ON ENTROPY FOR GENERAL QUANTUM SYSTEMS

W. A. MAJEWSKI AND L. E. LABUSCHAGNE

Abstract. In these notes we will give an overview and road map for a definition and characterization of (relative) entropy for both classical and quantum systems. In other words, we will provide a consistent treatment of entropy which can be applied within the recently developed Orlicz space based approach to large systems. This means that the proposed approach successfully provides a refined framework for the treatment of entropy in each of classical statistical physics, Dirac’s formalism of Quantum Mechanics, large systems of quantum statistical physics, and finally also for Quantum Field Theory.

Despite the efforts of many authors over a very long period of time, gaining a deeper understanding of entropy remains one of the most important and intriguing challenges in the physics of large systems - a challenge still receiving the close attention of many prominent authors. See for example [30]. In this endeavour the techniques available for the quantum framework still lack the refinement of those available for classical systems. On this point, Dirac’s formalism for Quantum Mechanics and von Neumann’s definition of entropy in the context of $B(H)$, does however provide a “template” for developing techniques for the description and study of entropy in the context of tracial von Neumann algebras. One possible way in which von Neumann’s ideas could be refined to provide a “good” description of states with well-defined entropy in the tracial case, was fully described in [19]. As is shown in that paper, a successful description of states with entropy can be achieved on condition that the more common framework for quantum theory based on the pair of spaces $(L^\infty, L^1)$, is replaced with a formalism based on the pair of Orlicz spaces $(L^{\cosh^{-1}}, L^{\log(L+1)})$. An important point worth noting (also pointed out in [19]), is that this axiomatic shift leaves Dirac’s formalism intact! However not all quantum systems correspond to tracial von Neumann algebras. (Note for example that the local algebras of Quantum Field Theory are type $III_1$.) Hence no formalism for describing and studying entropy is complete, if it cannot also find expression in a type III context. In these notes we will provide a formalism for describing (relative) entropy for the most general quantum systems. Our approach is to describe relative entropy in terms of modular dynamics, for which a common input stemming from the concept of Radon-Nikodym derivatives is crucial. As we shall show in section 4, the theory achieved dovetails well with existing concepts of relative entropy [1, 2], and also allows for a “density based” description which

Date: April 17, 2018.

2010 Mathematics Subject Classification. 46L51, 28D20, 46E30, 47L90, and 82B30.

Key words and phrases. noncommutative integration, entropy, Orlicz spaces, operator algebras, statistical thermodynamics.

The contribution of L. E. Labuschagne is based on research partially supported by the National Research Foundation (IPRR Grant 96128). Any opinion, findings and conclusions or recommendations expressed in this material, are those of the author, and therefore the NRF do not accept any liability in regard thereto.
faithfully mimics the classical formula. Then in section 5 we use the theory thus developed as a guideline for introducing a concept of entropy for single states of type III von Neumann, before concluding by indicating the way forward. As shall be seen, the definition for entropy achieved in section 5 harmonises perfectly with the above-mentioned Orlicz space formalism, and is a natural extension of the descriptions given in [19]. We emphasize that the aforementioned extension demands a regularization procedure which has recently been shown to fit the operator algebraic approach to Quantum Field Theory very naturally, cf [18].

1. Boltzmann’s H-functional and (classical) entropy

Let Γ be a phase space associated with a system. We fix a reference measure λ on Γ, usually it will be the Lebesgue measure. A function $f$ such that $f \in \{g; g \in L^1(\Gamma, d\lambda), g \geq 0, \int_\Gamma g d\lambda = 1\} \equiv \mathcal{G}$ defines a probability measure $d\mu = f d\lambda$. In the Boltzmann theory, such an $f$ has the interpretation of velocity distribution function, cf [29] see also [25]. On the other hand we note that $g \in \mathcal{G}$ can be written as

$$g = \frac{d\mu}{d\lambda}$$

where $\frac{d\mu}{d\lambda}$ stands for the Radon-Nikodym derivative. Hence, the Boltzmann $H$-functional can be written as

$$H(g) \equiv \int g \log(g) d\lambda = \int \frac{d\mu}{d\lambda} \log \left(\frac{d\mu}{d\lambda}\right) d\lambda = \mu \left(\log \left(\frac{d\mu}{d\lambda}\right)\right),$$

provided that the above integrals exist. In [17], [19], [20] we have argued that for states (probability measures) $\mu$ such that $\frac{d\mu}{d\lambda} \in L \log(L + 1) \cap L^1$, the functional $H(\cdot)$ is well defined.

Remark 1.1. As the (classical) continuous entropy $S$ differs from the functional $H$ only by sign, the above means that the entropy $S(\frac{d\mu}{d\lambda})$ is well defined if $\frac{d\mu}{d\lambda} \in L \log(L + 1) \cap L^1$.

Let $\mu$ and $\nu$ be probability measures over a set $X$, and assume that $\mu$ is absolutely continuous with respect to $\nu$. The relative entropy (also known as Kullback-Leibner divergence) is defined as

$$S(\mu|\nu) = \int_X \log \left(\frac{d\mu}{d\nu}\right) d\mu = \int_X \frac{d\mu}{d\nu} \log \left(\frac{d\mu}{d\nu}\right) d\nu \equiv \left\langle \log \frac{d\mu}{d\nu} \right\rangle_\mu,$$

provided that the integrals in the above formulas exist, where $\frac{d\mu}{d\nu}$ is the Radon-Nikodym derivative of $\mu$ with respect to $\nu$. Assume additionally that $\nu$ (so also $\mu$ is absolutely continuous with respect to the reference measure $\lambda$. Then

$$\frac{d\mu}{d\nu} = \frac{d\mu}{d\lambda} \cdot \frac{d\lambda}{d\nu},$$

and under some additional assumptions one has the more familiar formula for the relative entropy

$$S(\mu|\nu) = \int_X p \log \frac{p}{q} d\lambda,$$

where $p = \frac{d\mu}{d\lambda}$ and $q = \frac{d\nu}{d\lambda}$.
Intuitively, it is easily seen that for a discrete case, the entropy of a random variable $f$ on $X$ with a probability distribution $p(x)$ is related to how much $p(x)$ diverges from the uniform distribution on the support of $f$. In particular, putting $q = 1$ in the formula 1.5 one gets
\begin{equation}
S(\mu\|\tau) = H(p),
\end{equation}
where the (non-normalized) functional $\tau$ is defined by the reference measure $\lambda$. As “uniformity” can be related to the “most” chaotic state (each microstate is equally probable), the basic property of statistical entropy expressing how far the given state is from the most chaotic, is recovered.

To clarify this point as well as to gain some intuition for a noncommutative generalization, we turn to the algebraic approach to the just defined concepts. For a fixed measure space $(X, \Sigma, \lambda)$, let $L^\infty(X, \Sigma, \lambda) \equiv L^\infty$ denote the set of all $\lambda$-measurable, essentially bounded functions on $X$. The absolute continuity of $\mu$ with respect to $\lambda$ is equivalent to the condition that $\mu$ can be regarded as a normal functional on $L^\infty(X, \Sigma, \lambda)$, cf [4] Theorem 1 , p. 167. Since $L^\infty(X, \Sigma, \lambda)$ is the prototype of abelian von Neumann algebras, one can rewrite definitions as well as the basic properties of the above concepts in (abelian) von Neumann algebraic terms.

To this end, let $\vartheta_\mu(f) = \int_X f \cdot (\frac{d\mu}{d\lambda}) d\lambda(x)$ denote the functional over $L^\infty(X, \Sigma, \lambda)$, for the reference measure $\lambda$. In particular, the trace $\tau$ over $L^\infty(X, \Sigma, \lambda)$ is given by
\begin{equation}
\tau(f) = \int_X f d\lambda(x).
\end{equation}

It is worth pointing out that the existence of such a trace affords the possibility of dealing with uniform distribution (as was indicated above). In other words, such existence affords the possibility of discussing the relation between entropy and relative entropy! Consequently, the entropy formula can be given as
\begin{equation}
S(\mu) \equiv S(\vartheta_\mu) = \tau \left( \left( \frac{D\vartheta_\mu}{D\tau} \right) \log \left( \frac{D\vartheta_\mu}{D\tau} \right) \right) \equiv \int_X \left( \frac{D\vartheta_\mu}{D\tau} \right) \log \left( \frac{D\vartheta_\mu}{D\tau} \right) d\lambda(x)
= \left\langle \log \left( \frac{D\vartheta_\mu}{D\tau} \right) \right\rangle_\mu,
\end{equation}
while the relative entropy formula reads
\begin{equation}
S(\vartheta_\mu | \vartheta_\nu) = \left\langle \log \left( \frac{D\vartheta_\mu}{D\vartheta_\nu} \right) \right\rangle_\mu,
\end{equation}
where $\frac{D\vartheta_\mu}{D\vartheta_\nu}$ stands for the Radon-Nikodym derivative of functional $\vartheta_\mu$ with respect to the functional $\vartheta_\nu$, see the next section.

Remark 1.2. Classical equilibrium thermodynamics. To get some better intuition, let us consider specific case, when the velocity distribution function $\frac{d\mu}{d\lambda}$ is given by Maxwell-Boltzmann distribution
\begin{equation}
\frac{d\mu}{d\lambda} = Ze^{-\beta H} = e^{\log Z - \beta H} \equiv e^K,
\end{equation}
where $Z$ is the normalization constant, $\beta > 0$ (usually interpreted as “the inverse temperature”), and $H$ is the Hamiltonian of the system under consideration.
For such cases, the above formulas for entropies read
\begin{equation}
S\left( \frac{d\mu}{d\lambda} \right) \equiv S\left( \frac{D\mu}{D\tau} \right) = \langle \log(e^K) \rangle_\mu = \langle K \rangle_\mu,
\end{equation}
while for the relative entropy of \( \frac{d\mu}{d\lambda} = \frac{Z_1 e^{-\beta_1 H_1}}{Z_2 e^{-\beta_2 H_2}} \equiv e^{K_1}, \frac{d\nu}{d\lambda} = \frac{Z_2 e^{-\beta_2 H_2}}{Z_2 e^{-\beta_2 H_2}} \equiv e^{K_2} \), one has
\begin{equation}
S\left( \frac{d\mu}{d\lambda} | \frac{d\nu}{d\lambda} \right) = \langle \log \frac{e^{K_1}}{e^{K_2}} \rangle_\mu = \langle K_1 \rangle_\mu - \langle K_2 \rangle_\mu.
\end{equation}

It is important to note that (1.11) is in perfect agreement with the second law of thermodynamics; see section 32 in [14]. The above formulas can be rewritten as
\begin{equation}
S\left( \frac{d\mu}{d\lambda} \right) = -i \frac{d}{dt} \langle e^{itK} \rangle_\mu |_{t=0}
\end{equation}
and
\begin{equation}
S\left( \frac{d\mu}{d\lambda} | \frac{d\nu}{d\lambda} \right) = -i \frac{d}{dt} \langle e^{itK_1} e^{-itK_2} \rangle_\mu |_{t=0}.
\end{equation}

As it will be seen in the next Sections, the above recipe can easily be generalized and quantized.

To clarify the significance of derivatives and to proceed with our exposition we need some preliminaries, which for the reader’s convenience will be given in a separate section.

## 2. Algebraic preliminaries

As the concepts of entropy and relative entropy involve Radon-Nikodym derivatives, for the reader convenience, we here provide the relevant material on non-commutative Radon-Nikodym and cocycle derivatives, thus making our exposition self-contained. The theory of such cocycles goes back to [8], [9], [10]. In particular, Connes proved, see [8]

**Theorem 2.1.** Let \( \mathcal{M} \) be a von Neumann algebra and \( \phi, \psi \) faithful semifinite normal weights on \( \mathcal{M} \). Then there exists a \( \sigma \)-strongly continuous one parameter family \( \{u_t\} \) of unitaries in \( \mathcal{M} \) with the following properties:
- \( u_{t+t'} = u_t \sigma^\phi_t(u_{t'}) \), for all \( t, t' \in \mathbb{R} \),
- \( \sigma_t^\psi(x) = u_t \sigma^\phi_t(x) u_t^* \), for all \( x \in \mathcal{M}, t \in \mathbb{R} \),
- A unitary \( u \in \mathcal{M} \) satisfies \( \psi(x) = \phi(uau^*) \) for all \( x \in \mathcal{M} \), if and only if \( u_t = u^* \sigma^\phi_t(u) \) for all \( t \in \mathbb{R} \),

where \( \sigma_t^\phi (\sigma^\psi_t) \) stands for the modular evolution determined by \( \phi \) (\( \psi \) respectively).

**Definition 2.2.** The family of unitaries described by the above theorem is called the cocycle derivative of \( \phi \) with respect to \( \psi \) and is denoted by
\begin{equation}
(D \phi : D \psi)_t = u_t.
\end{equation}

To understand fully the next remark we need, cf [7], Theorem 5.3.10.

**Theorem 2.3.** (Takesaki) Let \( \mathcal{M} \) be a von Neumann algebra, and \( \omega \) a normal state on \( \mathcal{M} \). The following are equivalent:
(1) $\omega$ is a faithful as a state on $\pi_\omega(M)$, i.e. there exists a projector $E \in M \cap M'$ such that $\omega(I - E) = 0$ and $\omega|_{ME}$ is a faithful state.

(2) There exists a $\sigma$-weakly continuous one-parameter group $\sigma$ of $^*$-automorphisms of $M$ such that $\omega$ is $\sigma$-KMS state.

Moreover, $\sigma|_{ME}$ is the modular group of $M_E$ associated with $\omega$.

This theorem legitimises the application of KMS theory to our approach to quantum entropy. In other words, our scheme is related to quantum equilibrium thermodynamics. Now we are in position to present

Remark 2.4. We note that $u_0 = I$ (see the proof of Theorem 3.3, Chapter VIII in [27]). Further, let us take (formally) a derivation of (2.1) at $t = 0$. Then, denoting the infinitesimal generator of $\sigma^\psi$ ($\sigma^\varphi$ respectively) one gets

$$(2.3) \quad L^\psi(x) = \left. \frac{du_t}{dt} \right|_{t=0} x + L^\varphi(x) + x \left( \left. \frac{du_t}{dt} \right|_{t=0} \right)^*,$$

or equivalently

$$(2.4) \quad L^\psi(x) - L^\varphi(x) = \left. \frac{du_t}{dt} \right|_{t=0} x + x \left( \left. \frac{du_t}{dt} \right|_{t=0} \right)^*.$$  

Theorem (2.3) implies that the modular evolution for a fixed faithful normal state $\varphi$ on $M$, can be interpreted as Hamilton type dynamics for the equilibrium (KMS) state on $M$. This means that the derivative of $u_t$ at $t = 0$ determines the difference of two “equilibrium” type generators $L^\psi$ and $L^\varphi$. The important point to note here is the fact that in general, $L^\psi$ and $L^\varphi$ are unbounded derivations. Thus, the equality (2.3) is not well defined for each $x$. This clearly indicates that derivatives of $u_t$ should be studied carefully, and this will be done in the ensuing sections.

To say more, let $\psi$ be a perturbed $\varphi$-state, so $\psi \equiv \varphi^P$; for all details see section 5.4.1 in [7]. In particular, for $P \in M$ there exists an explicit form of $u_t$. Furthermore, it is easy to note that $L^\varphi x - L^\psi x = i[P,x]$, which is well defined. Consequently, comparing two states which differ by finite an energy perturbation, does not lead to any problem.

Finally, we note that KMS states can be characterized by passivity, see [24] and/or section 5.4.4 in [7]. We remind that among other things passivity ensures compatibility with the second law of thermodynamics. Therefore, our scheme based on Tomita-Takesaki theory, seems to be a natural quantization of the classical case presented in Remark 1.2.

The Radon-Nikodym theorem used in the previous section has generalizations to general von Neumann algebras. The first generalization, for traces, is extracted from Pedersen’s book [23], see Theorem 5.3.11 and remarks in 5.3.12.

**Theorem 2.5.** Let $\tau$ be a normal semifinite trace over $\mathfrak{M}$. For each normal semifinite weight $\psi$ on $\mathfrak{M}$ which is absolutely continuous with respect to $\tau$ in the sense that for any $a \in \mathfrak{M}$, the fact that $\tau(a^*a) = 0$ implies $\psi(a^*a) = 0$, there exists a unique positive operator $h$ on $\mathcal{H}_\tau$ ($\mathcal{H}_\tau$ is GNS space for $(\mathfrak{M},\tau)$) such that

$$(2.5) \quad \psi(x) = \tau(hx).$$

For a general von Neumann algebra $\mathfrak{M}$ and two normal faithful semifinite weights such that one dominates the other one has (see Theorem VIII.3.17 in [27])
Theorem 2.6. For a pair \( \vartheta, \psi \) of faithful semifinite normal weights on \( M \), the following conditions are equivalent:

1. There exists \( M > 0 \) such that
   \[
   \vartheta(x) \leq M \psi(x), \quad x \in M_+,
   \]

2. The cocycle derivative \( (D\vartheta : D\vartheta)_t \equiv u_t \) can be extended to an \( M \)-valued \( \sigma \)-weakly continuous bounded function on the horizontal strip \( D_{1/2} = \{ z \in \mathbb{C}; -1/2 \leq \text{Im}(z) \leq 0 \} \), which is holomorphic in the interior of the strip.

If these conditions hold, then

\[
\vartheta(x) = \psi(u^*_{-1/2} xu_{-1/2}), \quad x \in \left\{ \sum_{i=1}^{n} y_i^* x_i; \quad x_i, y_j \in n_\varphi \right\},
\]

where \( n_\varphi = \{ x \in M; \psi(x^* x) < \infty \} \).

Remark 2.7. We emphasize that a domination of one weight by another is a stronger property than “absolute continuity” described in Theorem 2.5, but the domination condition is in the same vein as the condition of absolute continuity. Also notice from part (2) of the above theorem, that \( |u^*_{-1/2}|^2 \) in a very real sense fulfills the role of the “density” of \( \psi \) with respect to \( \vartheta \).

One may ask whether there is a relation, based on the Connes characterization of unitary Radon-Nikodym cocycles, between cocycle derivatives and the relative modular operator. More precisely, see [1], [3], let \( \varphi, \vartheta \) be normal semifinite weights on \( M \), and \( \varphi \) be faithful. Then

\[
u_t \equiv (D\vartheta : D\vartheta)_t = \Delta^{it}_\varphi \Delta^{-it}_\varphi.
\]

In particular, if \( M \) is semifinite von Neumann algebra, \( \psi \) and \( \vartheta \) faithful semifinite normal weights, \( \tau \) a faithful, normal semifinite trace on \( M \), then one has (see [1], p.470) that there exist positive operators affiliated with \( M \) such that \( \psi(x) = \tau(\varphi, x) \), \( \vartheta(x) = \tau(\varphi, x) \) for each \( x \in M \), and

\[
(D\vartheta : D\psi)_t = \varphi^{it} \varphi^{-it}.
\]

Hence, on applying this equality to the abelian von Neumann algebra \( L^\infty \) (cf. the discussion at the end of the previous section), one has

\[
\frac{d}{dt} \left( D\varphi_\mu : D\varphi_\nu \right)_{t=0} = i \log f_\mu - i \log f_\nu,
\]

where \( \mu = f_\mu d\lambda, \nu = f_\nu d\lambda \), and \( f_\mu > 0, f_\nu > 0 \). Thus

\[
- i \left( \frac{d}{dt} \left( D\varphi_\mu : D\varphi_\nu \right)_{t=0} \right)_\mu = \int f_\mu \log f_\mu - f_\nu \log f_\nu d\lambda = \int \left( f_\mu \log f_\mu - f_\nu \log f_\nu \right) d\lambda,
\]

which is in perfect agreement with the definition of the relative entropy, cf. formula

We remind the reader that the proper basic structure for a description of large quantum systems, is a von Neumann algebra of type III. In other words, one is forced to deal with a von Neumann algebra which is not equipped with a nontrivial trace. Consequently, to be able to study entropy, access to the type of functional
calculus required for an effective description of uniform distribution would be a powerful tool, which can be accessed by passing to a larger super-algebra, i.e. to the crossed-product \( \mathcal{M} \). It is in this larger super-algebra that we have access to the functional calculus for \( \tau \)-measurable operators. If \( \mathfrak{M} \) together with a canonical faithful normal semifinite weight \( \omega \) is given on a Hilbert space \( \mathcal{H} \), then \( \mathcal{M} \) is the von Neumann algebra on the Hilbert space \( L^2(\mathcal{H}) \) generated by the following operators:

\[
\tag{2.12}
(\pi(x)\xi)(t) = \sigma_{-1}(x)\xi(t), \xi \in L^2(\mathcal{H}), t \in \mathbb{R}, x \in \mathfrak{M},
\]

\[
\tag{2.13}
(\lambda(s)\xi)(t) = x(t-s), \xi \in L^2(\mathcal{H}), t \in \mathbb{R}, x \in \mathfrak{M},
\]

where \( \sigma_t = \sigma_{it} \) stands for the modular automorphism.

Remark 2.8  
(1) \( \mathfrak{M} \) can be identified with its image \( \pi(\mathfrak{M}) \) in \( \mathcal{M} \).

(2) If \( \mathfrak{M} \) is type III then \( \mathcal{M} \) is a semifinite. Thus, on \( \mathcal{M} \) there is a semifinite normal faithful trace!

We wish to close this section with a deep result of Haagerup, see [13] Theorem 4.7 or/and [28] pp. 26-27. Let \( \psi, \vartheta \) be normal, faithful semifinite weights on \( \mathfrak{M} \). \( \check{\psi} \) and \( \check{\vartheta} \) stand for the corresponding dual weights on \( \mathcal{M} \). Then, for any \( t \in \mathbb{R} \)

\[
\tag{2.14}
(D\check{\psi}; D\check{\vartheta})_t = (D\psi; D\vartheta)_t.
\]

3. The von Neumann entropy and Dirac’s formalism.

In Dirac’s formalism, a (small) quantum system is described by an infinite dimensional Hilbert space \( \mathcal{H} \) and the von Neumann algebra \( B(\mathcal{H}) \). A normal state \( \psi \) on \( B(\mathcal{H}) \) has the form \( \psi(a) = \text{Tr}_\psi a \) where \( \rho_\psi \) is a positive trace class operator, with trace equal to 1, i.e. \( \text{Tr}(\rho_\psi) = 1 \). Here the set of states \( \mathcal{S} \) is given by \( \mathcal{S} = \{ \varrho \in B(\mathcal{H}); \varrho^* = \varrho, \varrho \geq 0, \text{Tr}(\varrho) = 1 \} \). Applying the non-commutative Radon-Nikodým theorem, see Theorem 1, pp. 469-470 in [11], one has

\[
\tag{3.1}
\varrho_\psi^t = (D\psi; D\text{Tr})_t.
\]

We remind that von Neumann entropy \( S(\varrho_\psi) \) was defined as

\[
\tag{3.2}
S(\varrho_\psi) = \text{Tr}(\varrho_\psi \log \varrho_\psi).
\]

This definition can be rewritten in Radon-Nikodým terms in the following way

\[
\tag{3.3}
S(\varrho_\psi) = -i\text{Tr}\left( \varrho_\psi \frac{d}{dt}(D\psi; D\text{Tr}) \bigg|_{t=0} \right) \equiv -i\psi\left( \frac{d}{dt}(D\psi; D\text{Tr}) \bigg|_{t=0} \right),
\]

and for \( \psi(\cdot) = \text{Tr}(\varrho_\psi \cdot), \varphi(\cdot) = \text{Tr}(\varrho_\varphi \cdot) \)

\[
\tag{3.4}
S(\psi|\varphi) = \text{Tr}(\varrho_\psi \log \varrho_\psi - \varrho_\psi \log \varrho_\varphi) = -i\psi\left( \frac{d}{dt}(D\psi; D\varphi) \bigg|_{t=0} \right),
\]

where we assumed that the states are faithful.

Remark 3.1. As was pointed out at the end of [19] Section 6, within Dirac’s formalism the Orlicz space scheme for selecting “good” states with well defined entropy, agrees with the standard approach to elementary quantum mechanics. Specifically in this setting the space \( L^1 \cap L(\log(L+1))(B(\mathcal{H})) \) is precisely the trace class operators \( L^1(B(\mathcal{H})) \). In fact for \( B(\mathcal{H}) \), all noncommutative measurable operators are already bounded; for details see cf. [19], and [21]. This behaviour is not unexpected as on the one hand Dirac’s formalism is designed for small systems, and on the
other hand, restricting to $B(\mathcal{H})$, noncommutative integration theory is oversimplified. The entropy for large systems will be examined in the next section.

4. General quantum case

Let us consider a general quantum system and let $\mathfrak{M}$ be a von Neumann algebra associated with the system. In general, for large systems, $\mathfrak{M}$ is a type III von Neumann algebra. Let $\omega$ be a normal semifinite faithful weight on $\mathfrak{M}$. The weight $\omega$ will play the role of a noncommutative probability reference measure. We denote by $\mathcal{M}$ the cross product of $\mathfrak{M}$ associated with the modular morphism $\sigma_\omega$ produced by $\omega$, cf. Section 2. By $\tilde{\omega}$ we denote the dual (and hence normal semifinite faithful) weight on $\mathcal{M}$, and $\tau$ stands for the canonical trace on $\mathcal{M}$. We remark that the modular automorphism group $\tilde{\sigma}$ produced by the dual weight $\tilde{\omega}$ has the form $\tilde{\sigma}(\cdot) = \lambda(t) \cdot \lambda(t)^*$ - for details see [20, 22] and [18]. By Stone’s theorem one has $\lambda(t) = h^it$. We note that $h$ can be identified with $-i\frac{\tilde{\omega}}{\omega}(D\omega : D\tau)|_{t=0}$ where $\tau$ is the canonical trace on $\mathcal{M}$. Based on the foregoing analysis we propose:

**Definition 4.1.** Let $\vartheta, \psi$ be faithful states on $\mathfrak{M}$. We define the relative entropy $S(\vartheta|\psi)$ to be $S(\vartheta|\psi) = \lim_{t \to 0} \frac{1}{t}(D\vartheta : D\psi)_t - 1$ if the limit exists, and assign a value of $\infty$ to $S(\vartheta|\psi)$ otherwise.

Let $\mathfrak{M}$ be a $\sigma$-finite von Neumann algebra in standard form described above, and let $\psi$ and $\phi$ be two faithful normal states with unit vector representatives $\Psi, \Phi \in \mathcal{H}$. The basic theory of Tomita-Takesaki theory easily extends to show that the densely defined anti-linear operator $S_{\phi,\psi}(a \Psi) = a^* \Phi$ is in fact closable. The operator $\Delta_{\phi,\psi}$ is then defined to be the modulus of the closure of $S_{\phi,\psi}$. In the same way that the “standard” modular operator may be used to generate the modular automorphism group of a given state, this operator in a very real sense encodes the manner in which the dynamics determined by the modular automorphism group of one state, differs from other. Using this fact, Araki then defined the relative entropy of $\phi$ to be $\varrho = \langle \Psi, \log(\Delta_{\phi,\psi})\Psi \rangle$. We refer the interested reader to [1, 2] and the references therein for a survey of the basic properties of this entropy. However despite the success of Araki’s approach, we prefer the above definition, since on the one hand it is more overtly based on modular dynamics, and on the other it more easily allows for the incorporation of crossed product techniques in the study of this entropy — as we shall subsequently see. The two approaches turn out to be equivalent – a fact which is the content of the next theorem. One of the crucial facts which help to establish this link, is the fact that the Connes cocycle derivative $(D\psi : D\phi)_t$ may be described in terms of $\Delta_{\phi,\psi}$ and $\Delta_\phi$ (see Appendix B of [3] for details). Another is that any normal state $\vartheta$ on a $\sigma$-finite von Neumann algebra $\mathfrak{M}$ in standard form, must have a vector representative [7, Theorem 2.5.31].

**Theorem 4.2.** Let $\mathfrak{M}$ be a $\sigma$-finite von Neumann algebra in the standard form described above, and let $\psi$ and $\phi$ be two faithful normal states with unit vector representatives $\Psi, \Phi \in \mathcal{H}$. Then $S(\psi|\phi)$ as defined in Definition 4.1 agrees exactly with Araki’s definition of relative entropy [1].

**Proof.** To start the proof we recall that in this case $(D\psi : D\phi)_t = \Delta_{\phi,\psi}^it \Delta_{\phi,\psi}^{-it}$. The chain rule for cocycle derivatives informs us that $1 = (D\psi : D\psi)_t = (D\psi : D\phi)_t(D\phi : D\psi)_t$, and hence that we also have that

$$(D\psi : D\phi)_t = (D\phi : D\psi)_t^{-1} = (\Delta_{\phi,\psi}^it \Delta_{\phi,\psi}^{-it})^{-1} = \Delta_{\psi,\phi}^it \Delta_{\psi,\phi}^{-it}.$$
Note that by construction we have that $\Psi$ is an eigenvector of $D(\phi)$ corresponding to the eigenvalue $1$. Hence $\Psi$ must then be an eigenvector of $D(-it)$ corresponding to the eigenvalue $1 - it = 1$. So for any $t$ we must then have that
\[
\langle \Psi, (D(\phi)t) \rangle = \langle \Psi, D(\phi)t \rangle
= \langle D(-it) \Psi, D(\phi) \Psi \rangle
= \langle D(-it) \Psi, D(\phi) \Psi \rangle.
\]
It is therefore trivially clear that
\[
\lim_{t \to 0} \frac{1}{t} \psi((D(\phi)t - 1) = \lim_{t \to 0} \frac{1}{t} \langle \Psi, (D(-it)t - 1) \Psi \rangle.
\]
Recall that Araki’s definition of entropy is
\[
S(\psi|\phi) = -\langle \Psi, \log(D(\phi)|\Psi \rangle
\]
where the latter term is understood to be
\[
- \int_0^\infty \log(\lambda) \, d\langle \Psi, e_\lambda \Psi \rangle
\]
(here $\lambda \to e_\lambda$ is the spectral resolution of $D(\phi)$). As Araki points out, the value of this integral is either real (in the case where log is integrable), or $\infty$ otherwise. In the case where log is not integrable, it once again follows from [1] that this setting log is always integrable on $[1, \infty)$, and hence that the non-integrability of log is derived from the fact that $-\int_0^1 \log(\lambda) \, d\langle \Psi, e_\lambda \Psi \rangle = \infty$.

First suppose that log is integrable. For any $t > 0$ and any $\lambda > 0$, we have that
\[
|\frac{1}{t}(\lambda^{it} - 1)| \leq |\frac{1}{t}(\lambda^{it/2} - 1)(\lambda^{it/2} + 1)| \leq \frac{2}{t}(\lambda^{it/2} - 1).
\]
Carrying on inductively, leads to the conclusion that $|\frac{1}{t}(\lambda^{it} - 1)| \leq \frac{2}{t}(\lambda^{it/2} - 1)$ for any $k \in \mathbb{N}$. If now we let $k \to \infty$, we obtain the inequality $|\frac{1}{t}(\lambda^{it} - 1)| \leq |\log(\lambda)|$, which holds for any $t > 0$ and any $\lambda > 0$. Hence we may apply the dominated convergence theorem to see that for any sequence $\{t_n\}$ converging to 0, we have that
\[
\lim_{n \to \infty} \frac{-i}{t_n} \langle \Psi, (D(-it_n) - 1) \rangle = \lim_{n \to \infty} \frac{-i}{t_n} \int_0^\infty (\lambda^{it_n} - 1) \, d\langle \Psi, e_\lambda \Psi \rangle
= \int_0^\infty \log(\lambda) \, d\langle \Psi, e_\lambda \Psi \rangle.
\]
This fact is enough to enable us to conclude that
\[
\lim_{t \to 0} \frac{-i}{t} \langle \Psi, (D(-it) - 1) \rangle = -\int_0^\infty \log(\lambda) \, d\langle \Psi, e_\lambda \Psi \rangle.
\]
Next suppose that log is not integrable. As was noted earlier, the fact that in this setting log is always integrable on $[1, \infty)$, ensures that non-integrability of log is equivalent to the statement that $\int_0^1 \log(\lambda) \, d\langle \Psi, e_\lambda \Psi \rangle = \int_0^1 \log(\lambda^{-1}) \, d\langle \Psi, e_\lambda \Psi \rangle = \infty$. In fact a slight modification of the above argument shows that for any sequence $\{t_n\}$ decreasing to 0, we then always have that
\[
\lim_{n \to \infty} \frac{-i}{t_n} \int_1^\infty (\lambda^{it_n} - 1) \, d\langle \Psi, e_\lambda \Psi \rangle = \int_1^\infty \log(\lambda) \, d\langle \Psi, e_\lambda \Psi \rangle \in \mathbb{R}.
\]
We therefore need to investigate this type of behaviour on the interval $[0, 1]$. 
For any sequence \( \{t_n\} \) decreasing to 0, we may use Fatou’s lemma to conclude that
\[
\infty = \liminf_{n} \int_{0}^{1} \frac{\sin(t_n \log(\lambda^{-1}))}{t_n} d(\psi, e_{\lambda} \psi).
\]
Since for any \( n \) we have that \( \Re[-i(\lambda^{-it_n} - 1)] = \sin(-t_n \log(\lambda)) = \sin(t_n \log(\lambda^{-1})) \),
this fact ensures that in this case \( \lim_{n \to \infty} \int_{0}^{1} \Re \left[ \frac{-i}{t_n} (\lambda^{-it_n} - 1) \right] d(\psi, e_{\lambda} \psi) \) does not exist as a real number. Hence neither does \( \lim_{t_n \to 0} \Re \left[ \frac{-i}{t_n} (\Delta_{\phi, \psi}^{-it_n} - 1) \psi \right] \) = \( \lim_{n \to \infty} \int_{0}^{\infty} \Re \left[ \frac{-i}{t_n} (\lambda^{-it_n} - 1) \right] d(\psi, e_{\lambda} \psi) \). This suffices to prove the theorem. \( \square \)

**Remark 4.3.** If \( \mathfrak{M} \) is commutative, then \( \mathfrak{M} = L^\infty(X, \mu) \), in which case \( \psi \) and \( \phi \) correspond to positive measures on \( X \). In particular (cf Theorem 2.1) there exists a Radon-Nikodym derivative \( h = \frac{d\phi}{d\mu} \), and \( h^it = (D\psi : D\phi)_t \). Therefore, the definition of classical relative entropy is also stemming from Definition 4.1. Finally, the definition of relative entropy for Dirac’s formalism also follows from Definition 4.1 (cf formula (4.3.1)).

To say more, we are going to invoke some results from the theory of \( L^p \)-spaces associated with von Neumann algebras. We note, cf. [28], Theorem 36, that
\[
(\lambda(\mathfrak{M}), L^2(\mathfrak{M}), J, L^2(\mathfrak{M})_+)
\]
is the standard form of \( \mathfrak{M} \), where the right action \( \lambda(\cdot) \) is defined as \( \lambda(a) a = a, a \in L^2(\mathfrak{M}) \) and \( J \) denotes the conjugate isometric involution \( a \mapsto a^* \) of \( L^2(\mathfrak{M}) \).

Next, we note, cf. [28], Proposition 4, that there is a bijection \( \phi \mapsto h_\phi \) of the set of all normal semifinite weights on \( \mathfrak{M} \) onto the set of all positive selfadjoint operators \( h \) affiliated with \( \mathfrak{M} \), and satisfying \( \theta_s(h) = e^{s}h \) for any \( s \in \mathbb{R} \). As we wish to deal with \( \tau \)-measurable operators we must restrict ourselves to normal functionals on \( \mathfrak{M} \). Then, the mapping \( \phi \mapsto h_\phi \) is an isometry of \( \mathfrak{M}_\tau \) onto \( L^1(\mathfrak{M}) \). Consequently, fixing \( \phi \in \mathfrak{M}_\tau^+ \), one gets \( h_\phi \in L^1(\mathfrak{M})_+ \). In particular, \( h_\phi^2 \in L^2(\mathfrak{M})_+ \), and
\[
(4.2) \quad \phi(x) = \text{tr}(h_\phi^\frac{x}{2} x h_\phi^\frac{x}{2}) = \langle h_\phi^\frac{x}{2}, x h_\phi^\frac{x}{2} \rangle_{L^2(\mathfrak{M})},
\]
where \( \text{tr} \) stands for a linear functional (having the trace property) on \( L^1(\mathfrak{M}) \), see Definition II.13 and Proposition II.21 in [28]. In other words, \( h_\phi^\frac{x}{2} \) is a vector in the natural cone \( L^2(\mathfrak{M})_+ \), and this vector represents the state \( \phi \).

Using the above framework, the proposed definition of entropy may be written as the claim that \( S(\psi|\phi) = \lim_{t \to 0} -\frac{1}{t} \text{tr}(h_\psi^{1/2} [(D\psi : D\phi)_t - 1] h_\psi^{1/2}) \) whenever the limits exists, with \( S(\psi|\phi) = \infty \) otherwise. The next objective in this section, is to show that this definition can very concretely be reformulated in a manner which is a faithful noncommutative analogue of the classical formula presented in Equation 1.5. However for this we will need some additional technology which we now review.

The first factor that suggests that such a formula may well be within reach is the fact that in the above framework we have that
\[
(D\theta : D\psi)_t = h_{\theta \psi}^{-it}
\]
where \( h_{\theta \psi} = \frac{D\theta}{D\psi} \) and \( h_{\psi} = \frac{D\psi}{D\psi} \). To see that the above claim is true, we may use Haagerup’s result and the cocycle chain rule to see that
\[
1 = (D\psi : D\psi)_t = (D\psi : D\tau)_t (D\tau : D\psi)_t.
\]
Equivalently
\[
(D\tau : D\tilde{\psi})_t = (D\tilde{\psi} : D\tau)_t^{-1}.
\]
But from section 2 we know that \((D\tilde{\psi} : D\tau)_t = h_\psi^{it}\). Hence \((D\tau : D\tilde{\psi})_t = h_\psi^{-it}\).
Since also \((D\tilde{\vartheta} : D\tau)_t = \vartheta_\psi^it\), we may once again use Haagerup’s result and the chain rule to see that
\[
(D\vartheta : D\psi)_t = (D\tilde{\vartheta} : D\tilde{\psi})_t = (D\tilde{\vartheta} : D\tau)_t(D\tau : D\tilde{\psi})_t = h_\psi^{it}h_\psi^{-it}.
\]

Another major factor to take into account is that \(tr\) is only defined on \(L^1(\mathfrak{M})\). Thus, to proceed with our objective of developing a noncommutative version of formula (1.5) we must show that in some sense \(h_\vartheta \log h_\vartheta - h_\vartheta \log h_\vartheta\) is in \(L^1\). As we shall see below, this can indeed be achieved in a limiting sense. Following Terp’s arguments, see [23] Lemma II.19, we consider the function
\[
(4.3) \quad S^0 \ni \alpha \mapsto h_\vartheta^\alpha h_\varphi^{1-\alpha} \in L^1(\mathfrak{M}),
\]
where obviously \(h_\vartheta, h_\varphi \in L^1(\mathfrak{M})\), and \(S\) is the closed complex strip \(\{\alpha \in \mathbb{C} : 0 \leq \Re(\alpha) \leq 1\}\) and \(S^0\) stands for the corresponding open strip. Terp’s Lemma II.19 easily adapts to show that the function (4.3) is analytic on \(S^0\). Taking the derivative, in the Banach space language, one gets that
\[
(4.4) \quad \alpha \mapsto h_\vartheta^\alpha \cdot \log h_\vartheta \cdot h_\varphi^{1-\alpha} - h_\vartheta^{\alpha} \cdot \log h_\varphi \cdot h_\varphi^{1-\alpha} \in L^1(\mathfrak{M})
\]
inside \(S^0\). More importantly, the analyticity ensures that this derivative varies continuously with respect to \(\alpha\) in \(L^1\)-norm. A fact which underlies the above very regular behaviour on this strip, is that for any \(0 < s, x^s \log(x)\) is very well behaved continuous function which is \(0\) at \(0\), and for which \(x^s \leq x^s \log(x) \leq x^{s+1}\) whenever \(x \geq e\). This fact can be used to show that for any positive \(\tau\)-measurable operator \(g, g^s \log(g)\) will again be \(\tau\)-measurable.

Using the above formula and letting \(\alpha \to 1\), leads us to the promised noncommutative analogue of formula (1.5). To understand how this is achieved, assume that \(\alpha = s + it\) where \(0 < s < 1\). Then the fact that
\[
h_\vartheta^\alpha h_\varphi^{1-\alpha} = h_\vartheta^s h_\varphi^{it} h_\varphi^{1-s} h_\varphi^{-it}
\]
leads to the conclusion that
\[
-i\frac{d}{dt}(h_\vartheta^s (D\vartheta : D\psi)_t h_\varphi^{1-s}) = -i\frac{d}{dt}(h_\vartheta^s h_\varphi^{it} h_\varphi^{-it} h_\varphi^{1-s}) = \frac{d}{d\alpha} h_\vartheta^\alpha h_\varphi^{1-\alpha} = h_\vartheta^\alpha \cdot \log h_\vartheta \cdot h_\varphi^{1-\alpha} - h_\vartheta^{\alpha} \cdot \log h_\varphi \cdot h_\varphi^{1-\alpha}.
\]
To see this, notice that in computing a limit of the form \(\lim_{\Delta\alpha \to 0} \frac{f(\alpha + \Delta\alpha) - f(\alpha)}{\Delta\alpha}\), we may as well assume that \(\Delta\alpha = i\Delta t\). In particular when computing the derivative anywhere along the line segment \(0 < s < 1, t = 0\), we always have that
\[
-i\frac{d}{dt}(h_\vartheta^s (D\vartheta : D\psi)_t h_\varphi^{1-s}) \bigg|_{t=0} = h_\vartheta^s \cdot \log h_\vartheta \cdot h_\varphi^{1-s} - h_\vartheta^{s} \cdot \log h_\varphi \cdot h_\varphi^{1-s} \in L^1(\mathfrak{M}).
\]
With the groundwork having been done, we are now ready to present the promised result.
Therefore as this construction see (10). In view of this we will simply write $z$ given $\{z \in \mathbb{C} : -\delta \leq \Im(z) \leq 0\}$, and analytic on the open strip $\{z \in \mathbb{C} : -\delta < \Im(z) < 0\}$.

In the above ordering the case $\delta = \frac{1}{2}$ corresponds exactly to Theorem 2.6.

**Theorem 4.5.** Let $\mathfrak{M}$ be a $\sigma$-finite von Neumann algebra in the standard form described above, and let $\vartheta$ and $\phi$ be two faithful normal states with unit vector representatives $h_\vartheta^{1/2}, h_\phi^{1/2} \in L^2(\mathfrak{M})$. If $\phi \leq \vartheta(\delta)$, then $S(\vartheta|\phi)$ is finite if and only if the limit

$$\lim_{s \to 0^+} \frac{1}{t} tr(h_\vartheta^* \cdot \log h_\vartheta \cdot h_\vartheta^{1-s} - h_\phi^* \cdot \log h_\phi \cdot h_\phi^{1-s})$$

exists, in which case they are equal.

**Proof.** First suppose that $S(\vartheta|\phi)$ is finite and let $\epsilon > 0$ be given. This means that $\lim_{t \to 0^+} \frac{1}{t} tr(h_\vartheta^{1/2} | h_\vartheta^{i\epsilon} - 1 | h_\vartheta^{i\epsilon} | h_\vartheta^{1/2})$ exists. Next let $s$ be given with $\frac{1}{2} < s < 1$. So for $t_0 > 0$ small enough, we will have that

- $\frac{1}{t} tr(h_\vartheta^{1/2} | h_\vartheta^{i\epsilon} - 1 | h_\vartheta^{i\epsilon} | h_\vartheta^{1/2})$ is within $\epsilon$ of $S(\vartheta|\phi)$,
- and $\frac{1}{t} h_\vartheta^* [h_\vartheta^{i\epsilon} - 1] h_\vartheta^{1-s}$ is within $\epsilon$ of $-i \frac{t}{\epsilon} (h_\vartheta^* (D\vartheta : D\vartheta) t h_\vartheta^{1-s})$.

Notice that by the properties of the trace functional $tr$ we have that $tr(h_\vartheta^* [h_\vartheta^{i\epsilon} - 1] h_\vartheta^{1-s}) = tr(h_\vartheta^{1/2} [h_\vartheta^{i\epsilon} - 1] h_\vartheta^{1-s} h_\vartheta^{1/2})$.

By assumption $\vartheta \leq \vartheta(\delta)$. This means that $t \to (D\vartheta : D\vartheta)$ extends to an $\mathfrak{M}$-valued function $f(z)$ which is point to weak*-continuous on the closed strip, $\{z \in \mathbb{C} : -\delta \leq \Im(z) \leq 0\}$, and analytic on the open strip $\{z \in \mathbb{C} : -\delta < \Im(z) < 0\}$. For each $z$ the value $f(z)$ is essentially just an extension of $h_\vartheta^* h_\vartheta^{i\epsilon}$. (For details of this construction see [13].) In view of this we will simply write $[h_\vartheta^* h_\vartheta^{i\epsilon}]$ for $f(z)$. So if we set $z = ir$ where $0 \leq r \leq \delta$, we obtain that as $r \to 0^+$ we will have that $[h_\vartheta^* h_\vartheta^{i\epsilon}]$ converges to 1 in the weak* topology on $\mathfrak{M}$.

Next notice that for $0 \leq 1 - s \leq \delta$, we have that $h_\vartheta^{1-s} h_\vartheta^{s-1/2} = [h_\vartheta^{1-s} h_\vartheta^{1-s}] h_\vartheta^{1/2}$. Therefore as $s \nearrow 1$ on the interval $[1 - \delta, 1]$, we must have that $\frac{1}{t} tr(h_\vartheta^{1/2} [h_\vartheta^{i\epsilon} - 1] h_\vartheta^{1-s} h_\vartheta^{1/2}) = \frac{1}{t} tr(h_\vartheta^* [h_\vartheta^{i\epsilon} - 1] [h_\vartheta^{1-s} h_\vartheta^{1-s}])$ converges to $\frac{1}{t} tr(h_\vartheta^* [h_\vartheta^{i\epsilon} - 1] [h_\vartheta^{1-s} h_\vartheta^{1-s}]) = \frac{1}{t} tr(h_\vartheta^{1/2} [h_\vartheta^{i\epsilon} - 1] h_\vartheta^{1/2})$.

There must therefore exist a $\delta > 0$ such that for any $s$ with $1 - \delta < s < 1$, the term $\frac{1}{t} tr(h_\vartheta^{1/2} [h_\vartheta^{i\epsilon} - 1] h_\vartheta^{1-s} h_\vartheta^{1/2})$ will be within $\epsilon$ of $\frac{1}{t} tr(h_\vartheta^{1/2} [h_\vartheta^{i\epsilon} - 1] h_\vartheta^{1/2})$. If we combine all the above observations, it follows that for any $s$ with $1 - \delta < s < 1$, $tr(h_\vartheta^* \cdot \log h_\vartheta \cdot h_\vartheta^{1-s} - h_\phi^* \cdot \log h_\phi \cdot h_\phi^{1-s})$ will be within $3\epsilon$ of $S(\vartheta|\phi)$. This proves the “only if” part of the theorem.

Next suppose that $S(\vartheta|\phi) = \infty$. From the proof of Theorem 4.2 is is clear that given $M > 0$ we may in this case find some $t_0 > 0$ such that $|\frac{1}{t_0} tr(h_\vartheta^{1/2} [h_\vartheta^{i\epsilon} - 1] h_\vartheta^{1/2})| \geq M$ with additionally (as before) $\frac{1}{t} h_\vartheta^* [h_\vartheta^{i\epsilon} - 1] h_\vartheta^{1-s}$ within $\epsilon$ of
for which we have that \( (h^s_\phi (D\vartheta : D\phi)_{1-s}) \) with respect to \( L^1 \)-norm. The constant \( \delta \) is selected as in the first part of the proof. Combining these estimates, now leads to the conclusion that \( |\text{tr}(h^s_\phi \cdot \log h_\vartheta \cdot h^{1-s}_\phi - h^s_\phi \cdot \log h_\phi \cdot h^{1-s}_\phi)| \geq M - 2\epsilon \) for all \( s \) with \( 1 - \delta < s < 1 \). Since both \( M > 0 \) and \( \epsilon > 0 \) were arbitrary, the limit \( \lim_{s \to 1} \text{tr}(h^s_\phi \cdot \log h_\vartheta \cdot h^{1-s}_\phi - h^s_\phi \cdot \log h_\phi \cdot h^{1-s}_\phi) \) can then clearly not exist. The theorem therefore follows. \( \square \)

5. AN ALTERNATIVE APPROACH TO THE GENERAL QUANTUM CASE

Here we propose a means for defining the entropy of a single state \( \vartheta \). This definition turns out to be exactly equivalent to von Neumann entropy in the tracial case. Some careful preparation for and justification of this definition is required. As a first step in identifying a suitable prescription for defining entropy of a single state, we first take some time to see what Theorem 4.5 looks like when the states are commuting affiliated operators. Let \( \vartheta \) be an Orlicz function and let \( \varphi_\vartheta \) be the fundamental function of \( L^\vartheta (\mathbb{R}) \) equipped with the Luxemburg norm. Then

\[
\chi_{(1,\infty)}(a \varphi_\vartheta(b)) = \chi_{(1,\infty)}(\varphi_\vartheta(a)b).
\]

**Proof.** Let \( \alpha, \beta > 0 \) be given. It is a known fact that \( \alpha \Psi(\beta) \leq 1 \iff \beta \leq \Psi^{-1}(\frac{1}{\alpha}) \). If we apply this fact to the Borel functional calculus for the commuting positive operators \( a \) and \( b \), we have that \( \chi_{(1,\infty)}(a \varphi_\vartheta(b)) = \chi_{(1,\infty)}(\varphi_\vartheta(a)b) \) as required. \( \square \)

The above lemma now enables us to make the following conclusion:

**Proposition 5.2.** Let \( \vartheta, \phi \) be faithful normal states on \( \mathfrak{M} \) with unit vector representatives \( h^{1/2}_\vartheta, h^{1/2}_\phi \), which commute in the sense that they satisfy one (and therefore all) of the criteria described in [27] Corollary VIII.3.6. Assume in addition that \( \phi \leq \vartheta(\delta) \) for some \( \delta > 0 \). With \( \varphi_\text{log} \) denoting the fundamental function of the space \( L \log(L + 1)(\mathbb{R}) \) (equipped with the Luxemburg norm), we then have that

- \( h_\vartheta \) and \( h_\phi \) are commuting operators affiliated to \( \mathfrak{M} \),
- \( f = h_\vartheta h_\phi^{-1} \) extends uniquely to an element of \( \mathfrak{M} \),
- and \( S(\vartheta, \phi) = \varphi(f \log(f)) = \inf_{\epsilon > 0} \epsilon \tau(\chi_{(1,\infty)}(\varphi_{\text{log}}(h_\vartheta)f) + \log(\epsilon)||h_\vartheta f||_1) \).

**Proof.** The first step is to show that \( h_\vartheta \) and \( h_\phi \) commute.

It is clear from the proof of [27] Corollary VIII.3.6, that the commutation of \( \vartheta \) and \( \phi \), ensures the existence of an operator \( h \) affiliated to the centraliser \( \mathfrak{M}_\phi \) of \( \phi \) for which we have that \( (D\phi : D\vartheta)_{t} = h^{it} \). But from the discussion preceding Theorem 4.5 we know that \( (D\vartheta : D\phi)_{t} = (D\vartheta : D\phi)_{t} = h^{it}h^{-it}_\phi \). In other words for each \( t \), \( h^{it} = h^{it}_\phi h^{-it}_\phi \).

On appealing to the properties of the cocycle derivative, we may now conclude that

\[
\begin{align*}
h^{it(1+s)} &= (D\vartheta : D\phi)_{1+t+s} \\
&= (D\vartheta : D\phi)_{t} \sigma_{\phi}^t(1(D\vartheta : D\phi)_{t}) \\
&= h^{is}h^{it}_\phi h^{-is}_\phi
\end{align*}
\]
or equivalently, \( h^{it} = h_\phi^{it} h^{it}_\phi \). So each \( h^{it} \) commutes with each \( h_\phi^{it} \).

But we saw earlier that \( h^{it} = h_\phi^{it} h^{it}_\phi \), or equivalently that \( h_\phi^{it} = h^{it} h_\phi^{it} \). Together these two facts ensure that each \( h_\phi^{it} \) commutes with each \( h_\phi^{it} \). We may now use the Borel functional calculus to conclude from these two facts that \( h_\phi \) and \( h_\phi \) themselves also commute. This proves the first bullet.

To see the second bullet, we note from the proof of Theorem 4.5 that the requirement that \( \phi \leq \theta(\delta) \), ensures that for \( r > 0 \) small enough, \( h_\phi^{it} h_\phi^{it-1} \) extends uniquely to an element of \( \mathfrak{M} \). Since \( h_\phi \) and \( h_\phi \) commute, this clearly ensures that the closure of \((h_\phi^{it-1} h_\phi^{it})^{1/r} = h_\phi h_\phi^{-1} \) also belongs to \( \mathfrak{M} \).

For the final bullet, note that by the Borel functional calculus, the commutation of \( h_\phi \) and \( h_\phi \), ensures that we may write the limit formula \( \lim_{s \to 1} tr(h_\phi^{i} \cdot \log h_\phi \cdot h_\phi^{i-1} - h_\phi^{i} \cdot \log h_\phi \cdot h_\phi^{i-1} \cdot s) \) as \( \lim_{s \to 1} tr(f^s \log(f) h_\phi) \) where \( f = h_\phi h_\phi^{-1} \). It also follows from the proof of Theorem 4.5 that there exists an interval \([\delta, 1] \) for which \( s \to f^s \) is point-weak* continuous. So given \( \rho \) with \( \delta < \rho < 1 \), we may write the limit formula as \( \lim_{s \to \rho} \rho tr(f^s f^{(1 - \rho)} \log(f) h_\phi) \). We may now use the continuous functional calculus to see that since \( f \in \mathfrak{M} \), we must have that \( f^{(1 - \rho)} \log(f) \in \mathfrak{M} \). But then \( [f^{(1 - \rho)} \log(f)] h_\phi \in L^1(\mathfrak{M}) \). The point-weak* continuity of the map \( r \to f^r \) on \([\delta, \rho] \), now ensures that \( \lim_{r \to \rho} \rho tr(f^r f^{(1 - \rho)} \log(f) h_\phi) = tr(f \log(f) h_\phi) = \phi(f \log(f)) \).

To prove the final equality, one firstly uses a similar argument to the one in [19, Proposition 6.8] to see that \( tr(h_\phi f \log(f)) = \inf_{r > 0} \rho tr(h_\phi f (f/\epsilon) \log((f/\epsilon) + 1)) + \log(\epsilon) tr(h_\phi f) \). On combining the preceding Lemma with [28, Lemma II.5 & Def II.13], we then have that

\[
tr(h_\phi (f/\epsilon) \log((f/\epsilon) + 1)) = \tau(\chi_{(1, \infty)}(h_\phi (f/\epsilon) \log((f/\epsilon) + 1)) = \tau(\chi_{(1, \infty)}(\varphi_{\log(h_\phi)} f/\epsilon)) = \tau(\chi_{(\epsilon, \infty)}(\varphi_{\log(h_\phi)} f)).
\]

This proves the final claim.

We are now finally ready to present the definition of the entropy \( S(\vartheta) \) of a faithful normal state \( \vartheta \). The basic idea is to use the above result as guide, for the kind of technical prescription that might work. Tempting as it may be to simply replace \( f = h_\phi h_\phi^{-1} \) with \( h_\phi \), and \( \phi \) with \( tr \), to obtain \( tr(h_\phi \log(h_\phi)) \) as a definition, this cannot possibly work. The problem with this prescription is that \( tr \) is only defined on \( L^1(\mathfrak{M}) \) where in the crossed product setting, all the elements \( h \) of \( L^1(\mathfrak{M}) \) have to satisfy the requirement that \( \theta_s(h) = e^{-sh} \) for all \( s \). Since \( h_\phi \in L^1(\mathfrak{M}) \) we do have that \( \theta_s(h_\phi) = e^{-sh_\phi} \). But then \( \theta_s(h_\phi \log(h_\phi)) = e^{-sh_\phi \log(e^{-sh_\phi})} \). However the final equality in the third bullet of Proposition 5.2 does present us with a means for overcoming this difficulty for a subspace of \( L^1(\mathfrak{M}) \). The subspace in question is the noncommutative Orlicz space \( L^1 \cap L \log(L + 1)(\mathfrak{M}) \).

Some analysis is necessary before we are able to present the definition. Note that classically \( L^1 \cap L \log(L + 1) \) is an Orlicz space produced by the Young’s function

\[
\Psi_{\text{end}}(t) = \max(t, t \log(t + 1)) = \begin{cases} 
    t & 0 \leq t \leq e - 1 \\
    t \log(t + 1) & e - 1 \leq t 
\end{cases}
\]

We start by describing how to construct the type III analogue of the space \( L^1 \cap L \log(L + 1) \). We will for simplicity of computation assume that each of \( L \log(L + 1)(0, \infty) \) and \( L^1 \cap L \log(L + 1)(0, \infty) \) are equipped with the Luxemburg norm. It is then an exercise to see that the fundamental function of \( L^1 \cap L \log(L + 1)(0, \infty) \)
is of the form \( \varphi_{\text{ent}}(t) = \max(t, \varphi_{\text{log}}(t)) \). It is this fundamental function that one uses to construct the type III analogue of \( L^1 \cap L \log(L+1) \) in accordance with the prescriptions given in [16] [18]. Let us for the sake of brevity denote this space by \( L^{\text{ent}}(\mathcal{M}) \). We now show that this space canonically embeds into both \( L^1(\mathcal{M}) \) and \( L \log(L+1)(\mathcal{M}) \).

From the above computations, it is clear that the functions \( \zeta_1(t) = \frac{1}{\varphi_{\text{ent}}(t)} \), and \( \zeta_{\text{log}}(t) = \frac{\varphi_{\text{log}}(t)}{\varphi_{\text{ent}}(t)} \) are both continuous and bounded above (by 1) on \((0, \infty)\). Hence for \( h = \frac{D_{\omega}}{D_{\tau}} \), the operators \( \zeta_1(h) \) and \( \zeta_{\text{log}}(h) \) are both contractive elements of \( \mathcal{M} \).

It is now an exercise to see that the prescriptions \( x \rightarrow \zeta_1(h)^{1/2}x \zeta_1(h)^{1/2} \) and \( x \rightarrow \zeta_{\text{log}}(h)^{1/2}x \zeta_{\text{log}}(h)^{1/2} \) respectively yield continuous embeddings of \( L^{\text{ent}}(\mathcal{M}) \) into \( L^1(\mathcal{M}) \) and \( L \log(L+1)(\mathcal{M}) \). Using these embeddings, we now make the following definition:

**Definition 5.3.** A state \( \vartheta \) on the von Neumann algebra \( \mathcal{M} \) is called *regular* if for some element \( g \) of \( [L \log(L+1) \cap L^1](\mathcal{M})^+ = L^{\text{ent}}(\mathcal{M})^+ \), \( \frac{D_{\omega}}{D_{\tau}} \) is of the form \( \zeta_1(h)^{1/2}g\zeta_1(h)^{1/2} \). For such a regular state we then define the entropy to be

\[
\tilde{S}(\vartheta) = \inf_{\epsilon > 0} [\epsilon \tau(\chi_{(\epsilon, \infty)}(D_{\omega}g\zeta_1(h)^{1/2})) + \log(\epsilon)\|\zeta_1(h)^{1/2}g\zeta_1(h)^{1/2}\|_1].
\]

(Here \( h \) is the density \( \frac{D_{\omega}}{D_{\tau}} \) of the dual weight \( \omega \).)

**Remark 5.4.** What we must clarify is the meaning of the term “regularization” used in the above definition. The density \( h = \frac{D_{\omega}}{D_{\tau}} \) is related, by Bisognano-Wichmann results [15, 16], to the equilibrium hamiltonian, cf Remark 2.11 in [18]. This gives a relation to thermodynamics of equilibrium states, cf Remark 2.4. Further, \( \frac{D_{\omega}}{D_{\tau}} \) is in \( L^1(\mathcal{M}) \) space. This and Definition 5.3 imply that the regularization procedure stems from the prescription leading to the construction of \( L^{\text{ent}}(\mathcal{M}) \) space, see Definition 3.4 in [18].

On the other hand, it is worth noting that the same procedure was used to define \( \tau \)-measurability of quantum field operators, see [18]. Consequently, the regularization procedure is based on the selection of such measurable operators which are good candidates for representing states and this selection is compatible with the new formalism of statistical mechanics. We remind that this new formalism is based on the distinguished pair of Orlicz spaces \( (L^{\cosh^{-1}}(\mathcal{M}), L \log(L+1)(\mathcal{M})) \), for details see [19, 20].

One has (cf [18])

**Corollary 5.5.** If \( \vartheta \) is a regular state, then \( \tilde{S}(\vartheta) \) is well defined (although possibly infinite valued).

**Remark 5.6.** The important point to note here is that to define entropy for large systems (so for type III von Neumann algebras) we were here working within the new formalism, which is based on the distinguished pair of Orlicz spaces \( (L^{\cosh^{-1}}, L \log(L+1)) \) - for details see [17, 19, 20]. In particular, the superalgebra \( \mathcal{M} \) was employed. In that way it is possible to define entropy for non-semifinite von Neumann algebras, and consequently to study thermodynamics for such systems. Furthermore, this should make clear in which way we avoided the problems discussed in [22] - see Theorem 6.10 of that monograph.

Now let \( \mathcal{M} \) be a semifinite von Neumann algebra and \( \omega = \tau_\omega \) a tracial state. Let \( \vartheta \) be a faithful normal state for which the Radon-Nikodym derivative \( a \) described
in Theorem 2.5 belongs to the tracial space \([L \log(L + 1) \cap L^1](\mathcal{M}, \tau_\omega)\) (see the prescription in for example section 1 of [16] to see how this space is defined. Then by Proposition 2.5 and Definition 3.4 of [16], \(a\) corresponds to an element \(g\) of \([L \log(L + 1) \cap L^1](\mathcal{M})^+ = L^\text{ent}(\mathcal{M})^+\), which is of the form \(g = a \otimes \varphi_{\text{ent}}(\epsilon^t)\). Again by [16] Proposition 2.5, the operators \(\zeta_{\log(h)}(h)^{1/2}g\zeta_{\log(h)}(h)^{1/2}\) and \(\zeta_{1}(h)^{1/2}g\zeta_{1}(h)^{1/2}\) are respectively of the form \(a \otimes \varphi_{\log}(\epsilon^t)\) and \(a \otimes \epsilon^t\). We may therefore apply [12] Proposition 1.7 and [16] Theorem 2.2] to see that we will for any \(\epsilon > 0\) have that

\[
\epsilon \tau(\chi_{(\epsilon, \infty)}(\zeta_{\log(h)}(h)^{1/2}g\zeta_{\log(h)}(h)^{1/2})) + \log(\epsilon)\|\zeta_{1}(h)^{1/2}g\zeta_{1}(h)^{1/2}\|_1
\]

\[
= \epsilon \tau(\chi_{(\epsilon, \infty)}(\zeta_{\log(h)}(h)^{1/2}g\zeta_{\log(h)}(h)^{1/2})) + \log(\epsilon)\tau(\chi_{(1, \infty)}(\zeta_{1}(h)^{1/2}g\zeta_{1}(h)^{1/2}))
\]

\[
= \tau_\omega(a \log(a/\epsilon + 1)) + \log(\epsilon)\tau_\omega(a)
\]

\[
= \tau_\omega(a \log(a + \epsilon)).
\]

So in this case the formula in the preceding definition corresponds exactly to the more familiar formula \(\hat{S}(\vartheta) = \inf_{\epsilon > 0} \tau_\omega(a \log(a + \epsilon\mathbf{1})) = \tau_\omega(a \log(a))\).

6. Conclusions

One of the challenges of contemporary physics is to derive the macroscopic properties of matter from the quantum laws governing the microscopic description of a system. On the other hand, thermodynamics being a prerequisite for (quantum) statistical physics, provides laws governing the behaviour of macroscopic variables. It is well known that entropy is a crucial concept for this scheme.

Knowing that statistical physics deals with large systems (so systems with infinite degrees of freedom) we proposed a concise approach to entropy. It was done in operator algebraic language. This language is indispensable as on the one hand it is the basis for noncommutative integration theory, and on the other von Neumann algebras of type III are acknowledged to be the correct formalism for large quantum systems. Consequently, using the algebraic approach, a consistent dynamical description of entropy was achieved.

It is worth pointing out that our results can be considered as the first step in getting genuine quantum thermodynamics for general quantum systems.

References

[1] H. Araki, Relative entropy of states of von Neumann algebras, *Publ. RIMS* 11 (1976) pp. 809-833
[2] H. Araki, Relative entropy of states of von Neumann algebras II, *Publ. RIMS* 13 (1977) pp. 173-192
[3] H. Araki, T. Masuda, Positive cones and \(L_p\)-spaces for von Neumann algebras, *Pub. of RIMS* 18 (1982), pp. 339-411
[4] S. K. Berberian, Measure and Integration, AMS Chelsea Publications (reprint edition), 2011.
[5] J. J. Bisognano, E. H. Wichmann, On the duality condition for Hermitean scalar fields, *J. Math. Phys.* 16, 985 (1975)
[6] J. J. Bisognano, E. H. Wichmann, On the duality condition for quantum fields, *J. Math. Phys.* 17, 303 (1976)
[7] O. Bratteli, D. Robinson, Operator algebras and Quantum Statistical Mechanics , Texts and Monographs in Physics, Springer Verlag; vol. I, 1979; vol. II, 1981
[8] A Connes, Une classification de facteurs de type III, *Ann. Sci. École Norm. Sup.*, 4-eme, Ser. 6 (1973), pp. 133-252
[9] A Connes, Sur le théorème de Radon-Nikodym pour les normaux fidelis semiﬁni, *Bull. Sci.Math. SecII* 97 (1973) pp. 253-258
[10] A. Connes, M. Takesaki, The flow of weights on factors III, *Tohoku Math. Journ.* **29** (1977) pp. 473-575
[11] A. Connes, *Noncommutative geometry*, Academic Press, 1994
[12] T Fack and H Kosaki, Generalized s-numbers of $\tau$-measurable operators, *Pacific J Math* **123**(1986), 269-300.
[13] U. Haagerup, Operator valued weights in von Neumann algebras, *J. Funct. Anal.* **32** (1979) pp. 175–206
[14] A. I. Khinchin, *Mathematical foundations of Statistical Mechanics*, Dover Publications Inc., 1949
[15] H Kosaki, A remark on Sakai’s quadratic Radon-Nikodym theorem, *Proc AMS* **116**(1992), 783–786.
[16] L. E. Labuschagne, A crossed product approach to Orlicz spaces, *Proc. London Math. Soc.* **107** (2013) 965-1003
[17] L. E. Labuschagne, W. A. Majewski, Maps on noncommutative Orlicz spaces, *Illinois J. Math.* 55, 1053-1081 (2011)
[18] L. E. Labuschagne, W. A. Majewski, Integral and differential structures for quantum field theory, *arXiv: 1702.0065v2 [math-ph]*
[19] W. A. Majewski, L. E. Labuschagne, On applications of Orlicz spaces to Statistical Physics, *Ann. H. Poincaré*, 15, 1197-1221 (2014)
[20] W. A. Majewski, L. E. Labuschagne, Why are Orlicz spaces useful for Statistical Physics? in *Noncommutative Analysis, Operator Theory and Applications;* Eds. D. Alpay et al. Birkhauser-Basel, Series: Linear Operators and Linear Systems, vol 252, 271-283 (2016)
[21] W. A. Majewski, On quantum statistical mechanics; A study guide. *Adv. Math. Phys.* 2017, ID 934717, arXiv 1608.06766v2 [math-ph]
[22] M. Ohya, D. Petz, *Quantum entropy and its use*, Texts and Monographs in Physics, Springer Verlag, 1993
[23] G. K. Pedersen, *$C^*$-algebras and their automorphism groups*, Academic Press, 1979
[24] W. Pusz, S. L. Woronowicz, Passive states and KMS states for general quantum systems, *Commun. math. Phys.*, **58**, 273-290 (1978)
[25] D. Ruelle, *Statistical Mechanics; Rigorous Results*, W. A. Benjamin, INC. New York Amsterdam 1969
[26] M. Takesaki, *Theory of Operator Algebras*, vol. I, Springer, 1979
[27] M. Takesaki, *Theory of Operator Algebras*, vol. II, Springer, 2003
[28] M. Terp, *$L^p$ spaces associated with von Neumann algebras*, Rapport No 3a, University of Copenhagen, (1981)
[29] C. J. Thompson, *Mathematical Statistical Mechanics*, The Macmillan Company, New York, 1972.
[30] E. Witten, Notes on some entanglement properties of quantum field theory, *arXiv:1803.04993v3 [hep-th]*

**Institute of Theoretical Physics and Astrophysics, The Gdańsk University, Wita Stwosza 57, Gdansk, 80-952, Poland and Unit for BMI, North-West-University, Potchefstroom, South Africa**

**E-mail address:** fizwam@univ.gda.pl

DST-NRF CoE in Math. and Stat. Sci., Unit for BMI, Internal Box 209, School of Comp., Stat., & Math. Sci., NWU, PVT. BAG X6001, 2520 Potchefstroom, South Africa

**E-mail address:** Louis.Labuschagne@nwu.ac.za