{B/A+B}$ in $s(A+B)B(\mathcal{H})s(A + B)$. Note that these are indeed $T_{\varphi/\varphi+\psi}, T_{\psi/\varphi+\psi}$, respectively. By Lemma D.3,

$$A\sigma B = (A + B)^{1/2}(T_{A/A+B}\sigma T_{B/A+B})(A + B)^{1/2},$$

which means that $\varphi\sigma\psi$ corresponds to $A\sigma B$, or $\varphi\sigma\psi$ coincides with $A\sigma B$.

The next theorem summarizes basic properties of the connection $\varphi\sigma\psi$.

Theorem D.6 *Let $\varphi, \psi, \varphi_i, \psi_i \in M_*^+$.*

(i) Joint monotonicity: *If $\varphi_1 \leq \varphi_2$ and $\psi_1 \leq \psi_2$, then $\varphi_1\sigma\psi_1 \leq \varphi_2\sigma\psi_2$.*
(ii) Transformer inequality: *For every $a \in M$,*

$$a^*(\varphi\sigma\psi)a \leq (a^*\varphi a)\sigma(a^*\psi a), \tag{D.5}$$

   where $(a^*\varphi a)(x) := \varphi(axa^*)$, $x \in M$. Moreover, if a is invertible, then equality hold in the above inequality.
(iii) Downward continuity: *If $\varphi_n \searrow \varphi$ and $\psi_n \searrow \psi$ in M_*^+, then $\varphi_n\sigma\psi_n \searrow \varphi\sigma\psi$, where $\varphi_n \searrow \varphi$ means that $\varphi_1 \geq \varphi_2 \geq \cdots$ and $\|\varphi_n - \varphi\| \to 0$. In particular, for any $\omega \in M_*^+$,*

$$\varphi\sigma\psi = \lim_{\varepsilon\searrow 0}(\varphi + \varepsilon\omega)\sigma(\psi + \varepsilon\omega) \quad \text{in the norm (decreasingly).} \tag{D.6}$$

(iv) Joint concavity:

$$(\varphi_1 + \varphi_2)\sigma(\psi_1 + \psi_2) \geq (\varphi_1\sigma\psi_1) + (\varphi_2\sigma\psi_2).$$

Proof (i) Let $\omega := \varphi_1 + \varphi_2 + \psi_1 + \psi_2$. It is easy to see that $T_{\varphi_1/\omega} \leq T_{\varphi_2/\omega}$ and $T_{\psi_1/\omega} \leq T_{\psi_2/\omega}$. Hence by Lemma D.4 we have

$$h_{\varphi_1\sigma\psi_1} = h_\omega^{1/2}(T_{\varphi_1/\omega}\sigma T_{\psi_1/\omega})h_\omega^{1/2} \leq h_\omega^{1/2}(T_{\varphi_2/\omega}\sigma T_{\psi_2/\omega})h_\omega^{1/2} = h_{\varphi_2\sigma\psi_2}.$$

(ii) This will be proved below after we give the integral expression of $\varphi\sigma\psi$.

(iii) Joint monotonicity in (i) says that $\varphi_n\sigma\psi_n$ is monotone decreasing as $n \to \infty$. Hence a positive linear functional ϕ on M is defined by $\phi(x) := \lim_{n\to\infty}(\varphi_n\sigma\psi_n)(x)$, $x \in M$. Since $\varphi_n\sigma\psi_n \geq \phi$, it is clear that ϕ is normal, so $\phi \in M_*^+$ and $\|\varphi_n\sigma\psi_n - \phi\| = (\varphi_n\sigma\psi_n - \phi)(1) \to 0$ as $n \to \infty$. To prove that $\phi = \varphi\sigma\psi$, let $\omega := \varphi_1 + \psi_1$. It follows that $T_{\varphi_n/\omega} \searrow T_1$ and $T_{\psi_n/\omega} \searrow T_2$ for some $T_1, T_2 \in M_+$. For every $x \in M_+$ one has

$$\varphi_n(x) = \text{tr}\,(h_\omega^{1/2}xh_\omega^{1/2})T_{\varphi_n/\omega} \searrow \text{tr}\,(h_\omega^{1/2}xh_\omega^{1/2})T_1$$

as $n \to \infty$, so that $\varphi(x) = \text{tr}\,xh_\omega^{1/2}T_1h_\omega^{1/2}$, implying $h_\varphi^{1/2} = h_\omega^{1/2}T_1h_\omega^{1/2}$. This implies that $T_1 = T_{\varphi/\omega}$ and similarly $T_2 = T_{\psi/\omega}$. Hence by Lemma D.4 one has

$$(\varphi_n\sigma\psi_n)(x) = \text{tr}\,(h_\omega^{1/2}xh_\omega^{1/2})(T_{\varphi_n/\omega}\sigma T_{\psi_n/\omega})$$
$$\searrow \text{tr}\,(h_\omega^{1/2}xh_\omega^{1/2})(T_{\varphi/\omega}\sigma T_{\psi/\omega}) = (\varphi\sigma\psi)(x)$$

as $n \to \infty$. Therefore, $\phi = \varphi \sigma \psi$.

(iv) With $\omega := \varphi_1 + \varphi_2 + \psi_1 + \psi_2$, it is easy to see that $T_{(\varphi_1+\varphi_2)/\omega} = T_{\varphi_1/\omega} + T_{\varphi_2/\omega}$ and $T_{(\psi_1+\psi_2)/\omega} = T_{\psi_1/\omega} + T_{\psi_2/\omega}$. Hence by Lemma D.4 one has

$$
\begin{aligned}
h_{(\varphi_1+\varphi_2)\sigma(\psi_1+\psi_2)} &= h_\omega^{1/2}((T_{\varphi_1/\omega} + T_{\varphi_2/\omega})\sigma(T_{\psi_1/\omega} + T_{\psi_2/\omega}))h_\omega^{1/2} \\
&\geq h_\omega^{1/2}((T_{\varphi_1/\omega}\sigma T_{\psi_1/\omega}) + (T_{\varphi_2/\omega}\sigma T_{\psi_2/\omega}))h_\omega^{1/2} \\
&= h_{\varphi_1\sigma\psi_1} + h_{\varphi_2\sigma\psi_2} = h_{\varphi_1\sigma\psi_1+\varphi_2\sigma\psi_2},
\end{aligned}
$$

where the inequality above is due to the joint concavity in [73, Theorem 3.5] (also [45, Corollary 3.2.3]). $\qquad \square$

Corresponding to (D.1) we have the integral expression of $\varphi\sigma\psi$ as follows:

Theorem D.7 *For every $\varphi, \psi \in M_*^+$,*

$$
(\varphi\sigma\psi)(x) = \alpha\varphi(x) + \beta\psi(x) + \int_{(0,\infty)} \frac{1+s}{s}((s\varphi) : \psi)(x)\, d\mu(s), \qquad x \in M.
$$

Proof Let $\omega := \varphi + \psi$. By applying (D.2) to $T_{\varphi/\omega}$ and $T_{\psi/\omega}$ we have

$$
\begin{aligned}
(\varphi\sigma\psi)(x) &= \mathrm{tr}\,(h_\omega^{1/2}xh_\omega^{1/2})(T_{\varphi/\omega}\sigma T_{\psi/\omega}) \\
&= \alpha\mathrm{tr}\,(h_\omega^{1/2}xh_\omega^{1/2})T_{\varphi/\omega} + \beta\mathrm{tr}\,(h_\omega^{1/2}xh_\omega^{1/2})T_{\psi/\omega} \\
&\quad + \int_{(0,\infty)} \frac{1+s}{s}\,\mathrm{tr}\,(h_\omega^{1/2}xh_\omega^{1/2})((sT_{\varphi/\omega}) : T_{\psi/\omega})\, d\mu(s) \\
&= \alpha\varphi(x) + \beta\psi(x) + \int_{(0,\infty)} \frac{1+s}{s}((s\varphi) : \psi)(x)\, d\mu(s),
\end{aligned}
$$

as asserted. $\qquad \square$

Remark D.8 In the above theorem the integral $\int_{(0,\infty)} \frac{1+s}{s}((s\varphi) : \psi)\, d\mu(s)$ is given in the weak sense (or the $\sigma(M_*, M)$-sense). But it exists in a stronger sense as follows. By Theorem D.6 (i) note that

$$
\|(s\varphi) : \psi\| = ((s\varphi) : \psi)(1) \leq ((s\omega) : \omega)(1) = \frac{s}{1+s}\omega(1)
$$

and hence $\int_{(0,\infty)} \frac{1+s}{s}\|(s\varphi) : \psi\|\, d\mu(s) < +\infty$. Moreover, it is easy to see that if $\varphi \sim \psi$ then the function $s \in (0,\infty) \mapsto \frac{1+s}{s}((s\varphi) : \psi) \in M_*^+$ is continuous in the norm. Thus, in view of (D.6), the integral expression

$$
\varphi\sigma\psi = \alpha\varphi + \beta\psi + \int_{(0,\infty)} \frac{1+s}{s}((s\varphi) : \psi)\, d\mu(s)
$$

exists as the Bochner integral in the norm.

The following variational expression is the variant of (D.4) for the parallel sum $\varphi : \psi$.

Theorem D.9 *For every* $\varphi, \psi \in M_*^+$ *and any* $x \in M$,

$$(\varphi : \psi)(x^*x) = \inf\{\varphi(y^*y) + \psi(z^*z) : y, z \in M, \; y + z = x\}. \qquad (\text{D.7})$$

Proof Let $\omega := \varphi + \psi$. From (D.6) and

$$\inf_{y+z=x} \{\varphi(y^*y) + \psi(z^*z)\} = \inf_{y+z=x} \inf_{\varepsilon>0} \{(\varphi + \varepsilon\omega)(y^*y) + (\psi + \varepsilon\omega)(z^*z)\}$$

$$= \inf_{\varepsilon>0} \inf_{y+z=x} \{(\varphi + \varepsilon\omega)(y^*y) + (\psi + \varepsilon\omega)(z^*z)\},$$

it suffices to prove (D.7) in the case $\varphi \sim \psi$. Let $e := s(\varphi) = s(\omega)$ and choose an $A \in eMe$ such that $h_\psi^{1/2} = Ah_\omega^{1/2} = h_\omega^{1/2}A^*$. Then A is invertible in eMe and $T_{\varphi/\omega} = A^*A$. Since $T_{\varphi/\omega} + T_{\psi/\omega} = e$, we find that

$$h_{\varphi:\psi} = h_\omega^{1/2}(T_{\varphi/\omega} : T_{\psi/\omega})h_\omega^{1/2}$$

$$= h_\omega^{1/2}T_{\varphi/\omega}(T_{\varphi/\omega} + T_{\psi/\omega})^{-1}T_{\psi/\omega}h_\omega^{1/2}$$

$$= h_\omega^{1/2}A^*A(e - A^*A)h_\omega^{1/2} = h_\varphi^{1/2}(e - AA^*)h_\varphi^{1/2}.$$

When $y, z \in M$ and $y + z = x$,

$$\varphi(y^*y) + \psi(z^*z) - (\varphi : \psi)(x^*x)$$

$$= \text{tr}\,(x - z)^*(x - z)h_\varphi + \text{tr}\,z^*zh_\psi - \text{tr}\,x^*x(h_\varphi - h_\varphi^{1/2}AA^*h_\varphi^{1/2})$$

$$= \text{tr}\,z^*zh_\varphi + \text{tr}\,z^*zh_\psi - \text{tr}\,x^*zh_\varphi - \text{tr}\,z^*xh_\varphi + \text{tr}\,x^*xh_\varphi^{1/2}AA^*h_\varphi^{1/2}$$

$$= \text{tr}\,z^*zh_\omega + \text{tr}\,A^*h_\varphi^{1/2}x^*xh_\varphi^{1/2}A - 2\text{Re}\,\text{tr}\,x^*zh_\varphi$$

$$= \|zh_\omega^{1/2}\|_2^2 + \|xh_\varphi^{1/2}A\|_2^2 - 2\text{Re}\,\text{tr}\,A^*h_\varphi^{1/2}x^*zh_\varphi^{1/2}A^{*-1}$$

$$= \|zh_\omega^{1/2}\|_2^2 + \|xh_\varphi^{1/2}A\|_2^2 - 2\text{Re}\,\langle xh_\varphi^{1/2}A, zh_\omega^{1/2}\rangle$$

thanks to $h_\omega^{1/2} = A^{-1}h_\varphi^{1/2} = h_\varphi^{1/2}A^{*-1}$. Therefore, by the Schwarz inequality, $\varphi(y^*y) + \psi(z^*z) - (\varphi : \psi)(x^*x) \geq 0$. Moreover, since $\overline{Mh_\omega^{1/2}} = L^2(M)e$ and $xh_\varphi^{1/2}A \in L^2(M)e$, one can choose a sequence $\{z_n\} \subset M$ such that $\|z_nh_\omega^{1/2} - xh_\varphi^{1/2}A\| \to 0$. Letting $y_n := x - z_n$ one has

$$\varphi(y_n^*y_n) + \psi(z_n^*z_n) - (\varphi : \psi)(x^*x)$$

$$= \|z_nh_\omega^{1/2}\|_2^2 + \|xh_\varphi^{1/2}A\|_2^2 - 2\text{Re}\,\langle xh_\varphi^{1/2}A, z_nh_\omega^{1/2}\rangle \longrightarrow 0,$$

so that (D.7) follows.   □

Now we are in a position to prove Theorem D.6 (ii).

Proof (Theorem D.6 (ii)) By Theorem D.7, to obtain (D.5), it suffices to prove it for the parallel sum $\varphi : \psi$. For every $x, a \in M$, using (D.7) twice we have

$$
\begin{aligned}
(a^*(\varphi : \psi)a)(x^*x) &= (\varphi : \psi)((xa^*)^*(xa^*)) \\
&= \inf\{\varphi(y^*y) + \psi(z^*z) : y, z \in M, \ y + z = xa^*\} \\
&\leq \inf\{\varphi((ya^*)^*(ya^*)) + \psi((za^*)^*(za^*)) : y, z \in M, \ y + z = x\} \\
&= \inf\{(a^*\varphi a)(y^*y) + (a^*\psi a)(z^*z) : y, z \in M, \ y + z = x\} \\
&= ((a^*\varphi a) : (a^*\psi a))(x^*x),
\end{aligned}
$$

implying the assertion.

Next, assume that a is invertible. Apply (D.5) to $a^*\varphi a$, $a^*\psi a$ and a^{-1} to see that

$$
a^{*-1}((a^*\varphi a)\sigma(a^*\psi a))a^{-1} \leq (a^{*-1}(a^*\varphi a)a^{-1})\sigma(a^{*-1}(a^*\psi a)a^{-1}) = \varphi\sigma\psi,
$$

so that $(a^*\varphi a)\sigma(a^*\psi a) \leq a^*(\varphi\sigma\psi)a$. Hence equality follows. □

Let $\mathcal{H}$ and $\mathcal{K}$ be separable Hilbert spaces. For any operator connection σ and any positive linear map $\alpha : B(\mathcal{H}) \to B(\mathcal{K})$, the following inequality is well-known:

$$
\alpha(A\sigma B) \leq \alpha(A)\sigma\alpha(B), \qquad A, B \in B(\mathcal{H}), \tag{D.8}
$$

which is essentially due to Ando [4] while proved only for the geometric mean and the parallel sum. Inequality (D.8) for general σ can be proved, for instance, by using the integral expression (D.2).

Proposition D.10 *Let $\gamma : N \to M$ be a normal positive linear map between von Neumann algebras. Then for every $\varphi, \psi \in M_*^+$,*

$$
(\varphi \circ \gamma)\sigma(\psi \circ \gamma) \geq (\varphi\sigma\psi) \circ \gamma.
$$

Proof Although γ is not necessarily unital here, with $\omega := \varphi + \psi$ one can define a normal positive map $\gamma_\omega^* : M \to e_0 M e_0$ (where $e_0 := s(\omega \circ \gamma)$ as in Lemma 8.3 (also Proposition 6.6). Then the result can be shown as in Lemma 8.6. Indeed, we have

$$
\begin{aligned}
h_{(\varphi\sigma\psi)\circ\gamma} &= \gamma_*(h_{\varphi\sigma\psi}) \\
&= \gamma_*(h_\omega^{1/2}(T_{\varphi/\omega}\sigma T_{\psi/\omega})h_\omega^{1/2}) \\
&= h_{\omega\circ\gamma}^{1/2}\gamma_\omega^*(T_{\varphi/\omega}\sigma T_{\psi/\omega})h_{\omega\circ\gamma}^{1/2} \\
&\geq h_{\omega\circ\gamma}^{1/2}(\gamma_\omega^*(T_{\varphi/\omega})\sigma\gamma_\omega^*(T_{\psi/\omega}))h_{\omega\circ\gamma}^{1/2} \\
&= h_{\omega\circ\gamma}^{1/2}(T_{\varphi\circ\gamma/\omega\circ\gamma}\sigma T_{\psi\circ\gamma/\omega\circ\gamma})h_{\omega\circ\gamma}^{1/2} \\
&= h_{(\varphi\circ\gamma)\sigma(\psi\circ\gamma)}.
\end{aligned}
$$

In the above, the third and the fourth equalities are due to Lemmas 8.3 and 8.5, respectively, which hold without the unitality assumption of γ. The above inequality follows from (D.8). □

The last result is the relation between the connection $\varphi\sigma\psi$ and the maximal f-divergence.

Proposition D.11 *For every operator connection σ and every $\varphi, \psi \in M_*^+$,*

$$\widehat{S}_{-k_\sigma}(\psi \| \varphi) = -(\varphi \sigma \psi)(1).$$

Proof When $\varphi \sim \psi$, by Definition 4.1 and Lemma D.4,

$$\widehat{S}_{-k_\sigma}(\psi \| \varphi) = \mathrm{tr}\,(h_\varphi^{1/2}(-k_\sigma)(T_{\psi/\varphi})h_\varphi^{1/2}) = -(\varphi \sigma \psi)(1).$$

For general $\varphi, \psi \in M_*^+$, by Definitions 4.3 and (D.6), with $\omega := \varphi + \psi$ we have

$$\begin{aligned}
\widehat{S}_{-k_\sigma}(\psi \| \varphi) &= \lim_{\varepsilon \searrow 0} \widehat{S}_{-k_\sigma}(\psi + \varepsilon\omega \| \varphi + \varepsilon\omega) \\
&= -\lim_{\varepsilon \searrow 0}((\varphi + \varepsilon\omega)\sigma(\psi + \varepsilon\omega))(1) = -(\varphi \sigma \psi)(1),
\end{aligned}$$

as asserted. $\square$

References

1. L. Accardi and C. Cecchini, Conditional expectations in von Neumann algebras and a theorem of Takesaki, *J. Funct. Anal.* **45** (1982), 245–273.
2. P. M. Alberti, A note on the transition probability over D^*-algebras, *Lett. Math. Phys.* **7** (1983), 25–32.
3. P. M. Alberti and A. Uhlmann, On Bures distance and $*$-algebraic transition probability between inner derived positive linear forms over W^*-algebras, *Acta Appl. Math.* **60** (2000), 1–37.
4. T. Ando, Concavity of certain maps on positive definite matrices and applications to Hadamard products, *Linear Algebra Appl.* **26** (1979) 203–241.
5. H. Araki, Relative Hamiltonian for faithful normal states of a von Neumann algebra, *Publ. Res. Inst. Math. Sci.* **9** (1973), 165–209.
6. H. Araki, Golden-Thompson and Peierls-Bogolubov inequalities for a general von Neumann algebras, *Comm. Math. Phys.* **34** (1973), 167–178.
7. H. Araki, Some properties of modular conjugation operator of von Neumann algebras and a non-commutative Radon-Nikodym theorem with a chain rule, *Pacific J. Math.* **50** (1974), 309–354.
8. H. Araki, Relative entropy of states of von Neumann algebras, *Publ. Res. Inst. Math. Sci.* **11** (1976), 809–833.
9. H. Araki, Relative entropy for states of von Neumann algebras II, *Publ. Res. Inst. Math. Sci.* **13** (1977), 173–192.
10. H. Araki and T. Masuda, Positive Cones and L_p-Spaces for von Neumann Algebras, *Publ. Res. Inst. Math. Sci.* **18** (1982), 339–411.
11. A. Arias, A. Gheondea and S. Gudder, Fixed points of quantum operations, *J. Math. Phys.* **43** (2002), 5872–5881.
12. K. M. R. Audenaert and N. Datta, α-z-relative entropies, *J. Math. Phys.* **56** (2015), 022202.
13. S. Beigi, Sandwiched Rényi divergence satisfies data processing inequality, *J. Math. Phys.* **54** (2013), 122202, 11 pp.
14. V. P. Belavkin and P. Staszewski, C^*-algebraic generalization of relative entropy and entropy, *Ann. Inst. H. Poincaré Phys. Théor.* **37**, (1982), 51–58.
15. J. Bergh and J. Löfström, *Interpolation Spaces: An Introduction*, Springer-Verlag, Berlin-Heidelberg-New York, 1976.
16. M. Berta, O. Fawzi and M. Tomamichel, On variational expressions for quantum relative entropies, *Lett. Math. Phys.* **107** (2017), 2239–2265.
17. M. Berta, V. B. Scholz and M. Tomamichel, Rényi divergences as weighted non-commutative vector valued L_p-spaces, *Ann. Henri Poincaré* **19** (2018), 1843–1867.
18. R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1996.
19. O. Bratteli, P. E. T. Jorgensen, A. Kishimoto and R. F. Werner, Pure states on O_d, *J. Operator Theory* **43** (2000), 97–143.

20. O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics 1*, 2nd ed., Springer-Verlag, New York, 1987.
21. O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics 2*, 2nd ed., Springer-Verlag, Berlin, 1997.
22. M.-D. Choi, A Schwarz inequality for positive linear maps on C^*-algebras, *Illinois J. Math.* **18** (1974), 565–574.
23. A. Connes, Sur le théoème de Radon-Nikodym pour les poids normaux fidéles semi-finis, *Bull. Sci. Math. (2)* **97** (1973), 253–258.
24. A. Connes, Une classification des facteurs de type III *Ann. Sci. École Norm. Sup. (4)* **6** (1973), 133–252.
25. A. Connes, Caractérisation des espaces vectoriels ordonnés sous-jacents aux algèbres de von Neumann, *Ann. Inst. Fourier (Grenoble)* **24** (1974), 121–155.
26. A. Connes, Classification of injective factors, Cases II_1, II_∞, III_λ, $\lambda \neq 1$, *Ann. of Math. (2)* **104** (1976), 73–115.
27. A. Connes and M. Takesaki, The flow of weights on factors of type III, *Tôhoku Math. J.* **29** (1977), 473–575.
28. N. Datta, Min- and max-relative entropies and a new entanglement monotone, *IEEE Trans. Inform. Theory* **55** (2009), 2816–2826.
29. J. Dixmier, Formes linéaires sur un anneau d'opérateurs, *Bull. Soc. Math. France* **81** (1953), 9–39.
30. M. J. Donald, Relative hamiltonians which are not bounded from above, *J. Funct. Anal.* **91** (1990), 143–173.
31. A. Ebadian, I. Nikoufar, and M. E. Gordji, *Perspectives of matrix convex functions*, Proc. Natl. Acad. Sci. USA **108** (2011), 7313–7314.
32. E. G. Effros, *A matrix convexity approach to some celebrated quantum inequalities*, Proc. Natl. Acad. Sci. USA **106** (2009), 1006–1008.
33. T. Fack and H. Kosaki, Generalized s-numbers of σ-measurable operators, *Pacific J. Math.* **123** (1986), 269–300.
34. R. L. Frank and E. H. Lieb, Monotonicity of a relative Rényi entropy, *J. Math. Phys.* **54** (2013), 122201, 5 pp.
35. U. Franz, F. Hiai and É. Ricard, Higher order extension of Löwner's theory: Operator k-tone functions, *Trans. Amer. Math. Soc.* **366** (2014), 3043–3074.
36. T. Furuta, $A \geq B \geq 0$ assures $(B^r A^p B^r)^{1/q} \geq B^{(p+2r)/q}$ for $r \geq 0$, $p \geq 0$, $q \geq 1$ with $(1 + 2r)q \geq p + 2r$, *Proc. Amer. Math. Soc.* **101** (1987), 85–88.
37. U. Haagerup, The standard form of von Neumann algebras, Notes, Copenhagen Univ., 1973.
38. U. Haagerup, The standard form of von Neumann algebras, *Math. Scand.* **37** (1975), 271–283.
39. U. Haagerup, On the dual weights for crossed products of von Neumann algebras I, Removing separability conditions, *Math. Scand.* **43** (1978/79), 99–118.
40. U. Haagerup, On the dual weights for crossed products of von Neumann algebras II, Application of operator-valued weights, *Math. Scand.* **43** (1978/79), 119–140.
41. U. Haagerup, L^p-spaces associated with an arbitrary von Neumann algebra, Colloq. Internat. CNRS, no. 274, CNRS, Paris, 1979, pp. 175–184.
42. U. Haagerup, On the uniqueness of injective III_1 factor, Preprint, 1985; arXiv:1606.03156 [math.OA].
43. F. Hansen, An operator inequality, *Math. Ann.* **246** (1980), 249–250.
44. M. Hayashi, *Quantum Information. An introduction*, Springer-Verlag, Berlin, 2006.
45. F. Hiai, Matrix Analysis: Matrix Monotone Functions, Matrix Means, and Majorization, *Interdisciplinary Information Sciences* **16** (2010), 139–248.
46. F. Hiai, Concavy of certain matrix trace and norm functions, *Linear Algebra Appl.* **439** (2013), 1568–1589.
47. F. Hiai, Concavy of certain matrix trace and norm functions. II, *Linear Algebra Appl.* **496** (2016), 193–220.
48. F. Hiai, Quantum f-divergences in von Neumann algebras I. Standard f-divergences, *J. Math. Phys.* **59** (2018), 102202, 27 pp.

49. F. Hiai, Quantum f-divergences in von Neumann algebras II. Maximal f-divergences, *J. Math. Phys.* **60** (2019), 012203, 30 pp.

50. F. Hiai, *Concise Lectures on Selected Topics of von Neumann Algebras*, An intensive course at Department of Mathematical Analysis, Budapest University of Technology and Economics, 2019.

51. F. Hiai and H. Kosaki, in preparation.

52. F. Hiai and M. Mosonyi, Different quantum f-divergences and the reversibility of quantum operations, *Rev. Math. Phys.* **29** (2017), 1750023, 80 pp.

53. F. Hiai, M. Mosonyi, D. Petz and C. Bény, Quantum f-divergences and error correction, *Rev. Math. Phys.* **23** (2011), 691–747. Erratum: Quantum f-divergences and error correction, *Rev. Math. Phys.* **29** (2017), 1792001, 2 pp.

54. F. Hiai and Y. Nakamura, Distance between unitary orbits in von Neumann algebras, *Pacific J. Math.* **138** (1989), 259–294.

55. F. Hiai, M. Ohya and M. Tsukada, Sufficiency, KMS condition and relative entropy in von Neumann algebras, *Pacific J. Math.* **96** (1981), 99–109.

56. F. Hiai and D. Petz, The Golden-Thompson trace inequality is complemented, *Linear Algebra Appl.* **181** (1993), 153–185.

57. V. Jaksic, Y. Ogata, Y. Pautrat and C.-A. Pillet, Entropic fluctuations in quantum statistical mechanics. An introduction, in Quantum Theory from Small to Large Scales, August 2010, Lecture Notes of the Les Houches Summer School, Vol. 95 (Oxford University Press, 2012), pp. 213   410; arXiv:1106.3786.

58. A. Jenčová, Reversibility conditions for quantum operations, *Rev. Math. Phys.* **24** (2012), 1250016, 26 pp.

59. A. Jenčová, Rényi relative entropies and noncommutative L_p-spaces, *Ann. Henri Poincaré* **19** (2018), 2513–2542.

60. A. Jenčová, Rényi relative entropies and noncommutative L_p-spaces II, arXiv:1707.00047 [quant-ph].

61. A. Jenčová, A remark on the relative modular operator, Private communication, 2017, September.

62. A. Jenčová and D. Petz, Sufficiency in quantum statistical inference, *Comm. Math. Phys.* **263** (2006), 259–276.

63. A. Jenčová and D. Petz, Sufficiency in quantum statistical inference. A survey with examples, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **9** (2006), 331–351.

64. A. Jenčová, D. Petz and J. Pitrik, Markov triplets on CCR-algebras, *Acta Sci. Math. (Szeged)* **76** (2010), 111–134.

65. R. V. Kadison, Strong continuity of operator functions, *Pacific J. Math.* **26** (1968), 121–129.

66. R. V. Kadison and J. R. Ringrose, *Fundamentals of the Theory of Operator Algebras, I. Elementary theory, II. Advanced theory*, Pure and Applied Mathematics, Vol. 100. Academic Press, 1983, 1986.

67. H. Kosaki, Linear Radon-Nikodym theorem for states on a von Neumann algebra, *Publ. Res. Inst. Math. Sci.* **18** (1981), 379–386.

68. H. Kosaki, Interpolation theory and the Wigner-Yanase-Dyson-Lieb concavity, *Comm. Math. Phys.* **87** (1982), 315–329.

69. H. Kosaki, Applications of the complex interpolation method to a von Neumann algebra: non-commutative L^p-spaces, *J. Funct. Anal.* **56** (1984), 29–78.

70. H. Kosaki, Relative entropy of states: a variational expression, *J. Operator Theory* **16** (1986), 335–348.

71. H. Kosaki, An inequality of Araki-Lieb-Thirring (von Neumann algebra case), *Proc. Amer. Math. Soc.* **114** (1992), 477–481.

72. H. Kosaki, A remark on Sakai's quadratic Radon-Nikodym theorem, *Proc. Amer. Math. Soc.* **116** (1992), 783–786.

73. F. Kubo and T. Ando, Means of positive linear operators, *Math. Ann.* **246** (1980), 205–224.

74. B. Kümmerer and R. Nagel, Mean ergodic semigroups on W^*-algebras, *Acta Sci. Math. (Szeged)* **41** (1979), 151–159.

75. A. Lesniewski and M. B. Ruskai, Monotone Riemannian metrics and relative entropy on noncommutative probability spaces, *J. Math. Phys.* **40** (1999), 5702–5724.

76. E. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture, *Adv. Math.* **11** (1973), 267–288.

77. A. Łuczak, Quantum sufficiency in the operator algebra framework, *Internat. J. Theoret. Phys.* **53** (2014), 3423–3433.

78. A. Łuczak, On a general concept of sufficiency in von Neumann algebras, *Probab. Math. Statist.* **35** (2015), 313–324.

79. K. Matsumoto, A new quantum version of f-divergence, In: *Reality and Measurement in Algebraic Quantum Theory*, Springer Proc. Math. Stat., 261, Springer, Singapore, 2018, pp. 229–273.

80. K. Matsumoto, On single-copy maximization of measures f-divergences between a given pair of quantum states, arXiv:1412.3676v5 [quant-ph], 2016.

81. M. Mosonyi, Convexity properties of the quantum Rényi divergences, with applications to the quantum Stein's lemma, In: 9th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2014), pp. 88–98, *LIPIcs. Leibniz Int. Proc. Inform.*, Vol. 27, 2014. arXiv:1407.1067 [quant-ph].

82. M. Mosonyi and F. Hiai, On the quantum Rényi relative entropies and related capacity formulas, *IEEE Trans. Inform. Theory* **57** (2011), 2474–2487.

83. M. Mosonyi and T. Ogawa, Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies, *Comm. Math. Phys.* **334** (2015), 1617–1648.

84. M. Mosonyi and T. Ogawa, Strong converse exponent for classical-quantum channel coding *Comm. Math. Phys.* **355** (2017), 373–426.

85. M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr and M. Tomamichel, On quantum Rényi entropies: A new generalization and some properties, *J. Math. Phys.* **54** (2013), 122203.

86. E. Nelson, Notes on non-commutative integration, *J. Funct. Anal.* **15** (1974), 103–116.

87. M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2000.

88. M. Ohya and D. Petz, *Quantum entropy and its use*, Springer-Verlag, Berlin, 1993; 2nd ed., 2004.

89. G. K. Pedersen, C^*-*Algebras and Their Automorphism Groups*, London Mathematical Society Monographs, Vol. 14. Academic Press, London-New York, 1979.

90. D. Petz, Quasi-entropies for states of a von Neumann algebra, *Publ. Res. Inst. Math. Sci.* **21** (1985), 787–800.

91. D. Petz, Quasi-entropies for finite quantum systems, *Rep. Math. Phys.* **23** (1986), 57–65.

92. D. Petz, Sufficient subalgebras and the relative entropy of states of a von Neumann algebra, *Comm. Math. Phys.* **105** (1986), 123–131.

93. D. Petz, Sufficiency of channels over von Neumann algebras, *Quart. J. Math. Oxford Ser. (2)* **39** (1988), 97–108.

94. D. Petz, A variational expression for the relative entropy, *Comm. Math. Phys.* **114** (1988), 345–349.

95. D. Petz, Discrimination between states of a quantum system by observations, *J. Funct. Anal.* **120** (1994), 82–97.

96. D. Petz, *Quantum Information Theory and Quantum Statistics*, Theoretical and Mathematical Physics. Springer-Verlag, Berlin, 2008.

97. D. Petz and M. B. Ruskai, Contraction of generalized relative entropy under stochastic mappings on matrices, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **1** (1998), 83–89.

98. G. A. Raggio, Comparison of Uhlmann's transition probability with the one induced by the natural positive cone of a von Neumann algebra in standard form, *Lett. Math. Phys.* **6** (1982), 223–236.

99. M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis*, 2nd ed., Academic Press, New York, 1980.

100. R. Remmert, *Classical Topics in Complex Function Theory*, Translated from the German, Graduate Texts in Mathematics, Vol. 172. Springer-Verlag, New York, 1998.

101. S. Sakai, *C*-algebras and W*-algebras*, Reprint of the 1971 edition, Classics in Mathematics. Springer-Verlag, Berlin, 1998.

102. I. E. Segal, A non-commutative extension of abstract integration, *Ann. of Math. (2)* **57** (1953), 401–457; Correction to "A non-commutative extension of abstract integration", **58** (1953), 595–596.

103. S. Strătilă, *Modular Theory in Operator Algebras*, Editura Academiei and Abacus Press, Tunbridge Wells, 1981.

104. S. Strătilă and L. Zsidó, *Lectures on Von Neumann Algebras*, Editura Academiei and Abacus Press, Tunbridge Wells, 1979.

105. M. Takesaki, *Tomita's Theory of Modular Hilbert Algebras and Its Applications*, Lecture Notes in Math., Vol. 128, Springer-Verlag, Berlin-New York, 1970.

106. M. Takesaki, Conditional expectations in von Neumann algebras, *J. Funct. Anal.* **9** (1972), 306–321.

107. M. Takesaki, Duality for crossed products and the structure of von Neumann algebras of type III, *Acta Math.* **131** (1973), 249–310.

108. M. Takesaki, *Theory of operator algebras. I.* Springer-Verlag, New York-Heidelberg, 1979.

109. M. Takesaki, *Theory of Operator Algebras. II*, Encyclopaedia of Mathematical Sciences, Vol. 125, Springer-Verlag, Berlin, 2003.

110. M. Terp, L^P spaces associated with von Neumann algebras, Notes, Copenhagen Univ., 1981.

111. M. Tomamichel, *Quantum Information Processing with Finite Resources. Mathematical Foundations*, SpringerBriefs in Mathematical Physics, Vol. 5, Springer, Cham, 2016.

112. J. Tomiyama, On the projection of norm one in W^*-algebras, *Proc. Japan Acad.* **33** (1957), 608–612.

113. H. Umegaki, Conditional expectation in an operator algebra, *Tôhoku Math. J. (2)* **6** (1954), 177–181.

114. H. Umegaki, Conditional expectation in an operator algebra, IV (entropy and information), *Kōdai Math. Sem. Rep.* **14** (1962), 59–85.

115. M. M. Wilde, *Quantum Information Theory*, Cambridge University Press, Cambridge, 2013; 2nd ed., 2017.

116. M. M. Wilde, A. Winter and D. Yang, Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative Entropy, *Comm. Math. Phys.* **331** (2014), 593–622.

Index