Commun. Math. Phys. 105, 13–34 (1986)

Communications in Mathematical Physics


# On The Relative Entropy

Matthew J. Donald
18 Holland Street, Cambridge CB4 3DL, Great Britain

**Abstract.** For $\mathscr{A}$ any subset of $\mathscr{B}(\mathscr{H})$ (the bounded operators on a Hilbert space) containing the unit, and $\sigma$ and $\rho$ restrictions of states on $\mathscr{B}(\mathscr{H})$ to $\mathscr{A}$, $\text{ent}_{\mathscr{A}}(\sigma|\rho)$—the entropy of $\sigma$ relative to $\rho$ given the information in $\mathscr{A}$—is defined and given an axiomatic characterisation. It is compared with $\text{ent}^S_{\mathscr{A}}(\sigma|\rho)$—the relative entropy introduced by Umegaki and generalised by various authors—which is defined only for $\mathscr{A}$ an algebra. It is proved that ent and $\text{ent}^S$ agree on pairs of normal states on an injective von Neumann algebra. It is also proved that ent always has all the most important properties known for $\text{ent}^S$: monotonicity, concavity, $w^*$ upper semicontinuity, etc.

## 1. Introduction

Given states $\sigma$ and $\rho$ on a von Neumann algebra $\mathscr{A}$, the entropy of $\sigma$ relative to $\rho$, written $\text{ent}_{\mathscr{A}}(\sigma|\rho)$, is a measure of how easy it is to distinguish the state $\rho$ from the state $\sigma$. As such it has, since first introduced by Umegaki [1], found application in quantum statistical mechanics [2, 3 (Sect. 6.2.3, pp 269–289), 4], quantum information theory [1, 5], the foundations of quantum mechanics [6], and the theory of von Neumann algebras [7, 8]. I shall give a brief sketch below of how I see the role of relative entropy in the foundations of quantum theory, as this is my own motivation for studying the subject. If my view of these matters is correct then the relative entropy is of fundamental importance to physics.

As a mathematical object the relative entropy is fascinating. It has been given three distinct but equivalent definitions: that of Araki [9, 10] who uses Tomita–Takesaki theory, that of Pusz and Woronowicz [11] who use their "functional calculus for sesquilinear forms" [12], and that of Uhlmann [13] who uses interpolation theory. The entropy defined by Araki and Uhlmann will be denoted by $\text{ent}^S_{\mathscr{A}}(\sigma|\rho)$ throughout this paper. For the equivalence of their definitions see [8]. The entropy of Pusz and Woronowicz will be denoted by $\text{ent}^{PW}_{\mathscr{A}}(\sigma|\rho)$. Its equivalence to $\text{ent}^S_{\mathscr{A}}(\sigma|\rho)$ will be proved in Sect. 4 of this paper.

The purpose of this paper is to give yet another definition. This new definition has several advantages. It is given by means of a set of axioms and is conceptually and mathematically simpler than the previous definitions. It gives a characterisation of the relative entropy which allows for a heuristic interpretation. This is significant, since there has been some controversy [14] over the "naturalness" of relative entropy as a tool for quantum statistical inference. This will be discussed further in Sect. 3. The most important advantage of the new definition however is that it allows a useful and substantial generalisation. Given states $\sigma$ and $\rho$ on an algebra $\mathscr{B}$ and some subset $\mathscr{A} \subset \mathscr{B}$ we shall define $\mathrm{ent}_{\mathscr{A}}(\sigma|_{\mathscr{A}}|\rho|_{\mathscr{A}})$—"the entropy of $\sigma$ relative to $\rho$ given the information in $\mathscr{A}$." The generalisation, of course, is that $\mathscr{A}$ need no longer be an algebra. Although I conjecture that

$$\mathrm{ent}_{\mathscr{A}}(\sigma|\rho) = \mathrm{ent}^{S}_{\mathscr{A}}(\sigma|\rho) \tag{1.1}$$

whenever the latter is defined, I can only prove this for the case when $\mathscr{A}$ is an injective von Neumann algebra and $\sigma$ and $\rho$ are normal. This equivalence theorem is proved in Sect. 8. However, because all the most useful properties of $\mathrm{ent}^S$ also hold for ent (see Sect. 6) and because the axioms (Sect. 2) are natural, if a counterexample to (1.1) should emerge I would take this as a failing of $\mathrm{ent}^S$ rather than of ent.

Most of the material in the remaining sections provides technical preliminaries. Sections 3 discusses alternatives to the axioms, Sect. 4 proves results about $\mathrm{ent}^S$. Sect. 5 about concave functions, and Sect. 7 about injective von Neumann algebras.

**Caveat.** *In order to have the interpretatively valuable property that increasing probability means increasing entropy, the convention of Bratteli and Robinson [3 (Sect. 6.2.3)] who define* $\mathrm{ent}_{\mathscr{A}}(\sigma|\rho)$ *so that* $\mathrm{ent}_{\mathscr{A}}(\sigma|\rho) \leqq 0$ *has been followed here. We also, of course, take* $\mathrm{ent}^S_{\mathscr{A}}(\sigma|\rho) \leqq 0$. *The convention* $\mathrm{ent}_{\mathscr{A}}(\sigma|\rho) \geqq 0$ *is frequently followed elsewhere, and some writers exchange the arguments $\sigma$ and $\rho$. Thus if* $\sigma = \sum_j s_j |\phi_j\rangle\langle\phi_j|$, $\rho = \sum_i r_i |\psi_i\rangle\langle\psi_i|$, *are the eigenfunction expansions of density matrices $\sigma$ and $\rho$ on a Hilbert space $\mathscr{H}$, with $\mathscr{B}(\mathscr{H})$ the algebra of all bounded operators on $\mathscr{H}$, then our definition will yield*

$$\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho) = \sum_{i,j} (-s_j \log s_j + s_j \log r_i) |\langle\phi_j|\psi_i\rangle|^2 . \tag{1.2}$$

Turning now to my intended application of the work of this paper, one way of stating "the" fundamental problem of the foundations of quantum mechanics is that although it appears that the state of the world can mainly be described by Heisenberg (i.e. time independent) states $\rho$ on $\mathscr{B}(\mathscr{H})$—for some $\mathscr{H}$—those states seem to change as the result of measurement. This is just the statement that time evolution at a measurement is not governed by the unitary group $e^{-itH}$ (where $H$ is the Hamiltonian). Now the set of possible states $\{\sigma_j : j \in J\}$ which can describe the world after the measurement are circumscribed by requirements of compatibility with the measurement device (a macroscopic pointer must point in a definite direction, a cat must be either alive or dead). These compatibility requirements can be defined by the values of the $\sigma_j$ on some subset $\mathscr{A}$ of the algebra $\mathscr{B}(\mathscr{H})$. In the examples mentioned, $J$ will then be the set of pointer positions, or the set {alive, dead}.

It is my belief that a complete and consistent interpretation of quantum theory can be given based on the equation

$$\frac{\text{Prob}\{\text{result of measurement is } \sigma_j\}}{\text{Prob}\{\text{result of measurement is } \sigma_k\}} = \frac{\exp(\text{ent}_{\mathscr{A}}(\sigma_j|\rho))}{\exp(\text{ent}_{\mathscr{A}}(\sigma_k|\rho))}, \tag{1.3}$$

where $\rho$ is the state describing the world before the measurement. Within the next few years I intend to publish work, already long in progress, substantiating this belief. Of course, the real difficulties lie in the construction of the sets $\mathscr{A}$ (which leads to a theory of consciousness) and of the sets $\{\sigma_j : j \in J\}$ (which leads to a many worlds theory), and in the interpretation of the formula (1.3) (for example, in general, $\sum_j \exp(\text{ent}_{\mathscr{A}}(\sigma_j|\rho))$ is not equal to one (!)). The present paper merely establishes some of the mathematical background to this work and can, of course, be read entirely independently of it. Indeed, I trust that the results will also prove useful for other applications of the relative entropy.

As already mentioned, one purpose of this paper is to extend the definition of $\text{ent}_{\mathscr{A}}(\sigma|\rho)$ to the case when $\mathscr{A}$ is not necessarily an algebra. This, of course, is very useful for measurement theoretic applications, since the result of a measurement is usually only to give values to a few operators and this may well not be sufficient to assign a unique state to some subalgebra. In order to avoid repetition in this context, we assume henceforth the following convention.

**Notational Convention.** *The notation* $\text{ent}_{\mathscr{A}}(\sigma|\rho)$ *includes the statement that there is a Hilbert space* $\mathscr{H}$ *with* $\mathscr{A} \subset \mathscr{B}(\mathscr{H})$, *and* $1 \in \mathscr{A}$, *and that there exist states* $\sigma'$ *and* $\rho'$ *on* $\mathscr{B}(\mathscr{H})$ *with* $\sigma = \sigma'|_{\mathscr{A}}$ *and* $\rho = \rho'|_{\mathscr{A}}$.

This convention requires that $\sigma(1) = \rho(1) = 1$. This is the case of relevance for the applications. It would make no difference to any of the results of this paper, but would lengthen some statements, if we assumed that ent was defined on general positive linear functionals $\sigma$ and $\rho$ by requiring the following relation to hold for all $\lambda > 0$, $\mu > 0$:

$$\text{ent}_{\mathscr{A}}(\lambda\sigma|\mu\rho) = \lambda\,\text{ent}_{\mathscr{A}}(\sigma|\rho) + (-\lambda\log\lambda + \lambda\log\mu)\,\sigma(1). \tag{1.4}$$

## 2. An Axiomatic Characterisation of the Relative Entropy

Equation (1.3) is compatible with a heuristic interpretation of $\text{ent}_{\mathscr{A}}(\sigma|\rho)$ which states that $\exp(\text{ent}_{\mathscr{A}}(\sigma|\rho))$ is the probability, per unit trial of the information in $\mathscr{A}$, of being able to mistake the state of the world for $\sigma$ despite the fact that it is actually $\rho$. In other words, $\exp(N\,\text{ent}_{\mathscr{A}}(\sigma|\rho))$ is (or is, in some sense, asymptotic as $N \to \infty$ to) the probability that $N$ tests on the state $\rho$ of the information in $\mathscr{A}$ give results compatible with the state being $\sigma$. For the special case when $\mathscr{A}$ is an abelian algebra a heuristic interpretation of this type is discussed by Bratteli and Robinson [3, p. 425 et seq.]. In particular, they prove that in $N$ independent repetitions of a trial, which has $M$ possible outcomes with probabilities $r_1, \ldots, r_M$, the event that, for each $j$, the $j^{th}$ outcome occurs with frequency $s_j$, has probability whose logarithm is asymptotic to

$$N\left\{\sum_{j=1}^{M}(-s_j\log s_j + s_j\log r_j)\right\} \tag{2.1}$$

(cf. Eq. (1.2)).

The following set of axioms for relative entropy can be constructed based on this heuristic interpretation. It is also, of course, based on the standard probabilistic interpretation of the states of quantum theory. The reader who finds the arguments for these axioms long-winded is encouraged to consider first the simple definition, presented at the beginning of Sect. 6, which they give rise to.

First consider the situation where one can make measurements on some set of operators $\mathscr{A}$, and one wishes to distinguish between states $\sigma$ and $\rho$ which, as required above, have extensions $\sigma'$ and $\rho'$ on some $\mathscr{B}(\mathscr{H}) \supset \mathscr{A}$. The existence of properties that one does not measure only affects one's belief about whether the state restricted to $\mathscr{A}$ is $\sigma$ or not by the requirement that any possible state must take some value on the operators corresponding to those properties. This yields:

**Axiom I.** $\quad \mathrm{ent}_{\mathscr{A}}(\sigma|\rho) = \sup\{\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho') : \sigma'|_{\mathscr{A}} = \sigma, \rho'|_{\mathscr{A}} = \rho\}.$

Now suppose that $\sigma = \sum_{j=1}^{J} s_j \sigma_j$, where $\sigma$ and each of the $\sigma_j$ are normal states on $\mathscr{B}(\mathscr{H})$, $0 \leqq s_j \leqq 1$, and the $\sigma_j$ have disjoint supports. In this case $\sigma$ can be viewed as a mixture of the independent states $\sigma_j$ with probabilities $s_j$. Thus, as $N \to \infty$, $\exp(N\,\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho))$ should be asymptotic to the probability that $N$ trials of a random variable $X$ with values in $\{1, 2, 3, \ldots\}$ and distribution $\mathrm{Prob}\{X=j\} = \exp(\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma_j|\rho))$ has a result with frequencies assigned according to the distribution $s_j$ (i.e. $X=j$ in $s_j N$ of the trials). The event $X=j$ is the event of making the mistake that the state of the world is $\sigma_j$ although it is actually $\rho$. Using Eq. (2.1) this yields:

**Axiom II.** *If* $\sigma = \sum_{j=1}^{J} s_j \sigma_j$, *where* $\{\sigma_j\}$, $\sigma$, *and* $\rho$ *are normal states on* $\mathscr{B}(\mathscr{H})$ *and where the* $\sigma_j$ *have disjoint supports, then*

$$\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho) = \sum_{j=1}^{J} (-s_j \log s_j + s_j \mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma_j|\rho)).$$

The restriction of this axiom to normal states on $\mathscr{B}(\mathscr{H})$ is because it is only in this case that it is clearly justifiable to apply the argument above and assume that $\sigma$ is physically a mixture.

Next consider the case when $\sigma$ is a pure normal state $\sigma = |\phi\rangle\langle\phi|$ on $\mathscr{B}(\mathscr{H})$, and when $\rho$ also normal, has the eigenfunction decomposition $\rho = \sum_i r_i |\psi_i\rangle\langle\psi_i|$. Assuming invariance under unitary maps, $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho)$ can depend only on $\{r_i, |\langle\phi|\psi_i\rangle|^2\}$. It is natural to require that $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(|\psi_i\rangle\langle\psi_i||\rho) = \log r_i$, since this follows from (2.1) and the assumption that $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\{\psi_i\rangle\langle\psi_i||\rho) = \mathrm{ent}_{\mathscr{A}}(|\psi_i\rangle\langle\psi_i||\rho)$, where $\mathscr{A}$ is the algebra generated by the projections $|\psi_j\rangle\langle\psi_j|$. We now give a value to $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(|\phi\rangle\langle\phi||_\rho)$ by assuming that it is as easy to mistake $\rho$ for $|\phi\rangle\langle\phi|$ during each of $N$ trials as it would be to mistake, for each $i$, $\rho$ for $|\psi_i\rangle\langle\psi_i|$ in a *predetermined* portion $|\langle\phi|\psi_i\rangle|^2$ of the trials. This assumption is to be distinguished carefully from the arguments for Axiom II. In the present axiom we compute a probability for a pure state by making an analogy, while in Axiom II we consider a mixed state $\sigma$ as if it actually behaved like each of its components for some of the time. Thus we would, for example, say that the mixed density matrix $\sum_i |\langle\phi|\psi_i\rangle|^2 |\psi_i\rangle\langle\psi_i|$ behaves as if it were actually mistaken for $|\psi_i\rangle\langle\psi_i|$ in a *random* portion $|\langle\phi|\psi_i\rangle|^2$ of trials.

Since we are only trying to give a heuristic argument for our axiom, we can assume that, for some $n$, $\langle\phi|\psi_i\rangle = 0$ for $i > n$ and that $|\langle\phi|\psi_i\rangle|^2 N$ is always an integer. Then the interpretation of $\exp(N(\text{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho))$ as a probability plus the assumption (spelt out for clarity) that it is as easy to mistake $\rho$ for $|\phi\rangle\langle\phi|$ during each of $N$ trials as it would be to mistake $\rho$ for $|\psi_1\rangle\langle\psi_1|$ during the first $|\langle\phi|\psi_1\rangle|^2 N$ of $N$ trials, to mistake $\rho$ for $|\psi_2\rangle\langle\psi_2|$ during the next $|\langle\phi|\psi_2\rangle|^2 N$ of $N$ trials, and so on up to the final $|\langle\phi|\psi_n\rangle|^2 N$ trials when we mistake $\rho$ for $|\psi_n\rangle\langle\psi_n|$, yields

$$\begin{aligned}\text{ent}_{\mathscr{B}(\mathscr{H})}(|\phi\rangle\langle\phi||\rho) &= \sum_{i=1}^{n} |\langle\phi|\psi_i\rangle|^2 \,\text{ent}_{\mathscr{B}(\mathscr{H})}(|\psi_i\rangle\langle\psi_i||\rho)\\ &= \sum_{i=1}^{n} |\langle\phi|\psi_i\rangle|^2 \log r_i,\end{aligned}$$

and produces:

**Axiom III.** *For $\sigma = |\phi\rangle\langle\phi|$ a pure normal state on $\mathscr{B}(\mathscr{H})$, and for $\rho$ a normal state on $\mathscr{B}(\mathscr{H})$,*

$$\text{ent}_{\mathscr{B}(\mathscr{H})}(|\phi\rangle\langle\phi||\rho) = \langle\phi|\log\rho|\phi\rangle.$$

Axioms II and III suffice to define $\text{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho)$ for all normal states $\sigma$ and $\rho$ on $\mathscr{B}(\mathscr{H})$ in agreement with formula (1.2). However as example 6.6 will show, it is necessary to consider $\text{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho)$ for $\sigma$ and $\rho$ non-normal in order to have a definition of $\text{ent}_{\mathscr{A}}(\sigma|\rho)$, which is $w^*$ upper semicontinuous for general sets $\mathscr{A}$, even if attention is restricted to functionals $\sigma$ and $\rho$ which have normal extensions on $\mathscr{B}(\mathscr{H})$. In order also to avoid the objection that non-normal states are unphysical we must give an axiom which makes manifest that they arise in the present theory merely as a mathematical tool. The following axiom seems to be the only natural way of doing this.

**Axiom IV.** *For any $(\sigma, \rho, \mathscr{A})$ (subject of course to the notational convention of Sect. 1) there exists a net $((\sigma'_\alpha, \rho'_\alpha))_{\alpha\in I}$ of pairs of normal states on $\mathscr{B}(\mathscr{H})$ converging $w^*$ to $(\sigma', \rho')$ (i.e. such that $\sigma'_\alpha(B) \to \sigma'(B)$ and $\rho'_\alpha(B) \to \rho'(B)$ for all $B \in \mathscr{B}(\mathscr{H})$), and such that setting $(\sigma_\alpha, \rho_\alpha) = (\sigma'_\alpha|_{\mathscr{A}}, \rho'_\alpha|_{\mathscr{A}})$, we have $(\sigma'|_{\mathscr{A}}, \rho'|_{\mathscr{A}}) = (\sigma, \rho)$, and*

$$\begin{aligned}\text{ent}_{\mathscr{A}}(\sigma|\rho) &= \text{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho')\\ &= \lim_{\alpha\in I} \text{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'_\alpha|\rho'_\alpha)\\ &= \lim_{\alpha\in I} \text{ent}_{\mathscr{A}}(\sigma_\alpha|\rho_\alpha).\end{aligned}$$

## 3. Remarks on the Axioms

It is of course not obvious that the axioms do constitute a definition. In this preliminary section we will assume that they do and make some less technical comments.

There has been a quite separate definition of something also called "relative entropy" by Benoist, Gudder, Marchand, and Wyss [15, 14]. Their relative entropy is not equivalent to that studied here. For the problem of statistical inference they study a process that in our notation reads as follows. Given a state $\rho$, which constitutes the a priori information, on an algebra $\mathscr{B}$ with a subalgebra $\mathscr{A}$ and a state $\sigma$ defined on $\mathscr{A}$ constituting the result of a measurement, define an inferred state $\hat{\sigma}$ on $\mathscr{B}$ as that state which attains

$$\sup\{\mathrm{ent}_{\mathscr{B}}(\sigma'|\rho): \sigma'|_{\mathscr{A}} = \sigma\}.$$

They claim that their relative entropy is natural because the state $\hat{\sigma}$ resulting from the inference is close to $\rho$ in some natural sense. In particular the Uhlmann transition probability from $\sigma'$ to $\rho$ is maximised at $\hat{\sigma}$, and $\hat{\sigma}$ can be defined by the Sakai operator relating $\sigma$ to $\rho$ whenever there exists $\lambda$ such that $\sigma \leqq \lambda\rho|_{\mathscr{A}}$. I disagree with their claim of naturalness, and base my disagreement on the opposing claim that closeness of $\hat{\sigma}$ to $\rho$ in this context should be defined rather by maximisation of a relative entropy obeying precisely the axioms of Sect. 2. Indeed, a prime reason for presenting those axioms is to make this point. Of course, like $\mathrm{ent}^S$, the B.G.M.W. relative entropy can only deal with measurements of all the operators in an algebra.

Turn now to considering an alternative to Axiom III. I have not given an axiom specifically defining $\mathrm{ent}_{\mathscr{A}}(\sigma|\rho)$ for an abelian algebra, since this definition follows from those I have presented (see Sect. 6). However it is interesting to give the argument which would derive such an axiom, and then to consider the resulting definition. For the remainder of this section attention will be confined to normal states. Suppose then that $\sigma$ and $\rho$ are such states defined on the abelian algebra $\mathscr{Z}$ generated by the mutually orthogonal projections $(P_n)_{n=1}^{\infty}$ with $\sum_{n=1}^{\infty} P_n = 1$.

**Property V.** $\mathrm{ent}_{\mathscr{Z}}(\sigma|\rho) = -\sum_{n=1}^{\infty} \sigma(P_n)\log\frac{\sigma(P_n)}{\rho(P_n)}$.

The argument for this property is that as $N \to \infty$ $N\,\mathrm{ent}_{\mathscr{Z}}(\sigma|\rho)$ should be asymptotic to the logarithm of the probability that $N$ trials of a random variable $X$ with values in $\{1, 2, 3, \ldots\}$ and distribution $\mathrm{Prob}\{X = n\} = \rho(P_n)$ has a result determined by the distribution $\sigma(P_n)$ (i.e. $X = n$ in $\sigma(P_n)N$ of the trials). The result then follows from Eq. (2.1).

Given the simplicity of this argument, it is of interest to attempt to derive Axiom III from Property V. I shall show very briefly that this attempt is unsuccessful as this casts further light on the axioms and lays to rest a natural conjecture.

The following definitions might seem natural:

$$\mathrm{ent}^a_{\mathscr{B}(\mathscr{H})}(\sigma|\rho) = \inf\left\{ -\sum_{n=1}^{N} \sigma(P_n)\log\frac{\sigma(P_n)}{\rho(P_n)} : (P_n)_{n=1}^{N} \text{ are mutually orthogonal projections in } \mathscr{B}(\mathscr{H}) \text{ with } \sum_{n=1}^{N} P_n = 1 \right\}$$

$$\mathrm{ent}^b_{\mathscr{B}(\mathscr{H})}(\sigma|\rho) = \inf\left\{ -\sum_{n=1}^{N} \sigma(A_n)\log\frac{\sigma(A_n)}{\rho(A_n)} : (A_n)_{n=1}^N \right.$$

are positive operators in $\mathscr{B}(\mathscr{H})$ with $\left.\sum_{n=1}^{N} A_n = 1\right\}$.

Clearly $$\mathrm{ent}^b_{\mathscr{B}(\mathscr{H})}(\sigma|\rho) \leqq \mathrm{ent}^a_{\mathscr{B}(\mathscr{H})}(\sigma|\rho).$$

We shall prove in the final section (Proposition 8.7) that

$$\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho) \leqq \mathrm{ent}^b_{\mathscr{B}(\mathscr{H})}(\sigma|\rho). \tag{3.1}$$

Unfortunately it is *not* true in general that

$$\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho) = \mathrm{ent}^a_{\mathscr{B}(\mathscr{H})}(\sigma|\rho), \tag{3.2}$$

or even that

$$\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho) = \mathrm{ent}^b_{\mathscr{B}(\mathscr{H})}(\sigma|\rho). \tag{3.3}$$

Both fail for the case $\mathscr{H} = \mathbb{C}^2$ and

$$\sigma = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad \text{and} \quad \rho = \begin{pmatrix} \frac{3}{8} & 0 \\ 0 & \frac{5}{8} \end{pmatrix}.$$

The proof that (3.2) fails is a straightforward numerical computation. The proof that (3.3) fails is a more complicated argument which because of its negative import it seems fruitless to expound in detail here. Briefly, after establishing some preliminary restrictions on the form of sequences $(A_n)_{n=1}^N$ over which one need minimise (w.l.o.g. one can take each $A_n$ of the form $A_n = a_n Q_n$ for $a_n$ real and positive and $Q_n$ a projection), one shows that for each $N$, $-\sum_{n=1}^{N} \sigma(A_n)\log(\sigma(A_n)/\rho(A_n))$ has a strict local minimum at $(\hat{A}_n)_{n=1}^N$, where $\hat{A}_1 = \hat{P}_1$, $\hat{A}_2 = \hat{P}_2$, $\hat{A}_n = 0$ $n > 2$, and where $(\hat{P}_n)_{n=1}^2$ is the pair of projections attaining the infimum defining $\mathrm{ent}^a_{\mathscr{B}(\mathscr{H})}(\sigma|\rho)$. It is then straightforward to use the local minimum for $2N$ to derive a global minimum for $N$. This shows, in particular, that for this pair of $\sigma$ and $\rho$,

$$\mathrm{ent}^a_{\mathscr{B}(\mathscr{H})}(\sigma|\rho) = \mathrm{ent}^b_{\mathscr{B}(\mathscr{H})}(\sigma|\rho). \tag{3.4}$$

Whether this is true in general I do not know.

The particular case is especially critical since using the unitary map $U = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, which gives $U\sigma U^* = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$ and $U\rho U^* = \rho$, it is possible to derive the standard value of $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho)$ just from unitary invariance, Axiom II, and the assumption that

$$\mathrm{ent}_{\mathscr{B}(\mathscr{H})}\left(\begin{pmatrix} \lambda & 0 \\ 0 & 1-\lambda \end{pmatrix}\middle|\begin{pmatrix} \mu & 0 \\ 0 & 1-\mu \end{pmatrix}\right) = -\lambda\log\frac{\lambda}{\mu} - (1-\lambda)\log\frac{(1-\lambda)}{(1-\mu)}. \tag{3.5}$$

Since (3.5) is satisfied by $\mathrm{ent}^a$ and $\mathrm{ent}^b$ as well as by ent and $\mathrm{ent}^S$, it follows that neither $\mathrm{ent}^a$ nor $\mathrm{ent}^b$ satisfy Axiom II. Thus neither $\mathrm{ent}^a$ nor $\mathrm{ent}^b$ provides a suitable alternative to ent.

## 4. Results on $\mathrm{ent}^S$

In order to verify that ent possesses the various properties which we will consider below (Sect. 6), it is necessary to rely heavily on earlier results and methods developed for $\mathrm{ent}^S$. Indeed we need to extend some of these results, and this will be done in this section. Thus it must be allowed that the simplicity claimed for ent is a simplicity only of definition and not, in general, of usage.

First we define $\mathrm{ent}^{PW}$ and by tightening a formal argument of Pusz and Woronowicz [11, 12] show that it is $w^*$ (i.e. $\sigma(\mathscr{A}^*, \mathscr{A})$) upper semicontinuous.

*Definition 4.1.* Let $\mathscr{A}$ be a von Neumann algebra acting on a Hilbert space $\mathscr{H}$, and let $\sigma$ and $\rho$ be states on $\mathscr{A}$. Define the P.-W. representation of $(\mathscr{A}, \mathscr{H}, \sigma, \rho)$ to be a tetrad $(\mathscr{K}, k, S, R)$ such that $\mathscr{K}$ is a Hilbert space, $k$ is a linear map of $\mathscr{A}$ onto a dense subset of $\mathscr{K}$, and $S$ and $R$ are commuting positive bounded operators on $\mathscr{K}$ such that for all $A_1, A_2 \in \mathscr{A}$,

$$(k(A_1), k(A_2)) = \rho(A_1^* A_2) + \sigma(A_2 A_1^*),$$

$$(k(A_1), Rk(A_2)) = \rho(A_1^* A_2),$$

$$(k(A_1), Sk(A_2)) = \sigma(A_2 A_1^*).$$

Define $\mathrm{ent}_{\mathscr{A}}^{PW}(\sigma|\rho) = (k(1), g(S, R)\, k(1))$, where $g(s, r) = -s \log s/r$.

Pusz and Woronowicz [11, 12] prove that this is a definition and that (with the current sign convention) $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}^{PW}(\sigma|\rho)$ is given by formula (1.2) for $\sigma$ and $\rho$ normal.

If $dE(s, r)$ is the joint spectral measure of $S$ and $R$ on $\mathscr{K}$ then, by the monotone convergence theorem and Fubini's theorem,

$$\begin{aligned}\mathrm{ent}_{\mathscr{A}}^{PW}(\sigma|\rho) &= \int \left(-s \log \frac{s}{r} - (r - s)\right)(k(1), dE(s, r)k(1)) \\ &= -\lim_{\substack{\delta \searrow 0 \\ \varepsilon \nearrow 1}} \int \int_{\delta}^{\varepsilon} \frac{(s-r)^2 t}{s(1-t) + rt} dt (k(1), dE(s, r)k(1)) \\ &= \inf\left\{-\int_{\delta}^{\varepsilon} \left(k(1), \frac{(S-R)^2 t}{S(1-t) + Rt} k(1)\right) dt \colon 1 > \varepsilon > \delta > 0\right\}.\end{aligned}$$

Let $\Theta_\delta^\varepsilon$ be the set of pairs $(A(t), B(t))$ such that $A\colon [\delta, \varepsilon] \to \mathscr{A}$ and $B\colon [\delta, \varepsilon] \to \mathscr{A}$ are piecewise constant functions with $A(t) + B(t) = 1$ for all $t \in [\delta, \varepsilon]$. Using $(k(1), (R - S)k(1)) = \rho(1) - \sigma(1) = 0$, it is mere algebra to show that, for $(A(t), B(t)) \in \Theta_\delta^\varepsilon$,

$$\begin{aligned}&-\int_{\delta}^{\varepsilon} \left(k(1), \frac{(S-R)^2 t}{S(1-t) + Rt} k(1)\right) dt \\ &= \int_{\delta}^{\varepsilon} \left((k(B(t)), (R-S)k(B(t))) + \frac{1}{t}(k(A(t)), Sk(A(t)))\right) dt - \int_{\delta}^{\varepsilon} \frac{1}{t}\left(\left(\frac{(R-S)tk(1)}{S(1-t) + Rt} - k(A(t))\right),\right. \\ &\quad \left. \cdot (S(1-t) + Rt)\left(\frac{(R-S)tk(1)}{S(1-t) + Rt} - k(A(t))\right)\right) dt. \qquad (4.2)\end{aligned}$$

Since $k(\mathscr{A})$ is dense in $\mathscr{K}$ and $t \mapsto (R - S)tk(1)/S(1 - t) + Rt$ is continuous on $[\delta, \varepsilon]$, for any $\zeta > 0$ there exists $(A(t), B(t)) \in \Theta_\delta^\varepsilon$ such that

$$\left\| \frac{(R-S)tk(1)}{S(1-t)+Rt} - k(A(t)) \right\| < \zeta \quad \text{for all} \quad t \in [\delta, \varepsilon]. \tag{4.3}$$

From the definitions, for $(A(t), B(t)) \in \Theta_\delta^\varepsilon$,

$$\int_\delta^\varepsilon \left( (k(B(t)), (R-S)k(B(t))) + \frac{1}{t}(k(A(t)), Sk(A(t))) \right) dt$$

$$= \int_\delta^\varepsilon \left( \sigma\left( \frac{1}{t} A(t)A(t)^* - B(t)B(t)^* \right) + \rho(B(t)^*B(t)) \right) dt$$

$$= \sigma\left( \int_\delta^\varepsilon \left( \frac{1}{t} A(t)A(t)^* - B(t)B(t)^* \right) dt \right) + \rho\left( \int_\delta^\varepsilon B(t)^*B(t)dt \right).$$

Let $\Gamma_\delta^\varepsilon = \Big\{ (X, Y)$: for some $(A(t), B(t)) \in \Theta_\delta^\varepsilon$,

$$X = \int_\delta^\varepsilon \left( \frac{1}{t} A(t)A(t)^* - B(t)B(t)^* \right) dt, \quad Y = \int_\delta^\varepsilon B(t)^*B(t)dt \Big\}.$$

Then, using the estimate (4.3) to show that the second term in Eq. (4.2), which is always negative, can be made arbitrarily close to zero,

$$-\int_\delta^\varepsilon \left( k(1), \frac{(S-R)^2 t}{S(1-t)+Rt} k(1) \right) dt = \inf\{\sigma(X) + \rho(Y): (X, Y) \in \Gamma_\delta^\varepsilon\},$$

and the following lemma has been proved:

**Lemma 4.4.** $\quad \mathrm{ent}_{\mathscr{A}}^{PW}(\sigma|\rho) = \inf\{\sigma(X) + \rho(Y): (X, Y) \in \bigcup_{0<\delta<\varepsilon<1} \Gamma_\delta^\varepsilon\}.$ □

Having exhibited $\mathrm{ent}_{\mathscr{A}}^{PW}(\sigma|\rho)$ as an infimum of real-valued, $w^*$ continuous, linear functions, the following is an immediate consequence:

**Proposition 4.5.** $\mathrm{ent}_{\mathscr{A}}^{PW}(\sigma|\rho)$ *is jointly concave and* $w^*$ *upper semicontinuous.* □

We now give a straightforward proof that $\mathrm{ent}_{\mathscr{A}}^{PW}(\sigma|\rho) = \mathrm{ent}_{\mathscr{A}}^{S}(\sigma|\rho)$.

With the notation of Pusz and Woronowicz, as introduced above, Uhlmann's definition of $\mathrm{ent}^S$ reads

$$= \lim_{x \searrow 0} \frac{(k(1), (S^{1-x}R^x - S)k(1))}{x}$$

$$= \lim_{x \searrow 0} \int f(x,s,r)(k(1), dE(s,r)k(1)) + \rho(1) - \sigma(1),$$

where

$$f(x,s,r) = \frac{s^{1-x}r^x - s}{x} - (r-s).$$

For $s \geqq 0$, $r \geqq 0$, and $1 \geqq x > 0$, we have $f(x,s,r) \leqq 0$, and $f(x,s,r)$ decreases monotonically to $f(0,s,r) = -s\log s/r - (r-s)$, so by the monotone convergence theorem,

$$\text{ent}^S_{\mathscr{A}}(\sigma|\rho) = \text{ent}^{PW}_{\mathscr{A}}(\sigma|\rho). \tag{4.6}$$

We will refer to both as $\text{ent}^S$ hereafter.

## 5. A Lemma on Concave Functions

The purpose of this section is to prove a lemma which we will use in the next section to give our definition of $\text{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho)$ for $\sigma$ or $\rho$ non-normal. This lemma may be well known, but I do not know of any reference for it.

Suppose given a $w^*$ upper semicontinuous concave function $G(\sigma, \rho)$ on pairs of normal states on $\mathscr{B}(\mathscr{H})$ and consider the extension $H$ of $G$ to pairs of arbitrary states on $\mathscr{B}(\mathscr{H})$ defined by

$$H(\sigma,\rho) = \sup\{\limsup_{\alpha\in I} G(\sigma_\alpha, \rho_\alpha):((\sigma_\alpha, \rho_\alpha))_{\alpha\in I} \text{ is a net with } \sigma_\alpha \text{ and } \rho_\alpha \text{ normal states on } \mathscr{B}(\mathscr{H}) \text{ and } \sigma_\alpha \xrightarrow{w^*} \sigma, \rho_\alpha \xrightarrow{w^*} \rho\}.$$

**Lemma 5.1.**
1) *$H$ is $w^*$ upper semicontinuous and concave, and $H(\sigma,\rho) = G(\sigma,\rho)$ for $\sigma$ and $\rho$ normal.*
2) *$H(\sigma,\rho) = \hat{F}(\sigma,\rho)$, where*

$$\hat{F}(\sigma,\rho) = \inf\{F(\sigma,\rho): F \text{ is } w^* \text{ upper semicontinuous and concave, and } F(\sigma,\rho) = G(\sigma,\rho) \text{ for } \sigma \text{ and } \rho \text{ normal}\}.$$

*Proof.* Suppose 1) is true. Then $\hat{F}$ exists and from the definition of $w^*$ upper semicontinuity (see Property $c$, Sect. 6) $H(\sigma,\rho) \leqq \hat{F}(\sigma,\rho)$, so 2) is true. It is thus sufficient to prove 1).

i) Let $((\sigma_\beta, \rho_\beta))_{\beta\in J}$ be a net such that $\sigma_\beta \xrightarrow{w^*} \sigma$, $\rho_\beta \xrightarrow{w^*} \rho$, and, for some $\delta > 0$, $H(\sigma_\beta, \rho_\beta) > H(\sigma,\rho) + \delta$ for all $\beta \in J$. Let $\mathscr{N}$ be an open $w^*$ neighbourhood of $(\sigma,\rho)$ in the space of pairs of states. There exists $\beta(\mathscr{N}) \in J$ such that $(\sigma_{\beta(\mathscr{N})}, \rho_{\beta(\mathscr{N})}) \in \mathscr{N}$. Since $\mathscr{N}$ is an open $w^*$ neighbourhood of $(\sigma_{\beta(\mathscr{N})}, \rho_{\beta(\mathscr{N})})$, there exist normal $\sigma_{\mathscr{N}}$ and $\rho_{\mathscr{N}}$ with $(\sigma_{\mathscr{N}}, \rho_{\mathscr{N}}) \in \mathscr{N}$ and

$$G(\sigma_{\mathscr{N}}, \rho_{\mathscr{N}}) > H(\sigma_{\beta(\mathscr{N})}, \rho_{\beta(\mathscr{N})}) - \tfrac{1}{2}\delta > H(\sigma,\rho) + \tfrac{1}{2}\delta.$$

This is a contradiction, since if $K$ is the net of open $w^*$ neighbourhoods of $(\sigma,\rho)$ ordered by inclusion then $(\sigma_{\mathscr{N}}, \rho_{\mathscr{N}})_{\mathscr{N}\in K} \xrightarrow{w^*} (\sigma,\rho)$.

Thus $H$ is $w^*$ upper semicontinuous.

ii) It is easy to observe that

$$H(\sigma,\rho) = \sup\{\lim_{\alpha\in I} G(\sigma_\alpha, \rho_\alpha): ((\sigma_\alpha, \rho_\alpha))_{\alpha\in I} \text{ is a net with } \sigma_\alpha \text{ and } \rho_\alpha \text{ normal, } \sigma_\alpha \xrightarrow{w^*} \sigma, \rho_\alpha \xrightarrow{w^*} \rho, \text{ and } \lim_{\alpha\in I} G(\sigma_\alpha, \rho_\alpha) \text{ exists}\}.$$

Let $(\sigma,\rho) = (\lambda_1\sigma^1 + \lambda_2\sigma^2, \lambda_1\rho^1 + \lambda_2\rho^2)$ with $\lambda_1, \lambda_2 \geqq 0$ and $\lambda_1 + \lambda_2 = 1$. Choose and fix $\varepsilon > 0$. By the remark above, there exist nets $((\sigma^1_\beta, \rho^1_\beta))_{\beta\in J} \xrightarrow{w^*} (\sigma^1, \rho^1)$ and

$((\sigma_\gamma^2,\rho_\gamma^2))_{\gamma\in K}\xrightarrow{w^*}(\sigma^2,\rho^2)$ satisfying $\lim_{\beta\in J} G(\sigma_\beta^1,\rho_\beta^1)\geqq H(\sigma^1,\rho^1)-\varepsilon$ and $\lim_{\gamma\in K} G(\sigma_\gamma^2,\rho_\gamma^2)\geqq H(\sigma^2,\rho^2)-\varepsilon$. Define $I=J\times K$ with $(\beta',\gamma')>(\beta,\gamma)\Leftrightarrow\beta'>\beta$ and $\gamma'>\gamma$, and set $\sigma^1_{(\beta,\gamma)}=\sigma^1_\beta$, $\rho^1_{(\beta,\gamma)}=\rho^1_\beta$, $\sigma^2_{(\beta,\gamma)}=\sigma^2_\gamma$, $\rho^2_{(\beta,\gamma)}=\rho^2_\gamma$. Then $((\sigma^i_\alpha,\rho^i_\alpha))_{\alpha\in I}\xrightarrow{w^*}(\sigma^i,\rho^i)$ for $i=1$, 2, and hence $((\lambda_1\sigma^1_\alpha+\lambda_2\sigma^2_\alpha,\lambda_1\rho^1_\alpha+\lambda_2\rho^2_\alpha))_{\alpha\in I}\xrightarrow{w^*}(\sigma,\rho)$.

Using the concavity of $G$, we obtain

$$\begin{aligned}H(\sigma,\rho)&\geqq\limsup_{\alpha\in I} G(\lambda_1\sigma^1_\alpha+\lambda_2\sigma^2_\alpha,\lambda_1\rho^1_\alpha+\lambda_2\rho^2_\alpha)\\&\geqq\limsup_{\alpha\in I}(\lambda_1 G(\sigma^1_\alpha,\rho^1_\alpha)+\lambda_2 G(\sigma^2_\alpha,\rho^2_\alpha))\\&\geqq\lambda_1 H(\sigma^1,\rho^1)+\lambda_2 H(\sigma^2,\rho^2)-2\varepsilon.\end{aligned}$$

Since $\varepsilon$ is arbitrary, this proves that $H$ is concave.

This argument was provided by the referee for whose assistance I am very grateful. It is much clearer than the argument I initially produced.

iii) $H(\sigma,\rho)=G(\sigma,\rho)$ for $\sigma$ and $\rho$ normal, since $G$ is $w^*$ upper semicontinuous. □

## 6. Definition and Properties of the Relative Entropy

Proposition 4.5, the fact that $\mathrm{ent}^S$ $(=\mathrm{ent}^{PW})$ satisfies Eq. 1.2, and Lemma 5.1 have shown that the following constitutes a definition of $\mathrm{ent}_{\mathscr{A}}(\sigma|\rho)$:

*Definition.*

6.1. $$\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho)=\sum_{i,j}(-s_j\log s_j+s_j\log r_i)|\langle\phi_j|\psi_i\rangle|^2$$

for $\sigma=\sum_j s_j\,|\phi_j\rangle\langle\phi_j|$ and $\rho=\sum_i r_i\,|\psi_i\rangle\langle\psi_i|$ normal states on $\mathscr{B}(\mathscr{H})$.

6.2. $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho)=\inf\{F(\sigma,\rho)$: $F$ is $w^*$ upper semicontinuous, concave, and given by (6.1) for $\sigma$ and $\rho$ normal$\}$.

6.3. $\mathrm{ent}_{\mathscr{A}}(\sigma|\rho)=\sup\{\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho')$: $\sigma'|_{\mathscr{A}}=\sigma$ and $\rho'|_{\mathscr{A}}=\rho\}$.

In this section the various properties of the relative entropy thus defined will be verified. It is clear that if the axioms are consistent then this is the unique relative entropy that they define. They are consistent since Axioms I and III obviously hold for ent as defined above, and Axioms II and IV will be shown to hold below. It should be noted that even were $\mathrm{ent}_{\mathscr{A}}(\sigma|\rho)=\mathrm{ent}^S_{\mathscr{A}}(\sigma|\rho)$ whenever the latter is defined, most of the results of this section would still be extensions of known results, since ent is a generalisation of $\mathrm{ent}^S$.

**a) Monotonicity.** *If* $\mathscr{A}_1\subset\mathscr{A}_2$, *then* $\mathrm{ent}_{\mathscr{A}_1}(\sigma|_{\mathscr{A}_1}|\rho|_{\mathscr{A}_1})\geqq\mathrm{ent}_{\mathscr{A}_2}(\sigma|\rho)$.

*Proof.* Immediate from Definition 6.3. □

**b) Concavity.** *For* $0\leqq\lambda_1,\lambda_2\leqq 1$ *and* $\lambda_1+\lambda_2=1$.

$$\mathrm{ent}_{\mathscr{A}}(\lambda_1\sigma_1+\lambda_2\sigma_2|\lambda_1\rho_1+\lambda_2\rho_2)\geqq\lambda_1\,\mathrm{ent}_{\mathscr{A}}(\sigma_1|\rho_1)+\lambda_2\,\mathrm{ent}_{\mathscr{A}}(\sigma_2|\rho_2).$$

*Proof.* By Proposition 4.5 and Lemma 5.1 this holds for $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho)$. In general, let $\sigma=\lambda_1\sigma_1+\lambda_2\sigma_2, \rho=\lambda_1\rho_1+\lambda_2\rho_2$ for $\lambda_1, \lambda_2$ as given.

$$\begin{aligned}\mathrm{ent}_{\mathscr{A}}(\sigma|\rho) &= \sup\{\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho'): \sigma' \text{ and } \rho' \text{ are states on } \mathscr{B}(\mathscr{H}) \\ &\qquad \text{with } \sigma'|_{\mathscr{A}}=\sigma \text{ and } \rho'|_{\mathscr{A}}=\rho\} \\ &\geqq \sup\{\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\lambda_1\sigma'_1+\lambda_2\sigma'_2|\lambda_1\rho'_1+\lambda_2\rho'_2): \sigma'_1, \sigma'_2, \rho'_1, \rho'_2 \text{ are states} \\ &\qquad \text{on } \mathscr{B}(\mathscr{H}) \text{ with } \sigma'_1|_{\mathscr{A}}=\sigma_1, \sigma'_2|_{\mathscr{A}}=\sigma_2, \rho'_1|_{\mathscr{A}}=\rho_1, \rho'_2|_{\mathscr{A}}=\rho_2\} \\ &\geqq \sup\{\lambda_1\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'_1|\rho'_1)+\lambda_2\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'_2|\rho'_2): \sigma'_1, \sigma'_2, \rho'_1, \rho'_2 \text{ are states} \\ &\qquad \text{on } \mathscr{B}(\mathscr{H}) \text{ with } \sigma'_1|_{\mathscr{A}}=\sigma_1, \sigma'_2|_{\mathscr{A}}=\sigma_2, \rho'_1|_{\mathscr{A}}=\rho_1, \rho'_2|_{\mathscr{A}}=\rho_2\} \\ &= \lambda_1\,\mathrm{ent}_{\mathscr{A}}(\sigma_1|\rho_1)+\lambda_2\,\mathrm{ent}_{\mathscr{A}}(\sigma_2|\rho_2).\end{aligned}$$

□

**Lemma 6.4** *For any* $(\sigma, \rho, \mathscr{A})$ *there exist* $\sigma'$, $\rho'$ *on* $\mathscr{B}(\mathscr{H})$ *such that* $\sigma'|_{\mathscr{A}}=\sigma$, $\rho'|_{\mathscr{A}}=\rho$ *and such that* $\mathrm{ent}_{\mathscr{A}}(\sigma|\rho)=\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho')$.

*Proof.* Let $(\sigma'_\alpha, \rho'_\alpha)_{\alpha\in I}$ be a net of states on $\mathscr{B}(\mathscr{H})$ such that $\sigma'_\alpha|_{\mathscr{A}}=\sigma$, $\rho'_\alpha|_{\mathscr{A}}=\rho$ for all $\alpha\in I$ and such that $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'_\alpha|\rho'_\alpha)\to\mathrm{ent}_{\mathscr{A}}(\sigma|\rho)$. Such a net certainly exists. After passing to a subnet, we may assume that $((\sigma'_\alpha, \rho'_\alpha))_{\alpha\in I}\overset{w^*}{\to}(\sigma', \rho')$. Of course, $\sigma'|_{\mathscr{A}}=\sigma$, $\rho'|_{\mathscr{A}}=\rho$.

By Proposition 4.5 and Lemma 5.1 $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}$ is $w^*$ upper semicontinuous, so

$$\mathrm{ent}_{\mathscr{A}}(\sigma|\rho)\geqq\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho')\geqq\limsup_{\alpha\in I}\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'_\alpha|\rho'_\alpha)=\mathrm{ent}_{\mathscr{A}}(\sigma|\rho).$$

**c) $w^*$ Upper Semicontinuity.** *For any net* $((\sigma_\alpha, \rho_\alpha))_{\alpha\in I}$ *defined on* $\mathscr{A}$ *with* $\sigma_\alpha(A)\to\sigma(A)$ *and* $\rho_\alpha(A)\to\rho(A)$ *for all* $A\in\mathscr{A}$,

$$\mathrm{ent}_{\mathscr{A}}(\sigma|\rho)\geqq\limsup_{\alpha\in I}\mathrm{ent}_{\mathscr{A}}(\sigma_\alpha|\rho_\alpha).$$

*Proof.* As just noted this holds for $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho)$. For the general case suppose that for some net $((\sigma_\alpha, \rho_\alpha))_{\alpha\in I}\overset{w^*}{\to}(\sigma, \rho)$ we have

$$\mathrm{ent}_{\mathscr{A}}(\sigma|\rho)<\limsup_{\alpha\in I}\mathrm{ent}_{\mathscr{A}}(\sigma_\alpha|\rho_\alpha). \tag{6.5}$$

Defining $((\sigma'_\alpha, \rho'_\alpha))_{\alpha\in I}$ on $\mathscr{B}(\mathscr{H})$ by Lemma 6.4 (i.e. $\sigma'_\alpha|_{\mathscr{A}}=\sigma_\alpha$, $\rho'_\alpha|_{\mathscr{A}}=\rho_\alpha$, $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'_\alpha|\rho'_\alpha)=\mathrm{ent}_{\mathscr{A}}(\sigma_\alpha|\rho_\alpha)$) there exists, by $w^*$ compactness, a $w^*$ convergent subnet $((\sigma'_\beta, \rho'_\beta))_{\beta\in J}$ of $((\sigma'_\alpha, \rho'_\alpha))_{\alpha\in I}$ chosen so as still to satisfy Inequality 6.5. Suppose that $(\sigma'_\beta, \rho'_\beta)\to(\sigma', \rho')$. Then $\sigma'|_{\mathscr{A}}=\sigma$, $\rho'|_{\mathscr{A}}=\rho$, so

$$\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho')\geqq\limsup_{\beta\in J}\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'_\beta|\rho'_\beta)=\limsup_{\beta\in J}\mathrm{ent}_{\mathscr{A}}(\sigma_\beta|\rho_\beta)>\mathrm{ent}_{\mathscr{A}}(\sigma|\rho).$$

This contradicts Definition 6.3. □

**Example 6.6.** Let $\{\psi_n: n\geqq 1\}$ be an orthonormal basis for a Hilbert space $\mathscr{H}$, and let

$$A=|\psi_1\rangle\langle\psi_1|+|\psi_2\rangle\langle\psi_2|+\sum_{n=3}^{\infty}\left(1-\frac{1}{n}\right)|\psi_n\rangle\langle\psi_n|,$$

$$B=|\psi_1\rangle\langle\psi_1|-\sum_{n=3}^{\infty}|\psi_n\rangle\langle\psi_n| \text{ and } \mathscr{A}=\{A,B,1\}.$$

Then
1) $\sigma'$ is a state on $\mathscr{H}$ such that $\sigma'(A)=\sigma'(B)=1$ if and only if $\sigma'=|\psi_1\rangle\langle\psi_1|$.
2) $\rho'$ is a *normal* state on $\mathscr{H}$ such that $\rho'(A)=1$, $\rho'(B)=0$ if and only if $\rho'=|\psi_2\rangle\langle\psi_2|$.
3) $\rho_n'(A)=1-1/2n$, $\rho_n'(B)=0$ is satisfied both by

$$\rho_n'=\frac{7}{4n}|\psi_1\rangle\langle\psi_1|+\left(1-\frac{7}{2n}\right)|\psi_2\rangle\langle\psi_2|+\frac{3}{4n}|\psi_3\rangle\langle\psi_3|+\frac{1}{n}|\psi_4\rangle\langle\psi_4|$$

and by

$$\rho_n''=\tfrac{1}{2}|\psi_1\rangle\langle\psi_1|+\tfrac{1}{2}|\psi_n\rangle\langle\psi_n|.$$

Now

$$\operatorname{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho')=-\infty,\quad \operatorname{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho_n')=-\log\left(\frac{4n}{7}\right),$$

and

$$\operatorname{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho_n'')=-\log 2.$$

This shows that $w^*$ upper semicontinuity does not hold if Axiom I is replaced by

$$\operatorname{ent}^c_{\mathscr{A}}(\sigma|\rho)=\sup\{\operatorname{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho')\colon \sigma' \text{ and } \rho' \text{ are normal states on } \mathscr{B}(\mathscr{H}) \text{ with } \sigma'|_{\mathscr{A}}=\sigma \text{ and } \rho'|_{\mathscr{A}}=\rho\}.$$ □

The consequence of this example is most of the analytical details in this paper.

### d) Axiom IV Holds

*Proof.* Given $(\sigma,\rho,\mathscr{A})$, define $(\sigma',\rho')$ as in Lemma 6.4. A net $((\sigma_\alpha',\rho_\alpha'))_{\alpha\in I}$ of normal states such that $(\sigma_\alpha',\rho_\alpha')\to(\sigma',\rho')$ satisfying $\lim_{\alpha\in I}\operatorname{ent}_{\mathscr{B}(\mathscr{H})}(\sigma_\alpha'|\rho_\alpha')=\operatorname{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho')$ exists by Lemma 5.1. Then by $w^*$ upper semicontinuity and monotonicity,

$$\begin{aligned}\operatorname{ent}_{\mathscr{A}}(\sigma|\rho)&\geqq\limsup_{\alpha\in I}\operatorname{ent}_{\mathscr{A}}(\sigma_\alpha|\rho_\alpha)\geqq\liminf_{\alpha\in I}\operatorname{ent}_{\mathscr{A}}(\sigma_\alpha|\rho_\alpha)\\&\geqq\liminf_{\alpha\in I}\operatorname{ent}_{\mathscr{B}(\mathscr{H})}(\sigma_\alpha'|\rho_\alpha')=\operatorname{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho')=\operatorname{ent}_{\mathscr{A}}(\sigma|\rho),\end{aligned}$$

where

$$\sigma_\alpha=\sigma_\alpha'|_{\mathscr{A}} \quad\text{and}\quad \rho_\alpha=\rho_\alpha'|_{\mathscr{A}}.$$ □

**Lemma 6.7.** $\operatorname{ent}_{\mathscr{A}}(\sigma|\rho)\leqq\operatorname{ent}^S_{\mathscr{A}}(\sigma|\rho)$ *whenever the latter is defined.*

*Proof.* By Proposition 4.5 and Definition 6.2, we always have

$$\operatorname{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho)\leqq\operatorname{ent}^S_{\mathscr{B}(\mathscr{H})}(\sigma|\rho).$$

Now let $\sigma$ and $\rho$ be defined on a von Neumann algebra $\mathscr{A}$. Define $\sigma'$ and $\rho'$ as in

Lemma 6.4. Then, using the monotonicity of $\mathrm{ent}^S$ (which is immediate from Lemma 4.4, or see [11, 13]) $\mathrm{ent}_{\mathscr{A}}(\sigma|\rho)=\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho')\leqq \mathrm{ent}^S_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho')\leqq \mathrm{ent}^S_{\mathscr{A}}(\sigma|\rho)$. □

e) **Property V of Sect. 3.** *Let $\sigma$ and $\rho$ be normal states defined on the abelian algebra $\mathscr{Z}$ generated by the mutually orthogonal projections $(P_n)_{n=1}^{\infty}$ with $\sum_{n=1}^{\infty} P_n=1$. Then $\mathrm{ent}_{\mathscr{Z}}(\sigma|\rho)=-\sum_{n=1}^{\infty}\sigma(P_n)\log(\sigma(P_n)/\rho(P_n))$.*

*Proof.* It is easy to check directly from the definition of $\mathrm{ent}^{PW}$ that $\mathrm{ent}^S_{\mathscr{Z}}(\sigma|\rho)=-\sum_{n=1}^{\infty}\sigma(P_n)\log(\sigma(P_n)/\rho(P_n))$ (or see [8]). Defining $\hat{\sigma}$ and $\hat{\rho}$ as density matrices on $\mathscr{B}(\mathscr{H})$ by $\hat{\sigma}=\sum_n \sigma(P_n)P_n$ and $\hat{\rho}=\sum_n \rho(P_n)P_n$ we have $\hat{\sigma}|_{\mathscr{Z}}=\sigma$, $\hat{\rho}|_{\mathscr{Z}}=\rho$ and $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\hat{\sigma}|\hat{\rho})=-\sum_{n=1}^{\infty}\sigma(P_n)\log(\sigma(P_n)/\rho(P_n))$ by Definition 6.1. The result now follows from Definition 6.3 and Lemma 6.7. □

**Lemma 6.8.** *If $\sigma$ and $\rho$ are states on a von Neumann algebra $\mathscr{A}$, then $\mathrm{ent}_{\mathscr{A}}(\sigma|\rho)\leqq -\frac{1}{2}\|\sigma-\rho\|^2$*

*Proof.* By Lemma 6.7 this is a consequence of the corresponding result for $\mathrm{ent}^S$. This is proved for $\sigma$ and $\rho$ normal in [7] (Theorem 3.1), and can be extended to general $\sigma$ and $\rho$ using Lemma 3.1 of [8]. □

f) **Non-Triviality.** *For arbitrary $(\sigma,\rho,\mathscr{A})$, if $\sigma\neq\rho$, then $\mathrm{ent}_{\mathscr{A}}(\sigma|\rho)<0$. More specifically.*

$$\mathrm{ent}_{\mathscr{A}}(\sigma|\rho)\leqq \inf\{-|\sigma(A)-\rho(A)|^2/2\|A\|^2 : A\in\mathscr{A}\}.$$

*Proof.* Let $\sigma'$ and $\rho'$ be states on $\mathscr{B}(\mathscr{H})$ with $\sigma'|_{\mathscr{A}}=\sigma$ and $\rho'|_{\mathscr{A}}=\rho$. Then, for $A\in\mathscr{A}$,

$$|\sigma(A)-\rho(A)|=|\sigma'(A)-\rho'(A)|\leqq\|\sigma'-\rho'\|\,\|A\|,$$

so the result is a consequence of Definition 6.3 and Lemma 6.8. □

g) **Axiom II Holds.**

*Proof.* Let $\{\sigma_j\}$, $\sigma$, and $\rho$ be as in the axiom. Since the $\sigma_j$ have disjoint supports, we may write $\sigma_j=\sum_k q_{jk}\,|\phi_{jk}\rangle\langle\phi_{jk}|$, where the set $\{\phi_{jk}: j=1,2,3,\dots,k=1,2,3,\dots\}$ forms an orthonormal basis of $\mathscr{H}$. Then applying Definition 6.1,

$$\begin{aligned}\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho)&=\sum_i\sum_j\sum_k(-s_jq_{jk}\log s_jq_{jk}+s_jq_{jk}\log r_i)|\langle\phi_{jk}|\psi_i\rangle|^2\\&=\sum_i\sum_j\sum_k s_j(-q_{jk}(\log s_j+\log q_{jk})+q_{jk}\log r_i)|\langle\phi_{jk}|\psi_i\rangle|^2\\&=\sum_j(-s_j\log s_j+s_j\,\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma_j|\rho)).\end{aligned}$$
□

Using [10, Theorem 3.6], one can prove more generally that for all von

Neumann algebras $\mathscr{A}$ and all states $\sigma$ and $\rho$ on $\mathscr{A}$ with $\sigma = \sum_j s_j \sigma_j$, where the $\sigma_j$ are states on $\mathscr{A}$ with disjoint support, we have $\mathrm{ent}^S_{\mathscr{A}}(\sigma|\rho) = \sum_j (-s_j \log s_j + s_j \mathrm{ent}^S_{\mathscr{A}}(\sigma_j|\rho))$. I do not know whether this equation holds in this generality for ent. In particular I do not even know if it holds for non-normal states on $\mathscr{B}(\mathscr{H})$. The following example, the details of which are left as an exercise for the very serious reader, shows that it certainly does not generalise indefinitely:

**Example 6.9.** Let $\mathscr{H} = \mathbb{C}^4$ and choose $\mathscr{A}$ so that $\sigma^1$, $\sigma^2$, and $\rho$ are determined on $\mathscr{A}$ by the following relations on the components of the corresponding density matrices on $\mathbb{C}^4$:

$$\sigma^1_{11} + \sigma^1_{22} = 1, \quad \sigma^1_{12} + \sigma^1_{34} = 0, \quad \sigma^1_{13} = \sigma^1_{14} = \sigma^1_{23} = \sigma^1_{24} = 0,$$
$$\sigma^2_{33} + \sigma^2_{44} = 1, \quad \sigma^2_{12} + \sigma^2_{34} = 0, \quad \sigma^2_{13} = \sigma^2_{14} = \sigma^2_{23} = \sigma^2_{24} = 0,$$
$$\rho_{11} + \rho_{22} = \tfrac{1}{2}, \quad \rho_{12} + \rho_{34} = \tfrac{1}{4}, \quad \rho_{13} = \rho_{14} = \rho_{23} = \rho_{24} = 0,$$

(also, of course, $\sigma^1(1) = \sigma^2(1) = \rho(1) = 1$ and all the other relations required for $\sigma^1$, $\sigma^2$, and $\rho$ to extend to states on $\mathbb{C}^4$). Define $\sigma = \frac{1}{2}\sigma^1 + \frac{1}{2}\sigma^2$. Then I claim that

$$\mathrm{ent}_{\mathscr{A}}(\sigma^1|\rho) = \mathrm{ent}_{\mathscr{A}}(\sigma^2|\rho) = -\log 2, \text{ while } \mathrm{ent}_{\mathscr{A}}(\sigma|\rho) = -\tfrac{1}{2}\log\tfrac{4}{3}.$$

Thus

$$-\tfrac{1}{2}\log\tfrac{1}{2} - \tfrac{1}{2}\log\tfrac{1}{2} + \tfrac{1}{2}\mathrm{ent}_{\mathscr{A}}(\sigma^1|\rho) + \tfrac{1}{2}\mathrm{ent}_{\mathscr{A}}(\sigma^2|\rho) = 0 \neq \mathrm{ent}_{\mathscr{A}}(\tfrac{1}{2}\sigma^1 + \tfrac{1}{2}\sigma^2|\rho). \qquad \square$$

**h) Uhlmann's Inequality.** *Let $\lambda: \mathscr{A}_1 \to \mathscr{A}_2$ be such that there is a normal linear map $\lambda'$: $\mathscr{B}(\mathscr{H}_1) \to \mathscr{B}(\mathscr{H}_2)$ (where $\mathscr{A}_1 \subset \mathscr{B}(\mathscr{H}_1)$ and $\mathscr{A}_2 \subset \mathscr{B}(\mathscr{H}_2)$) with $\lambda'(B^*) = \lambda'(B)^*$, $\lambda'(B)^*\lambda'(B) \leqq \lambda'(B^*B)$ for all $B \in \mathscr{B}(\mathscr{H}_1)$, $\lambda'|_{\mathscr{A}_1} = \lambda$, and $\lambda(1) = 1$. Then, for all $\sigma$ and $\rho$ on $\mathscr{A}_2$, $\mathrm{ent}_{\mathscr{A}_1}(\sigma\circ\lambda|\rho\circ\lambda) \geqq \mathrm{ent}_{\mathscr{A}_2}(\sigma|\rho)$.*

*Proof.* Uhlmann proved [13, Prop. 18] that if $\mathscr{A}_1$, $\mathscr{A}_2$ are *-algebras and $\lambda: \mathscr{A}_1 \to \mathscr{A}_2$ with $\lambda(A^*) = \lambda(A)^*$, $\lambda(A)^*\lambda(A) \leqq \lambda(A^*A)$ $(A \in \mathscr{A}_1)$, $\lambda(1) = 1$, then $\mathrm{ent}^S_{\mathscr{A}_1}(\sigma\circ\lambda|\rho\circ\lambda) \geqq \mathrm{ent}^S_{\mathscr{A}_2}(\sigma|\rho)$. Unlike Uhlmann I need to assume that $\lambda$ is normal, and I do not know if property *h* holds if this assumption is dropped. However, the results of Lindblad [6] are a special case of the present result.

Assuming the hypothesis, let $((\sigma'_\alpha, \rho'_\alpha))_{\alpha\in I}$ $(\overset{w^*}{\to}(\sigma', \rho'))$ be a net of pairs of normal states on $\mathscr{B}(\mathscr{H}_2)$ constructed as in Axiom IV applied to $(\sigma, \rho, \mathscr{A}_2)$. Then $((\sigma'_\alpha\circ\lambda', \rho'_\alpha\circ\lambda'))_{\alpha\in I}$ is a net of pairs of normal states on $\mathscr{B}(\mathscr{H}_1)$,

$$(\sigma'_\alpha\circ\lambda', \rho'_\alpha\circ\lambda') \overset{w^*}{\to} (\sigma'\circ\lambda', \rho'\circ\lambda'), \quad \text{and} \quad (\sigma'\circ\lambda'|_{\mathscr{A}_1}, \rho'\circ\lambda'|_{\mathscr{A}_1}) = (\sigma\circ\lambda, \rho\circ\lambda).$$

Now, using Uhlmann's result (for normal states on $\mathscr{B}(\mathscr{H})$ ent $= \mathrm{ent}^S$),

$$\mathrm{ent}_{\mathscr{A}_1}(\sigma\circ\lambda|\rho\circ\lambda) \geqq \mathrm{ent}_{\mathscr{B}(\mathscr{H}_1)}(\sigma'\circ\lambda'|\rho'\circ\lambda') \qquad \text{(property a)}$$
$$\geqq \limsup_{\alpha\in I} \mathrm{ent}_{\mathscr{B}(\mathscr{H}_1)}(\sigma'_\alpha\circ\lambda'|\rho'_\alpha\circ\lambda') \qquad \text{(property c)}$$
$$\geqq \limsup_{\alpha\in I} \mathrm{ent}_{\mathscr{B}(\mathscr{H}_2)}(\sigma'_\alpha|\rho'_\alpha) = \mathrm{ent}_{\mathscr{A}_2}(\sigma|\rho). \qquad \square$$

There remains but one major property of $\mathrm{ent}^S$ that we have not discussed and generalised in some form to ent. That property "Araki's property" is essentially equivalent to the equivalence theorem to be proved in the remaining sections. It is therefore deferred to the end of Sect. 8.

## 7. A Property Equivalent to Injectivity

This section is a brief technical interlude on the theory of von Neumann algebras.

*Definition 7.1.* A von Neumann algebra $(\mathscr{A}, \mathscr{H})$ has Property 7.1 if there exists a net $(\lambda_\alpha)_{\alpha\in I}$ of normal completely positive finite rank maps $\lambda_\alpha: \mathscr{B}(\mathscr{H}) \to \mathscr{A}$ such that $\lambda_\alpha(1) = 1$ and such that $\lambda_\alpha|_{\mathscr{A}}$ tends to the identity map on $\mathscr{A}$ in the topology of simple $w^*$ convergence. (I.e. $\omega(\lambda_\alpha(A)) \to \omega(A)$ for all $\omega \in \mathscr{A}_*$, $A \in \mathscr{A}$).

In the next section we shall see that for $\sigma$ and $\rho$ normal states on an algebra $\mathscr{A}$ with Property 7.1 we have $\mathrm{ent}_{\mathscr{A}}(\sigma|\rho) = \mathrm{ent}^S_{\mathscr{A}}(\sigma|\rho)$. In this section we shall prove that injectivity is equivalent to Property 7.1. Although we will not actually need our maps $\lambda_\alpha$ to have finite rank, the proof with this property omitted is not significantly simpler. For an introduction to the theory of injective algebras see [16, §10.22–10.31, pp 143–149].

*Definition 7.2.* A von Neumann algebra $(\mathscr{A}, \mathscr{H})$ is semidiscrete if there exists a net $(\lambda_\alpha)_{\alpha\in I}$ of normal completely positive finite rank maps $\lambda_\alpha: \mathscr{A} \to \mathscr{A}$ such that $\lambda_\alpha(1) = 1$ and such that $\lambda_\alpha$ tends to the identity map on $\mathscr{A}$ in the topology of simple $w^*$ convergence.

This property was introduced by Effros and Lance [17]. For its equivalence with injectivity see [17, 18]. In fact, Effros and Lance showed that semidiscreteness of $\mathscr{A}$ is equivalent to various properties of tensor products of $\mathscr{A}$. In this section we will show that Property 7.1 and these properties are equivalent by showing that we can go through Effros and Lance's paper carefully and in detail making the necessary extensions as we go, to show that they could have worked throughout with Property 7.1 instead of semidiscreteness. The first result however is not a direct translation of their work:

**Proposition 7.3.** *If $\tau$: $(\mathscr{A}_1, \mathscr{H}_1) \to (\mathscr{A}_2, \mathscr{H}_2)$ is an isomorphism of von Neumann algebras, then $(\mathscr{A}_1, \mathscr{H}_1)$ has Property 7.1 if and only if $(\mathscr{A}_2, \mathscr{H}_2)$ has Property 7.1.*

*Proof.* Suppose that $(\mathscr{A}_2, \mathscr{H}_2)$ has Property 7.1 and that $(\lambda_\alpha: \mathscr{B}(\mathscr{H}_2) \to \mathscr{A}_2)_{\alpha\in I}$ is the relevant net of maps. It is sufficient by the standard canonical form theorem for von Neumann algebra isomorphisms (e.g. [19, théorème I.4.3, p. 55]) to treat the following special cases:
i) $\tau$ is a spatial isomorphism: $\tau(A) = UAU^*$ for $U: \mathscr{H}_1 \to \mathscr{H}_2$ a surjective isometry. In this case $\lambda'_\alpha(B) = U^* \lambda_\alpha(UBU^*)\, U$ $(B \in \mathscr{B}(\mathscr{H}_1))$ is a suitable net of maps on $\mathscr{B}(\mathscr{H}_1)$.
ii) There is a Hilbert space $\mathscr{K}$ such that $\tau: (\mathscr{A}, \mathscr{H}) \to (\mathscr{A} \bar{\otimes} 1_{\mathscr{K}}, \mathscr{H} \otimes \mathscr{K})$ is the amplification. In this case define

$$\lambda'_\alpha(B) = \tau^{-1}(\lambda_\alpha(B \otimes 1_{\mathscr{K}})) \quad \text{for} \quad B \in \mathscr{B}(\mathscr{H}).$$

iii) There is a projection $e' \in \mathscr{A}'$ with central support 1 such that $\tau$: $(\mathscr{A}, \mathscr{H}) \to (\mathscr{A}_{e'}, e'\mathscr{H})$ is the canonical induction. In this case define $\lambda'_\alpha(B) = \tau^{-1}(\lambda_\alpha(B_{e'}))$ for all

$B\in\mathscr{B}(\mathscr{H})$. The confirmation in each of these cases that the $\lambda'_\alpha$ are suitable, is easy. □

Now we need to prove the following succession of results:

A) *If* $(\mathscr{A}_\beta, \mathscr{H}_\beta)$ *has Property 7.1 for each* $\beta\in J$, *then so does* $(\bigoplus_{\beta\in J}\mathscr{A}_\beta, \bigoplus_{\beta\in J}\mathscr{H}_\beta)$.

B) *If for all normal states* $\rho$ *on* $\mathscr{A}$, $(\pi_\rho(\mathscr{A}), \mathscr{H}_\rho)$ *has Property 7.1, then so does* $(\mathscr{A}, \mathscr{H})$. (*Here* $\pi_\rho$ *is the* G.N.S. *representation of* $\mathscr{A}$ *corresponding to* $\rho$).

C) *If* $(\mathscr{A}_1, \mathscr{H}_1)$ *and* $(\mathscr{A}_2, \mathscr{H}_2)$ *have Property 7.1, then so does* $(\mathscr{A}_1\bar{\otimes}\mathscr{A}_2, \mathscr{H}_1\otimes\mathscr{H}_2)$.

D) *If* $(\mathscr{A}, \mathscr{H})$ *has Property 7.1 and* $e\in\mathscr{A}$ *is a projection, then* $(e\mathscr{A}e, e\mathscr{H})$ *has Property 7.1.*

E) *If* $(\mathscr{A}, \mathscr{H})$ *has Property 7.1, then so does the commutant* $(\mathscr{A}', \mathscr{H})$.

These results all correspond to parts of results in Sect. 3 of [17], and in each case an almost identical proof works, except that the proof of $E$ needs Proposition 7.3.

It follows from $B$ and $E$ that to prove that $(\mathscr{A}, \mathscr{H})$ has Property 7.1, it is sufficient to show that for each normal state $\rho$ on $\mathscr{A}$, $(\pi_\rho(\mathscr{A})', \mathscr{H}_\rho)$ has Property 7.1. It is for such algebras that, using the theory of tensor products, we actually construct the maps $\lambda_\alpha$.

Thus, let $(\mathscr{A}, \mathscr{H})$ be an injective von Neumann algebra, and let $\rho$ be a normal state on $\mathscr{A}$. We now must assume the notation of [17, Sects. 1 and 2]. In particular, for von Neumann algebras $\mathscr{A}$ and $\mathscr{B}$, Effros and Lance define sets $\Gamma=\min(\mathscr{A}\otimes\mathscr{B})$, $\mathrm{bin}(\mathscr{A}\otimes\mathscr{B})$ of positive linear functionals on $\mathscr{A}\otimes\mathscr{B}$, and corresponding $C^*$ algebra tensor products $\mathscr{A}\otimes_\Gamma\mathscr{B}$. $S_\Gamma(\mathscr{A}\otimes\mathscr{B})$ is the set of restrictions to $\mathscr{A}\otimes\mathscr{B}$ of states on $\mathscr{A}\otimes_\Gamma\mathscr{B}$. $\theta_\rho:\pi_\rho(\mathscr{A})'\to\mathscr{A}_*$ is defined by $\theta_\rho(C)(A)=(\Omega_\rho, \pi_\rho(A)C_{\Omega\rho})$ for $C\in\pi_\rho(\mathscr{A})'$ and $A\in\mathscr{A}$, and there is a corresponding positive linear functional $\sigma$ on $\mathscr{A}\otimes\pi_\rho(\mathscr{A})'$ defined by $\sigma(A\otimes C)=\theta_\rho(C)(A)$.

$\sigma\in\mathrm{bin}(\mathscr{A}\otimes\pi_\rho(\mathscr{A})')$ and so, since $\mathscr{A}$ is semidiscrete, $\sigma\in S_{\min}(\mathscr{A}\otimes\pi_\rho(\mathscr{A})'$ ([17, Theorem 4.1]). Thus $\sigma$ has a unique extension to a state $\bar{\sigma}$ on $\mathscr{A}\otimes_{\min}\pi_\rho(\mathscr{A})'$. Since this is a subalgebra of $\mathscr{A}\otimes_{\min}\mathscr{B}(\mathscr{H}_\rho)$, $\bar{\sigma}$ has an extension to a state $\bar{\bar{\sigma}}$ on $\mathscr{A}\otimes_{\min}\mathscr{B}(\mathscr{H}_\rho)$. Set $\sigma'=\bar{\bar{\sigma}}|_{\mathscr{A}\otimes\mathscr{B}(\mathscr{H}_\rho)}\in S_{\min}(\mathscr{A}\otimes\mathscr{B}(\mathscr{H}_\rho))$.

Now we follow the proof of (ii)⇒(i) of Theorem 4.1 of [17]. Since $S_{\min}(\mathscr{A}\otimes\mathscr{B}(\mathscr{H}_\rho))=\overline{\min(\mathscr{A}\otimes\mathscr{B}(\mathscr{H}_\rho))\cap\mathrm{bin}(\mathscr{A}\otimes\mathscr{B}(\mathscr{H}_\rho))}^{wd}$, there exists a net $(\sigma'_\alpha)_{\alpha\in I}\subset\min(\mathscr{A}\otimes\mathscr{B}(\mathscr{H}_\rho))\cap\mathrm{bin}(\mathscr{A}\otimes\mathscr{B}(\mathscr{H}_\rho))$, such that $T_{\sigma'_\alpha}\to T_{\alpha'}$ in the topology of simple $w^*$ convergence, where $T_{\sigma'_\alpha}$ (respectively $T_{\sigma'}$): $\mathscr{B}(\mathscr{H}_\rho)\to\mathscr{A}^*$ are the maps defined by $T_{\sigma'_\alpha}(B)(A)=\sigma'_\alpha(A\otimes B)$ (respectively $T_{\sigma'}(B)(A)=\sigma'(A\otimes B)$) for $B\in\mathscr{B}(\mathscr{H}_\rho)$ and $A\in\mathscr{A}$.

However, $\sigma'$ need not belong to $\mathrm{bin}(\mathscr{A}\otimes\mathscr{B}(\mathscr{H}_\rho))$, so we cannot apply [17, Lemmata 4.2–4.4] directly. Instead, let $\sigma_\alpha=\sigma'_\alpha|_{\mathscr{A}\otimes\pi_\rho(\mathscr{A})'}$, and let $\Lambda=\mathrm{co}\{T_{\sigma_\alpha}:\alpha\in I\}$ considered as a set of maps from $\pi_\rho(\mathscr{A})'$ to $\mathscr{A}_*$.

$\Lambda$ has the same closure ($\bar{\Lambda}$) in the topologies of simple weak and simple norm convergence, and $\theta_\rho\in\bar{\Lambda}$.

Now [17, Lemmata 4.2–4.4] can be applied and show that there exists a net $(\omega_\beta)_{\beta\in J}\subset\Lambda$ such that $T_{\omega_\beta}\to\theta_\rho$ (simple norm convergence), and such that $T_{\omega_\beta}(1)=\rho$. Moreover, each $\omega_\beta$ has the form

$$\omega_\beta(A\otimes C)=\sum_{i=1}^{I}p_i(\sigma_{\alpha_i}(t_iAt_i\otimes C)+f^i_-(t_iAt_i)\tau_i(C))\quad\text{for}\quad A\in\mathscr{A}\quad\text{and}\quad C\in\pi_\rho(\mathscr{A})',$$

where $p_i \in [0, 1]$ with $\sum_{i=1}^{I} p_i = 1$, $t_i \in \mathcal{A}$ with $0 \leqq t_i \leqq 1$, $f_-^i \in \mathcal{A}_{*,+}$, and $\tau_i$ is a normal state on $\pi_\rho(\mathcal{A})'$.

Let $\tau_i'$ be any extension of $\tau_i$ to a normal state on $\mathcal{B}(\mathcal{H}_\rho)$. Define

$$\omega_\beta'(A \otimes B) = \sum_{i=1}^{I} p_i(\sigma_{\alpha_i}'(t_i A t_i \otimes B) + f_-^i(t_i A t_i)\tau_i'(B))$$

for $A \in \mathcal{A}$ and $B \in \mathcal{B}(\mathcal{H}_\rho)$. Clearly $\omega_\beta'|_{\mathcal{A} \otimes \pi_\rho(\mathcal{A})'} = \omega_\beta$ and $\omega_\beta' \in \min(\mathcal{A} \otimes \mathcal{B}(\mathcal{H}_\rho)) \cap \operatorname{bin}(\mathcal{A} \otimes \mathcal{B}(\mathcal{H}_\rho))$.

Thus $T_{\omega_\beta'}(1) = \rho$, so that the map $\lambda_\beta = \theta_\rho^{-1} T_{\omega_\beta'}$ from $\mathcal{B}(\mathcal{H}_\rho) \to \pi_\rho(\mathcal{A})'$ is well defined. $\lambda_\beta$ is also completely positive, has finite rank, and $\lambda_\beta(1) = 1$ ([17, Lemma 1.5]).

Now for $B \in \mathcal{B}(\mathcal{H}_\rho)$ and $A_1, A_2 \in \mathcal{A}$,

$$(\pi_\rho(A_1)\Omega_\rho, \lambda_\beta(B)\pi_\rho(A_2)\Omega_\rho) = T_{\omega_\beta'}(B)(A_1^* A_2) = \omega_\beta'(A_1^* A_2 \otimes B)$$

from which it follows that $\lambda_\beta$ is normal, and for $C \in \pi_\rho(\mathcal{A})'$,

$$\begin{aligned}(\pi_\rho(A_1)\Omega_\rho, (\lambda_\beta(C) - C)\pi_\rho(A_2)\Omega_\rho) &= T_{\omega_\beta'}(C)(A_1^* A_2) - \theta_\rho(C)(A_1^* A_2) \\ &= T_{\omega_\beta}(C)(A_1^* A_2) - \theta_\rho(C)(A_1^* A_2),\end{aligned}$$

from which it follows that $\lambda_\beta$ tends to the identity map on $\pi_\rho(\mathcal{A})'$ in the topology of simple $w^*$ convergence.

This completes the proof that $\pi_\rho(\mathcal{A})'$ has Property 7.1 from which it follows, as discussed, that each injective algebra has Property 7.1, and finally, since Property 7.1 clearly implies semidiscreteness, it follows that injectivity and Property 7.1 are equivalent.

## 8. The Equivalence Theorem

There are three distinct parts to this section. Throughout it $\mathcal{A}$ will denote a von Neumann algebra. In the first part we will show that $\operatorname{ent}_{\mathcal{A}}(\sigma|\rho)$ is representation invariant, in the second that ent and $\operatorname{ent}^S$ always agree on abelian algebras, and in the third, by using the results of Sect. 7, that ent and $\operatorname{ent}^S$ agree on normal states on injective algebras.

*8A.* First we derive some consequences of Uhlmann's inequality—Property *h* of Sect. 6. In each case a more direct proof can be given by applying Axiom IV and making an appropriate explicit bound.

8.1. Let $e \in \mathcal{B}(\mathcal{H})$ be a projection. For $B \in \mathcal{B}(\mathcal{H})$, write $B_e = eB|_{\mathcal{B}(e\mathcal{H})} \in \mathcal{B}(e\mathcal{H})$ and write $\mathcal{A}_e = \{A_e : A \in \mathcal{A}\}$. Define $\lambda_1 : \mathcal{A} \to \mathcal{A}_e$ by $\lambda_1(A) = A_e$ and extend to $\lambda_1' : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(e\mathcal{H})$ by $\lambda_1'(B) = B_e$. Uhlmann's inequality applies ($\lambda_1'$ is completely positive), so $\operatorname{ent}_{\mathcal{A}}(\sigma \circ \lambda_1 | \rho \circ \lambda_1) \geqq \operatorname{ent}_{\mathcal{A}_e}(\sigma|\rho)$ for all $\sigma$ and $\rho$ on $\mathcal{A}_e$. Note that, in general, $\mathcal{A}_e$ need not be an algebra.

8.2. Now suppose $e \in \mathcal{A}$. Let $\tau$ be a normal state on $\mathcal{B}(e\mathcal{H})$. Define $\lambda_2 : \mathcal{A}_e \to \mathcal{A}$ by $\lambda_2(A) = eAe + (1-e)\tau(A)$, and extend to $\lambda_2' : \mathcal{B}(e\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ by $\lambda_2'(B) = eBe + (1-e)\tau(B)$. Again Uhlmann's inequality applies, so that

$$\operatorname{ent}_{\mathcal{A}_e}(\sigma \circ \lambda_2 | \rho \circ \lambda_2) \geqq \operatorname{ent}_{\mathcal{A}}(\sigma|\rho) \text{ for all } \sigma \text{ and } \rho \text{ on } \mathcal{A}.$$

8.3. Since $\lambda_1 \circ \lambda_2$ is the identity map on $\mathscr{A}_e$, combining 8.1 and 8.2 gives $\mathrm{ent}_{\mathscr{A}}(\sigma\circ\lambda_1 | \rho\circ\lambda_1) = \mathrm{ent}_{\mathscr{A}_e}(\sigma|\rho)$ for all $\sigma$ and $\rho$ on $\mathscr{A}_e$.
8.4. Let $\mathscr{K}$ be a Hilbert space, and $\lambda_4: \mathscr{A} \to \mathscr{A} \otimes 1_{\mathscr{K}}$ be the amplification defined by $\lambda_4(A) = A \otimes 1_{\mathscr{K}}$. Extend to $\lambda_4'$: $\mathscr{B}(\mathscr{H}) \to \mathscr{B}(\mathscr{H} \otimes \mathscr{K})$ by $\lambda_4'(B) = B \otimes 1_{\mathscr{K}}$. Again $\mathrm{ent}_{\mathscr{A}}(\sigma\circ\lambda_4|\rho\circ\lambda_4) \geqq \mathrm{ent}_{\mathscr{A}\otimes 1_{\mathscr{K}}}(\sigma|\rho)$ for all states $\sigma$ and $\rho$ on $\mathscr{A} \otimes 1_{\mathscr{K}}$.
8.5. The final case of a spatial isomorphism between von Neumann algebras is totally trivial.

An immediate corollary of 8.1 applied to the case $e \in \mathscr{A}'$, 8.4, 8.5, and the standard decomposition theorem ([19, théorème I.4.3 p55]) is:

**Theorem 8.6.** *Let $\tau$: $(\mathscr{A}_1, \mathscr{H}_1) \to (\mathscr{A}_2, \mathscr{H}_2)$ be a mapping between von Neumann algebras and $\sigma$ and $\rho$ be states on $\mathscr{A}_2$. If $\tau$ is a normal homomorphism of $\mathscr{A}_1$ onto $\mathscr{A}_2$, then*

$$\mathrm{ent}_{\mathscr{A}_1}(\sigma\circ\tau|\rho\circ\tau) \geqq \mathrm{ent}_{\mathscr{A}_2}(\sigma|\rho), \text{ and so if } \tau \text{ is an isomorphism}$$
$$\mathrm{ent}_{\mathscr{A}_1}(\sigma\circ\tau|\rho\circ\tau) = \mathrm{ent}_{\mathscr{A}_2}(\sigma|\rho). \qquad \square$$

As a consequence of 8.3, we prove inequality 3.1:

**Proposition 8.7.** *Let $\sigma$ and $\rho$ be states on $\mathscr{B}(\mathscr{H})$ and $(A_n)_{n=1}^N$ (with $N$ finite) be positive operators on $\mathscr{B}(\mathscr{H})$ with $\sum_{n=1}^N A_n = 1$.* Then

$$\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho) \leqq -\sum_{n=1}^N \sigma(A_n) \log \frac{\sigma(A_n)}{\rho(A_n)}.$$

*Proof.* First let $\mathscr{Z}$ be an abelian algebra generated by $N$ orthogonal projections $(P_n)_{n=1}^N$. $\sigma|_{\mathscr{Z}}$ and $\rho|_{\mathscr{Z}}$ are normal, so by monotonicity and property $e$ of Sect. 6,

$$\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho) \leqq -\sum_{n=1}^N \sigma(P_n) \log \frac{\sigma(P_n)}{\rho(P_n)}. \tag{8.8}$$

Now let $(A_n)_{n=1}^N$ be operators as in the hypothesis. Then by a theorem of Naimark (see [20, Theorem 9.3.2, p. 142]) there exists a Hilbert space $\mathscr{K}$ containing $\mathscr{H}$ as a subspace, with $e$: $\mathscr{K} \to \mathscr{H}$ the orthogonal projection, and mutually orthogonal projections

$$(P_n)_{n=1}^N \quad \text{such that} \quad A_n = eP_n|_{\mathscr{B}(e\mathscr{K})}. \tag{8.9}$$

By 8.3 applied to this projection $e \in \mathscr{B}(\mathscr{K})$, $\mathrm{ent}_{\mathscr{B}(\mathscr{K})}(\sigma\circ\lambda_1|\rho\circ\lambda_1) = \mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma|\rho)$. But 8.9 gives $\sigma\circ\lambda_1(P_n) = \sigma(A_n)$, $\rho\circ\lambda_1(P_n) = \rho(A_n)$, so applying 8.8 to $\sigma\circ\lambda_1$, $\rho\circ\lambda_1$ gives the result. $\square$

*8B.*

**Theorem 8.10.** *Let $\mathscr{Z}$ be an abelian von Neumann algebra, and let $\sigma$ and $\rho$ be states on $\mathscr{Z}$. Then* $\mathrm{ent}_{\mathscr{Z}}(\sigma|\rho) = \mathrm{ent}^S_{\mathscr{Z}}(\sigma|\rho)$.

*Proof.* Let $\Lambda$ be the directed set of sets $\{P_1, \ldots, P_N\}$ where, $N$ is finite, the $(P_n)_{n=1}^N$ are mutually orthogonal projections in $\mathscr{Z}$ with $\sum_{n=1}^N P_n = 1$, and the ordering is by refinement.

For $\ell = \{P_1, \ldots, P_N\} \in \Lambda$, define $\sigma'_\ell$ as the normal state on $\mathscr{B}(\mathscr{H})$ which has density matrix $\sum_{n=1}^{N} \sigma(P_n)\, P_n$. Define $\rho'_\ell$ similarly.

Let $\mathscr{Z}_\ell$ be the subalgebra of $\mathscr{Z}$ generated by $\ell$. Define $\sigma_\ell = \sigma'_\ell|_{\mathscr{Z}_\ell}$ and $\rho_\ell = \rho'_\ell|_{\mathscr{Z}_\ell}$. Clearly $\sigma_\ell = \sigma|_{\mathscr{Z}_\ell}$ and $\rho_\ell = \rho|_{\mathscr{Z}_\ell}$. By property $e$ of Sect. 6 and monotonicity,

$$\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'_\ell|\rho'_\ell) = \mathrm{ent}_{\mathscr{Z}_\ell}(\sigma_\ell|\rho_\ell) \geqq \mathrm{ent}_{\mathscr{Z}}(\sigma|\rho). \tag{8.11}$$

Let $((\sigma'_\ell, \rho'_\ell))_{\ell \in L}$ be a $w^*$ convergent subnet of $((\sigma'_\ell, \rho'_\ell))_{\ell\in\Lambda}$ converging to $(\sigma', \rho')$. Then $(\sigma'|_{\mathscr{Z}}, \rho'|_{\mathscr{Z}}) = (\sigma, \rho)$, so

$$\mathrm{ent}_{\mathscr{Z}}(\sigma|\rho) \geqq \mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho') \geqq \limsup_{\ell\in L} \mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'_\ell|\rho'_\ell). \tag{8.12}$$

Since $\sigma'_\ell$ and $\rho'_\ell$ are normal, and 8.11 and 8.12 also hold for $\mathrm{ent}^S$ (the same arguments apply), we have

$$\begin{aligned}\mathrm{ent}_{\mathscr{Z}}(\sigma|\rho) &= \limsup_{\ell\in L} \mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'_\ell|\rho'_\ell) = \limsup_{\ell\in L} \mathrm{ent}^S_{\mathscr{B}(\mathscr{H})}(\sigma'_\ell|\rho'_\ell)\\ &= \mathrm{ent}^S_{\mathscr{Z}}(\sigma|\rho).\end{aligned}$$ □

*8C.*

**Theorem 8.13.** *If $(\mathscr{A}, \mathscr{B}(\mathscr{H}))$ is an injective von Neumann algebra then, for all normal states $\sigma$ and $\rho$ on $\mathscr{A}$,*

$$\mathrm{ent}_{\mathscr{A}}(\sigma|\rho) = \mathrm{ent}^S_{\mathscr{A}}(\sigma|\rho).$$

*Proof.* From Sect. 7 $(\mathscr{A}, \mathscr{B}(\mathscr{H}))$ has Property 7.1. Let $(\lambda_\alpha)_{\alpha\in I}$ be the relevant net of maps. Define $\sigma'_\alpha = \sigma\circ\lambda_\alpha$ and $\rho'_\alpha = \rho\circ\lambda_\alpha$. These are normal states on $\mathscr{B}(\mathscr{H})$. Let $((\sigma'_\beta, \rho'_\beta))_{\beta\in J}$ be a $w^*$ convergent subnet converging to $(\sigma', \rho')$. By hypothesis, $\sigma'|_{\mathscr{A}} = \sigma$ and $\rho'|_{\mathscr{A}} = \rho$.

Using Uhlmann's original inequality for $\mathrm{ent}^S$ ([13, Prop. 18]), which is quoted in Sect. 6,

$$\begin{aligned}\mathrm{ent}^S_{\mathscr{B}(\mathscr{H})}(\sigma'_\beta|\rho'_\beta) &\geqq \mathrm{ent}^S_{\mathscr{A}}(\sigma|\rho)\\ &\geqq \mathrm{ent}_{\mathscr{A}}(\sigma|\rho) && \text{(Lemma 6.7)}\\ &\geqq \mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'|\rho') && \text{(Property a)}\\ &\geqq \limsup_{\beta\in J} \mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma'_\beta|\rho'_\beta) && \text{(Property c)}\\ &= \limsup_{\beta\in J} \mathrm{ent}^S_{\mathscr{B}(\mathscr{H})}(\sigma'_\beta|\rho'_\beta) && \text{(Definition 6.1).}\end{aligned}$$

The result follows. □

A property, which is an immediate consequence of this result, is particularly important, since the applications of relative entropy to statistical mechanics rely heavily on it, (cf. [21]):

**i) Araki's Property.** *Let $(\mathscr{A}_\alpha)_{\alpha\in I}$ be a monotone increasing net of injective von Neumann algebras with $\mathscr{A} = (\bigcup_{\alpha\in I} \mathscr{A}_\alpha)''$. Let $\sigma$ and $\rho$ be normal states on $\mathscr{A}$, and let*

$(\sigma_\alpha, \rho_\alpha) = (\sigma|_{\mathscr{A}_\alpha}, \rho|_{\mathscr{A}_\alpha})$. *Then*

$$\mathrm{ent}_{\mathscr{A}}(\sigma|\rho) = \lim_{\alpha \in I} \mathrm{ent}_{\mathscr{A}_\alpha}(\sigma_\alpha|\rho_\alpha).$$

*Proof.* This is a consequence of monotonicity, [10, Theorem 3.9], and the fact that the hypothesis implies that $\mathscr{A}$ is injective (see [16, Prop. 10.25, p. 144]).

## 9. Postscript

Since this paper was first submitted, I have received a preprint [22] from Narnhofer and Thirring in which they have, independently, suggested an extension of $\mathrm{ent}^S_{\mathscr{A}}$ to linear subspaces of $\mathscr{A}$. It seems desirable, for completeness, to comment briefly on their work. Adopting the notational conventions of this paper, their definition runs as follows:

Let $\mathscr{A}$ be a U.H.F. von Neumann algebra, let $\sigma$ and $\rho$ be faithful normal states on $\mathscr{A}$, and let $\mathscr{B}$ be a linear subspace of $\mathscr{A}$ with $1 \in \mathscr{B}$. Let $(\mathscr{H}_\sigma, \pi_\sigma, \xi(\sigma))$ be the G.N.S. representation of $\sigma$ and let $\xi(\rho)$ be the cyclic separating vector representative of $\rho$ in the natural positive cone of $\mathscr{H}_\sigma$. Let $\Delta_{\mathscr{A}}(\rho, \sigma) = \Delta_{\xi(\rho), \xi(\sigma)}$ be the relative modular operator. (See [3, p. 278] and [23, Chap. 2.5] for these definitions.) Let $P_{\mathscr{B}}\colon \mathscr{H}_\sigma \to \pi_\sigma(\mathscr{B})\xi(\sigma)$ be the orthogonal projection. Then Narnhofer and Thirring's relative entropy is defined by

$$\mathrm{ent}^{NT}_{\mathscr{B},\mathscr{A}}(\sigma|\rho) = \lim_{\varepsilon \searrow 0}\left( \xi(\sigma), \log\left( \frac{\varepsilon + P_{\mathscr{B}}\Delta_{\mathscr{A}}(\rho,\sigma)P_{\mathscr{B}}}{1 + \varepsilon P_{\mathscr{B}}\Delta_{\mathscr{A}}(\rho,\sigma)P_{\mathscr{B}}} \right)\xi(\sigma) \right).$$

Narnhofer and Thirring show that this possesses the usual properties of concavity, monotonicity, and semicontinuity. In doing so, they provide new and elegant proofs of these results for $\mathrm{ent}^S_{\mathscr{A}}$. They also show that $\mathrm{ent}^{NT}_{\mathscr{B},\mathscr{A}}(\sigma|\rho) = \mathrm{ent}^S_{\mathscr{B}}(\sigma|\rho) = \mathrm{ent}^S_{\mathscr{B}}(\sigma|_{\mathscr{B}}|\rho|_{\mathscr{B}})$ in the special case that $\mathscr{B}$ is a subalgebra of $\mathscr{A}$. However, the following example, which is the simplest I could construct, shows that for general subspaces $\mathscr{B}$, $\mathrm{ent}^{NT}_{\mathscr{B},\mathscr{A}}(\sigma|\rho)$ is not equal to $\mathrm{ent}_{\mathscr{B}}(\sigma|_{\mathscr{B}}|\rho|_{\mathscr{B}})$. Thus $\mathrm{ent}^{NT}_{\mathscr{B},\mathscr{A}}$ does not satisfy the axioms of Sect. 2. The example also shows that $\mathrm{ent}^{NT}_{\mathscr{B},\mathscr{A}}(\sigma|\rho)$ does not depend solely on $\sigma|_{\mathscr{B}}$ and $\rho|_{\mathscr{B}}$. This is a property which $\mathrm{ent}_{\mathscr{B}}(\sigma|_{\mathscr{B}}|\rho|_{\mathscr{B}})$ does possess, and which I believe is fundamental for a satisfactory definition of the relative entropy on the subspace $\mathscr{B}$.

*Example.* Let $\mathscr{A} = \mathbb{C}^3$ considered as an abelian algebra. Let $\sigma = (1/3, 1/3, 1/3)$ (i.e. $\sigma((a_1, a_2, a_3)) = 1/3\,(a_1 + a_2 + a_3)$) and let $\rho_1 = (1/2, 1/4, 1/4)$, $\rho_2 = (2/5, 2/5, 1/5)$. Let $\mathscr{B}$ be the linear subspace of $\mathscr{A}$ spanned by $1 = (1,1,1)$ and $B = (1,2,4)$. Note that $\rho_1|_{\mathscr{B}} = \rho_2|_{\mathscr{B}}$, since $\rho_1(1) = \rho_2(1) = 1$ and $\rho_1(B) = \rho_2(B) = 2$.

Then, after a long calculation, I claim that, $\mathrm{ent}^{NT}_{\mathscr{B},\mathscr{A}}(\sigma|\rho_1) = -0.03615$, $\mathrm{ent}^{NT}_{\mathscr{B},\mathscr{A}}(\sigma|\rho_2) = -0.04173$, while $\mathrm{ent}_{\mathscr{B}}(\sigma|_{\mathscr{B}}|\rho_1|_{\mathscr{B}}) = \mathrm{ent}_{\mathscr{B}}(\sigma|_{\mathscr{B}}|\rho_2|_{\mathscr{B}}) = -0.02657$.

Roughly speaking, $\mathrm{ent}^{NT}_{\mathscr{B},\mathscr{A}}(\sigma|\rho_1)$ and $\mathrm{ent}^{NT}_{\mathscr{B},\mathscr{A}}(\sigma|\rho_2)$ differ because $P_{\mathscr{B}}\Delta_{\mathscr{A}}(\rho_i,\sigma)P_{\mathscr{B}}$, $i = 1, 2$ depends on $\rho_i(B^2)$ and $\rho_1(B^2) = 11/2$, $\rho_2(B^2) = 26/5$. □

## References

1. Umegaki, H.: Conditional expectation in an operator algebra, IV (entropy and information). Kōdai Math. Sem. Rep. **14**, 59–85 (1962)

2. Araki, H.: On uniqueness of KMS states of one-dimensional quantum lattice systems. Commun. Math. Phys. **44**, 1–7 (1975)
3. Bratteli, O., Robinson, D. W.: Operator algebras and quantum statistical mechanics, Vol. II. Berlin, Heidelberg, New York: Springer 1981
4. Fröhlich, J., Pfister, C.: On the absence of spontaneous symmetry breaking and crystalline ordering in two dimensional systems. Commun. Math. Phys. **81**, 277–298 (1981)
5. Ohya, M.: On compound state and mutual information in quantum information theory. IEEE Trans. Inf. Theory **29**, 770–774 (1983)
6. Lindblad, G.: Expectations and entropy inequalities for finite quantum systems. Commun. Math. Phys. **39**, 111–119 (1974)
7. Hiai, F., Ohya, M. Tsukada, M.: Sufficiency, KMS condition and relative entropy in von Neumann algebras. Pac. J. Math. **96**, 99–109 (1981)
8. Hiai, F., Ohya, M., Tsukada, M.: Sufficiency and relative entropy in *-algebras with applications in quantum systems. Pac. J. Math. **107**, 117–140 (1983)
9. Araki, H.: Relative entropy of states of von Neumann algebras. Publ. Res. Inst. Math. Sci. **11**, 809–833 (1976)
10. Araki, H.: Relative entropy for states of von Neumann algebras II. Publ. Res. Inst. Math. Sci. **13**, 173–192 (1977)
11. Pusz, W., Woronowicz, S. L.: Form convex functions and the WYDL and other inequalities. Lett. Math. Phys. **2**, 505–512 (1978)
12. Pusz, W., Woronowicz, S. L.: Functional calculus for sesquilinear forms and the purification map. Rep. Math. Phys. **8**, 159–170 (1975)
13. Uhlmann, A.: Relative entropy and the Wigner–Yanase–Dyson–Lieb concavity in an interpolation theory. Commun. Math. Phys. **54**, 21–32 (1977)
14. Benoist, R. W., Marchand, J.-P., Wyss, W.: A note on relative entropy. Lett. Math. Phys. **3**, 169–173 (1979)
15. Gudder, S., Marchand, J.-P., Wyss, W.: Bures distance and relative entropy. J. Math. Phys. **20**, 1963–1966 (1979)
16. Strătilă, S.: Modular theory in operator algebras. Tunbridge Wells: Abacus 1981
17. Effros, E. G., Lance, E. C.: Tensor products of operator algebras. Adv. Math. **25**, 1–34 (1977)
18. Connes, A.: On the equivalence between injectivity and semi-discreteness for operator algebras. Colloq. Int. C.N.R.S. **274**, 107–112 (1977)
19. Dixmier, J.: Les algèbres d'opérateurs dans l'espace hilbertien, (algèbres de von Neumann). Paris: Gauthier–Villars. $2^{nd}$ ed. (1969)
20. Davies, E. B.: Quantum theory of open systems. London, New York, San Francisco: Academic Press 1976
21. Araki, H.: On the equivalence of the KMS condition and the variational principle for quantum lattice systems. Commun. Math. Phys. **38**, 1–10 (1974)
22. Narnhofer, H., Thirring, W.: From relative entropy to entropy. Fizika **17**, 257–265 (1985)
23. Bratteli, O., Robinson, D.W.: Operator algebras and quantum statistical mechanics, Vol. I. Berlin: Heidelberg, New York, Springer 1979

Communicated by H. Araki

Received June 17, 1985; in revised form October 18, 1985