Invent. math. 132, 433–455 (1998)

Inventiones mathematicae

Maximality of entropy in finite von Neumann algebras

Uffe Haagerup[1], **Erling Størmer**[2]

[1] Department of Mathematics and Computer Science, Odense University, Campusvej 55, DK-5230 Odense M, Denmark
[2] Department of Mathematics, University of Oslo, P.O. Box 1053 Blindern, N- 0316 Oslo, Norway

Oblatum 24-IV-1997 & 6-V-1997

Abstract. It is shown that the entropy function $H(N_1,\dots,N_k)$ on finite dimensional von Neumann subalgebras of a finite von Neumann algebra attains its maximal possible value $H(\bigvee\limits_{\ell=1}^{k} N_\ell)$ if and only if there exists a maximal abelian subalgebra A of $\bigvee\limits_{\ell=1}^{k} N_\ell$ such that $A=\bigvee\limits_{\ell=1}^{k}(A\cap N_\ell)$.

1 Introduction

Nonabelian entropy of automorphisms of finite von Neumann algebras as put forward in [2] is not yet well understood. If R is a von Neumann algebra with a faithful normal finite trace τ the definition of entropy is based on a function $H(N_1,\dots,N_k)$ on finite dimensional von Neumann subalgebras $N_1,\dots,N_k$ of R, just like the entropy in the classical case is based on the entropy $H(\bigvee_{\ell=1}^k \mathscr{P}_\ell)$ of finite measurable partitions. The function $H(N_1,\dots,N_k)$ satisfies many of the same properties as $H(\bigvee_{\ell=1}^k P_\ell)$; in particular it is increasing, and $H(N_1,\dots,N_k)\le H(\bigvee_{\ell=1}^k N_\ell)$, where $\bigvee_{\ell=1}^k N_\ell$ denotes the von Neumann algebra generated by $N_1,\dots,N_k$. However, some crucial properties are false; for example it is not in general additive on tensor products, and if $H(N_1,\dots,N_k)=\sum_{\ell=1}^k H(N_\ell)$ we cannot conclude independence of the N_ℓ's in any natural sense. Only when there is "enough" commutativity between the N_ℓ's can we expect nice behaviour of the function H. More specifically, if there exists a masa – maximal abelian von Neumann subalgebra – A of $\bigvee_{\ell=1}^k N_\ell$ such that $A=\bigvee_{\ell=1}^k(A\cap N_\ell)$, then we have the nice formula

$$H(N_1,\dots,N_k)=H(A\cap N_1,\dots,A\cap N_k)=H(A)=H\left(\bigvee_{\ell=1}^{k} N_\ell\right)\ ,$$

thus yielding a very useful criterion for computing entropy.

In the present paper we prove the converse to the last result i.e. if we have maximality, viz $H(N_1,\ldots,N_k)=H(\bigvee_{\ell=1}^k N_\ell)$, then $\bigvee_{\ell=1}^k N_\ell$ is finite dimensional, and A as above exists. A consequence is that

$$H(M_1\otimes N_1,\ldots,M_k\otimes N_k)=H(M_1,\ldots,M_k)+H(N_1,\ldots,N_k)$$

when both the M_ℓ's and the N_ℓ's satisfy the maximality condition. If we furthermore have that $H(N_1,\ldots,N_k)=\sum_{\ell=1}^k H(N_\ell)$ then the masa A has the further properties that the algebras $A\cap N_\ell$'s are masas in the N_ℓ's and are independent.

Our result indicates two things; firstly that it is in some cases possible to describe relative positions of algebras $N_1,\ldots,N_k$ from values of the function $H(N_1,\ldots,N_k)$. Secondly, if α is a τ-invariant automorphism of R and $N_\ell=\alpha^\ell(N)$ for a fixed algebra N with $R=\bigvee_{-\infty}^{\infty}N_\ell$, then if the maximality condition prevails for $H(N_1,\ldots,N_k)$ for all k, then it should be within reach to prove analogues of some of the classical theorems for generators like the Shannon, Breiman, McMillan theorem.

The paper is organized as follows. In section 2 we prove some analytic results needed in the sequel. In section 3 show an inequality which is crucial for the proof. It can be described as follows. In [2, eq. (8)] it was shown that if $x_{ij}\in R^+$ with $\sum x_{ij}=1$ then

$$\sum_{ij}\tau\eta(x_{ij})\le\sum_i\tau\eta\left(\sum_j x_{ij}\right)+\sum_j\tau\eta\left(\sum_i x_{ij}\right)\ ,$$

where η is the function $\eta(t)=-t\log t$ for $t>0$, and $\eta(0)=0$. We improve this inequality and give explicit estimates for the difference between the right side and the left side. In section 4 we review the basic theory of the entropy function H and prove some general results needed in the proof of the main result in section 5.

The authors are indepted to E. Alfsen and M.B. Ruskai for valuable comments.

2 Analytic preliminaries

In this section we collect some facts which will be needed in the subsequent sections. It was shown by Nakamura and Umegaki [7] that the function $-\eta(t)$ is operator convex. Choi proved in [1, Lem. 3.6] that every operator convex function h on an open interval is strictly operator convex, except for the trivial case when h is a polynomial of degree at most 1. We shall need the following slight extension of Choi's result, giving at the same time a new proof of the operator concavity of $\eta(t)$.

Lemma 2.1 *The function* $h(t)=t\log t,\ t>0,\ h(0)=0,$ *is strictly operator convex on* $B(H)^+$, *i.e. for* $x,y\in B(H)^+$,

$$h(\tfrac{1}{2}(x+y)) \leq \tfrac{1}{2}h(x) + \tfrac{1}{2}h(y) \ ,$$

with equality if and only if $x = y$.

Proof. Note that if x, y are invertible we could have applied Choi's result directly. For the general case we use the following simple lemma of Choi [1, Lem. 3.5]. If $z, w \in B(H)_+$ are invertible, then

$$(\tfrac{1}{2}(z+w))^{-1} \leq \tfrac{1}{2}(z^{-1} + w^{-1}) \tag{1}$$

and equality holds if and only if $z = w$. From the integral representation

$$\log t = \int_0^\infty \Big(\frac{1}{1+a} - \frac{1}{t+a}\Big) da$$

we have for every $t \geq 0$,

$$\begin{aligned} h(t) &= \int_0^\infty \Big(\frac{t}{1+a} - \frac{t}{t+a}\Big) da \\ &= \int_0^\infty \Big(\frac{t}{1+a} - 1 + \frac{a}{t+a}\Big) da \end{aligned}$$

Hence for all $x, y \in B(H)^+$,

$$\begin{aligned} &\frac{1}{2}h(x) + \frac{1}{2}h(y) - h\Big(\frac{x+y}{2}\Big) \\ &= \int_0^\infty \Big(\frac{1}{2}(x+a1)^{-1} + \frac{1}{2}(y+a1)^{-1} - \Big(\frac{x+y}{2} + a1\Big)^{-1}\Big) a da \end{aligned}$$

as a $B(H)$-valued integral. By (1)

$$\frac{1}{2}(x+a1)^{-1} + \frac{1}{2}(y+a1)^{-1} - \Big(\frac{x+y}{2} + a1\Big)^{-1} \geq 0$$

for $a > 0$, and equality holds (for each fixed $a > 0$) if and only if $x = y$. This proves the lemma. □

Lemma 2.2 *Let R be a von Neumann algebra with a faithful normal tracial state τ. Let $a, k \in R^+$, $b \in R$, and assume $a + tk \geq 0$ and invertible for $t \in [0,1]$. Then the function $t \to \tau(b \log(a + tk))$ is differentiable with derivative*

$$\frac{d}{dt}\tau(b\log(a+tk)) = \int_0^\infty \tau(b(\alpha 1+a+tk)^{-1}k(\alpha 1+a+tk)^{-1})d\alpha \ .$$

Proof. By [6, eq. 3.6]

$$\frac{d}{dt}\tau(b\log(a+tk))\big|_{t=0} = \int_0^\infty \tau(b(\alpha 1+a)^{-1}k(\alpha 1+a)^{-1})d\alpha \ .$$

Therefore

$$\begin{aligned}\frac{d}{dt}\tau(b\log(a+tk))\big|_{t=t_0} &= \frac{d}{ds}\tau(b\log((a+t_0k)+sk))\big|_{s=0} \\ &= \int_0^\infty \tau(b(\alpha 1+a+t_0k)^{-1}k(\alpha 1+a+t_0k)^{-1})d\alpha \ .\end{aligned}$$

□

Lemma 2.3 *With R and τ as in Lemma 2.2 suppose $k=k^*\in R$, $y\in R^+$, and $y+tk\geq 0$ and invertible for all $t\in[0,1]$. Then the function $t\to\tau\eta(y+tk)$ is differentiable with derivative*

$$\frac{d}{dt}\tau\eta(y+tk) = -\tau(k(\log(y+tk)+1)) \ ,$$

where $\eta(s)=-s\log s$, $s>0$, $\eta(0)=0$.

Proof. Let $x(t)=y+tk$. By Lemma 2.2

$$\begin{aligned}\frac{d}{dt}\tau\eta(x(t)) &= -\tau\left(\left(\frac{d}{dt}x(t)\right)\log x(t) + x(t)\frac{d}{dt}\log x(t)\right) \\ &= -\tau(k\log x(t)) + \int_0^\infty \tau(x(t)(\alpha 1+x(t))^{-1}k(\alpha 1+x(t))^{-1})d\alpha \\ &= -\tau(k\log x(t)) + \int_0^\infty \tau((\alpha 1+x(t))^{-1}x(t)(\alpha 1+x(t))^{-1}k)\ d\alpha\end{aligned}$$

Now

$$\int_0^\infty \frac{s}{(\alpha 1+s)^2}d\alpha = \frac{-s}{\alpha+s}\Big|_{\alpha=0}^{\infty} = 1 \qquad \text{for } s>0 \ .$$

Thus by spectral calculus

$$\int_0^\infty (\alpha 1 + x(t))^{-1} x(t) (\alpha 1 + x(t))^{-1} d\alpha = 1 \ ,$$

so that

$$\frac{d}{dt}\tau\eta(x(t)) = -\tau(k(\log(x(t)) + 1)) \ .$$

Lemma 2.4 *Let* R *and* τ *be as in Lemma* 2.2 *and denote by* R^+_{inv} *the set of invertible elements in* R^+. *Let* $x_1, x_2 \in R^+_{\mathrm{inv}}$. *Then the function* $f\colon R^+_{\mathrm{inv}} \to \mathbb{R}$ *defined by*

$$f(y) = \tau(y(\log y - \log x_1 - \log x_2 - 1))$$

is strictly convex and has minimum at

$$y_0 = \exp(\log x_1 + \log x_2) \ .$$

In particular, $f(y) > f(y_0)$ *for all* $y \in R^+_{\mathrm{inv}}$, $y \neq y_0$.

Proof. Let $y \in R^+_{\mathrm{inv}}$ with $y \neq y_0$, and set

$$y(t) = (1-t)y_0 + ty = y_0 + t(y - y_0), \qquad t \in [0,1] \ .$$

Then $y'(t) = y - y_0$, and there exists $\varepsilon > 0$ such that $y(t) \geq \varepsilon 1$ for all $t \in [0,1]$. By Lemmas 2.2 and 2.3 $f(y(t))$ is a twice differentiable function of t, and we have

$$\begin{aligned} \frac{d}{dt} f(y(t)) &= \frac{d}{dt}\tau(-\eta(y(t))) - \tau(y'(t)(\log x_1 + \log x_2 + 1)) \\ &= \tau((y-y_0)[(\log y(t) + 1) - (\log x_1 + \log x_2 + 1)]) \\ &= \tau((y-y_0)(\log y(t) - \log x_1 - \log x_2)) \ . \end{aligned} \tag{2}$$

By Lemma 2.1 and faithfulness of τ the function $x \to f(x)$, $x \in R^+_{\mathrm{inv}}$, is strictly convex, hence the function $g(t) = f(y(t))$ is a strictly convex function of t since $y \neq y_0$. By (2)

$$g'(t) = \tau((y-y_0)(\log y(t) - \log x_1 - \log x_2)) \ ,$$

and so

$$g'(0) = \tau((y-y_0)(\log y_0 - \log x_1 - \log x_2)) = 0 \ .$$

Since g is strictly convex, 0 is the unique minimum of g, hence $g(t) > g(0)$ for $t > 0$. In particular, $f(y_0) = g(0) < g(1) = f(y)$. □

The previous lemma will be used in section 3, the next in section 4.

Lemma 2.5 *Let $g_n\colon [0,1] \to S \subset \mathbb{R}^k$ be a Borel function for $n \in \mathbb{N}$, where S is a compact set. Then there exist measure preserving Borel isomorphisms $\alpha_n\colon [0,1] \to [0,1]$ such that $(g_n \circ \alpha_n)_{n\in\mathbb{N}}$ has a subsequence which converges pointwise a.e.*

Proof. There is no loss of generality to assume $S = [0,1]^k$. We consider first the case $k = 1$. Choose $g'_n\colon [0,1] \to S$ such that g'_n takes only finite number of values, and $\|g_n - g'_n\|_\infty < {}^1/_n$. By suitable choice of α_n we can assume

$$h_n = g'_n \circ \alpha_n\colon [0,1] \to [0,1]$$

is an increasing function for each $n \in \mathbb{N}$. Put

$$V = \{h\colon [0,1] \to [0,1]\colon\ h \text{ is increasing}\}\ .$$

Then V is compact in the topology of pointwise convergence. Furthermore, since each increasing function has at most a countable number of points of discontinuity, each function in V is Borel. Choose by compactness a subnet (h_{n_α}) of (h_n) which converges pointwise to a function $h \in V$. Put

$$T = (\mathbb{Q} \cap [0,1]) \cup \{x \in [0,1]\colon\ h \text{ is discontinuous at } x\}\ .$$

Then T is countable and dense in $[0,1]$. Choose a subsequence (h_{n_i}) of (h_{n_α}) which converges pointwise to h on T. We assert that $h_{n_i} \to h$ pointwise on all of $[0,1]$. Indeed, if $x \in [0,1] \setminus T$ and $\varepsilon > 0$ choose $z, z' \in [0,1] \setminus T$ such that $z' < x < z$ and $h(z) - h(x) < {}^\varepsilon/_2$, and $h(x) - h(z') < {}^\varepsilon/_2$. Choose $y, y' \in T$ with $z' < y' < x < y < z$, and choose i_0 such that $i \ge i_0$ implies $|h_{n_i}(y) - h(y)| < {}^\varepsilon/_2$ and $|h_{n_i}(y') - h(y')| < {}^\varepsilon/_2$. Since h_{n_i} is increasing we have when $h_{n_i}(x) - h(x) \ge 0$ that

$$\begin{aligned} |h_{n_i}(x) - h(x)| &\le |h_{n_i}(y) - h(y)| + |h(y) - h(x)| \\ &\le {}^\varepsilon/_2 + |h(z) - h(x)| \\ &< {}^\varepsilon/_2 + {}^\varepsilon/_2 = \varepsilon\ . \end{aligned}$$

If $h_{n_i}(x) - h(x) \le 0$ we argue similarly with y' and z'. Since $\|h_n - g_n \circ \alpha_n\|_\infty = \|(g'_n - g_n) \circ \alpha_n\|_\infty \to 0$ it follows that

$$g_{n_i} \circ \alpha_{n_i} = h_{n_i} \to h \qquad \text{pointwise a.e.}$$

This proves the Lemma for $k = 1$.

We now assume $k \ge 2$ and $S = [0,1]^k$. Choose $\phi\colon [0,1] \to [0,1]^k$ which is continuous and surjective, a k-dimensional Peano path. We apply [4, Thm. 14.3.6] to $X = [0,1]^k$, $Y = [0,1]$, $\mathscr{S} = \{(\phi(p), p) : p \in [0,1]\}$, and π the projection $X \times Y \to X$. Then $\pi(\mathscr{S}) = [0,1]^k$. By the theorem there exists a Borel map $\psi : [0,1]^k \to [0,1]$ such that $(q, \psi(q)) \in \mathscr{S}$ for $q \in [0,1]^k$. Hence if $q = \phi(p)$ with $p \in [0,1]$ then $(\phi(p), p) = (q, \psi(q))$. In particular $\psi \circ \phi = \text{id.}$, so ψ is in particular surjective. But then with $p = \psi(q)$, $\phi \circ \psi(q) = q$, so $\phi \circ \psi = \text{id}$.

Let now (g_n) be the sequence in the lemma. Then $\widetilde{g}_n = \psi \circ g_n : [0,1] \to [0,1]$. From the case $k = 1$ there is $\alpha_n : [0,1] \to [0,1]$ which is a Borel isomorphism such that $\widetilde{g}_n \circ \alpha_n$ has a converging subsequence. But $g_n = \phi \circ \widetilde{g}_n$, and ϕ is continuous, so $(g_n \circ \alpha_n)$ has a subsequence which converges pointwise a.e. $\square$

In the course of the proof we showed the following

Corollary 2.6 *There exists a Borel measurable map $\psi : [0,1]^k \to [0,1]$ such that whenever $(g_n)_{n \in \mathbb{N}}$ is a sequence of Borel functions, $g_n : [0,1] \to [0,1]^k$, with $\psi \circ g_n$ an increasing function on $[0,1]$, then $(g_n)_{n \in \mathbb{N}}$ has a subsequence which converges pointwise a.e.*

3 An inequality

Let R be a finite von Neumann algebra with a faithful normal trace τ such that $\tau(1) = 1$. We follow [2] and use the notation

$$S_k = \Big\{ (x_{i_1 \ldots i_k}) : x_{i_1 \ldots i_k} \in R^+ \text{and equal to 0 except for a finite number of indices}, \sum_{i_1 \ldots i_k} x_{i_1 \ldots i_k} = 1 \Big\} .$$

$$x^\ell_{i_\ell} = \sum_{i_1 \ldots i_{l-1} i_{l+1} \ldots i_k} x_{i_1 \ldots i_k} .$$

As a consequence of Lieb's result [6] that relative entropy is a jointly convex function it was shown in [2, eq. (8)] that we have the following inequality. If $(x_{ij}) \in S_2$ then

$$\sum_{ij} \tau\eta(x_{ij}) \le \sum_i \tau\eta(x^1_i) + \sum_j \tau\eta(x^2_j) .$$

We shall in the present section improve this inequality via a proof which does not make use of relative entropy.

Lemma 3.1 *Let a and b be self-adjoint operators in R. Then we have*

$$\tau(e^a e^b) - \tau(e^{a+b}) \ge \frac{1}{2} \| [e^{a/2}, e^{b/2}] \|_2^2 .$$

Proof. It is well-known that

$$e^{a+b} = \lim_{p\to\infty} (e^{a/p}e^{b/p})^p, \qquad p \in \mathbb{N} \ ,$$

where the limit is in norm (cf. [9], proof of Theorem VIII.29). Straightforward computation yields

$$\tau(e^a e^b) - \tau((e^{a/2}e^{b/2})^2) = \frac{1}{2}\|[e^{a/2}, e^{b/2}]\|_2^2 \ . \tag{3}$$

Now

$$\begin{aligned}\tau((e^{a/2}e^{b/2})^2) &= \tau((e^{a/4}e^{b/4})(e^{a/4}e^{b/4})^*(e^{a/4}b^{b/4})(e^{a/4}e^{b/4})^*) \\ &= \|e^{a/4}e^{b/4}\|_4^4 \ ,\end{aligned}$$

where $\|x_p\| = \tau(|x|^p)^{1/p}$. By the generalized Hölder inequality [5, Corollary 3.2] we have for $p \in \mathbb{N}$

$$\|x_1 \dots x_p\|_1 \le \|x_1\|_p \dots \|x_p\|_p, \qquad x_1, \dots, x_p \in R \ .$$

Hence with $p = 4$, $x_1 = x_2 = x_3 = x_4 = e^{a/4}e^{b/4}$ we have

$$\tau((e^{a/4}e^{b/4})^4) \le \|e^{a/4}e^{b/4}\|_4^4 = \tau(e^{a/2}e^{b/2})^2 \ ) \ .$$

Inductively we obtain

$$\begin{aligned}\tau((e^{a/2^k}e^{b/2^k})^{2^k}) &\le \left\|e^{a/2^k}e^{b/2^k}\right\|_{2^k}^{2^k} \\ &= \tau((e^{a/2^{k-1}}e^{b/2^{k-1}})^{2^{k-1}}) \\ &\le \tau((e^{a/2}e^{b/2})^2) \ .\end{aligned}$$

Since the left side converges to $\tau(e^{a+b})$ we find

$$\tau(e^{a+b}) \le \tau((e^{a/2}e^{b/2})^2)$$

which combined with (3) completes the proof. □

Theorem 3.2 *Let* $(x_{ij}) \in S_2$, $i = 1, \dots, m$, $j = 1, \dots, n$. *Then we have*

$$\sum_{i=1}^m \tau\eta(x_i^1) + \sum_{j=1}^n \tau\eta(x_j^2) - \sum_{i=1}^m\sum_{j=1}^n \tau\eta(x_{ij}) \ge \frac{1}{2}\sum_{i=1}^m\sum_{j=1}^n \left\|[(x_i^1)^{1/2}, (x_j^2)^{1/2}]\right\|_2^2$$

Proof. Assume first that all x_{ij}, $i = 1, \dots, m$, $j = 1, \dots, n$, are invertible elements in R^+. Then we have

$$\begin{aligned}
&\sum_i \tau\eta(x_i^1) + \sum_j \tau\eta(x_j^2) - \sum_{ij} \tau\eta(x_{ij}) \\
&\quad = \sum_{ij} \tau(x_{ij}(\log x_{ij} - \log x_i^1 - \log x_j^2)) \\
&\quad = 1 + \sum_{ij} \tau(x_{ij}(\log x_{ij} - \log x_i^1 - \log x_j^2 - 1)) \ .
\end{aligned}$$

Set $y_{ij} = \exp(\log x_i^1 + \log x_j^2)$. Then by Lemma 2.4 the above quantity is greater than

$$1 + \sum_{ij} \tau(y_{ij}(\log y_{ij} - \log x_i^1 - \log x_j^2 - 1)) = 1 - \sum_{ij} \tau(y_{ij})$$

Put $a_i = \log x_i^1$, $b_j = \log x_j^2$. Then $x_i^1 = e^{a_i}$, $x_j^2 = e^{b_j}$, $y_{ij} = e^{a_i+b_j}$. Since $\sum\limits_i x_i^1 = \sum\limits_j x_j^2 = 1$ we get by Lemma 3.1

$$\begin{aligned}
1 - \sum_{ij} \tau(y_{ij}) &= \sum_{ij} \tau(e^{a_i} e^{b_j}) - \sum_{ij} \tau(e^{a_i+b_j}) \\
&\geq \frac{1}{2} \sum_{ij} \| [e^{a_i/2}, e^{b_j/2}] \|_2^2 \\
&= \frac{1}{2} \sum_{ij} \left\| [(x_i^1)^{1/2}, (x_j^2)^{1/2}] \right\|_2^2 \ ,
\end{aligned}$$

which completes the proof when all x_{ij} are invertible. The general case follows by approximation. Indeed, set

$$\widetilde{x}_{ij} = (1-\varepsilon)x_{ij} + \frac{\varepsilon}{mn} 1, \qquad 0 < \varepsilon < 1 \ .$$

Then $\widetilde{x}_{ij}$ is invertible, $\sum\limits_{ij} \widetilde{x}_{ij} = 1$, and

$$\begin{aligned}
\widetilde{x}_i^1 &= (1-\varepsilon)x_i^1 + \frac{\varepsilon}{m} 1, \\
\widetilde{x}_j^2 &= (1-\varepsilon)x_j^2 + \frac{\varepsilon}{n} 1 \ .
\end{aligned}$$

By continuity of the function η on [0,1] we are done. □

Corollary 3.3 *Let* $(x_{i_1 \cdots i_k}) \in S_k$. *Then*

$$\sum_{\ell=1}^{k} \sum_{i_\ell} \tau\eta(x_{i_\ell}^{\ell}) - \sum_{i_1 \ldots i_k} \tau\eta(x_{i_1 \ldots i_k}) \geq \max_{\ell \neq m} \tfrac{1}{2} \sum_{i_\ell i_m} \| [(x_{i_\ell}^{\ell})^{1/2}, (x_{i_m}^{m})^{1/2}] \|_2^2 \ .$$

Proof. By Theorem 3.2 the corollary holds for $k = 2$. Let $k > 2$ and assume the corollary holds for $k - 1$. Fix $l \neq m$ in $\{1, \dots k\}$. For simplicity of notation we may assume $l, m < k$. Put

$$y_{i_1 \dots i_{k-1}} = \sum_{i_k} x_{i_1 \dots i_k} \quad .$$

Then $(y_{i_1 \dots i_{k-1}}) \in S_{k-1}$ and $y^j_{i_j} = x^j_{i_j}$. Let

$$\alpha = \tfrac{1}{2} \sum_{i_\ell i_m} \| [(x^\ell_{i_\ell})^{1/2}, (x^m_{i_m})^{1/2}] \|_2^2 = \tfrac{1}{2} \sum_{i_\ell i_m} \| [(y^\ell_{i_\ell})^{1/2}, (y^m_{i_m})^{1/2}] \|_2^2$$

Then by induction hypothesis

$$\sum_{\ell=1}^{k-1} \sum_{i_\ell} \tau\eta(y^\ell_{i_\ell}) - \sum_{i_1 \dots i_{k-1}} \tau\eta(y_{i_1 \dots i_{k-1}}) \geq \alpha \quad ,$$

so that

$$\alpha \leq \sum_{\ell=1}^{k} \sum_{i_\ell} \tau\eta(x^\ell_{i_\ell}) - \sum_{i_k} \tau\eta(x^k_{i_k}) - \sum_{i_1 \dots i_{k-1}} \tau\eta\left(\sum_{i_k} x_{i_1 \dots i_k} \right) \quad . \tag{4}$$

Let $\Phi : \mathbb{N}^{k-1} \to \mathbb{N}$ be a bijection, and put $j = \Phi(i_1 \dots i_{k-1})$. Then by Theorem 3.2 (or rather [2, eq. (8)]) we get with $x_{ji_k} = x_{i_1 \dots i_k}$,

$$\begin{aligned} \sum_{i_1 \dots i_k} \tau\eta(x_{i_1 \dots i_k}) &= \sum_{ji_k} \tau\eta(x_{ji_k}) \\ &\leq \sum_j \tau\eta(x^1_j) + \sum_{i_k} \tau\eta(x^k_{i_k}) \\ &= \sum_{i_1 \dots i_{k-1}} \tau\eta\left(\sum_{i_k} x_{i_1 \dots i_k} \right) + \sum_{i_k} \tau\eta(x^k_{i_k}) \quad . \end{aligned}$$

Thus by (4) the corollary follows. □

4 Entropy

Throughout this section R is a finite von Neumann algebra with a faithful normal trace τ with $\tau(1) = 1$. If N is a von Neumann subalgebra of R we denote by E_N the unique τ-invariant faithful normal conditional expectation of R onto N. Let notation be as in section 3. If $N_1, \dots, N_k$ are finite dimensional von Neumann subalgebras of R we follow [2] and define

$$H(N_1,\dots,N_k)=\sup_{(x_{i_1\dots i_k})\in S_k}\left\{\sum_{i_1\dots i_k}\eta\tau(x_{i_1\dots i_k})-\sum_{\ell=1}^{k}\sum_{i_\ell}\tau\eta E_{N_\ell}(x^{\ell}_{i_\ell})\right\}\ .$$

If we want to express which trace we use we write $H_\tau(N_1,\dots,N_k)$. Then $H(N_1,\dots,N_k)\geq 0$, symmetric in its arguments and satisfies the following properties.

(A) $H(N_1,\dots,N_k)\leq H(P_1,\dots,P_k)$ if $N_\ell\subset P_\ell$
(B) $H(N_1,\dots,N_k,N_{k+1},\dots,N_p)\leq H(N_1,\dots,N_k)+H(N_{k+1},\dots,N_p)$
(C) $P_1,\dots,P_k\subset P\Rightarrow H(P_1,\dots,P_k,P_{k+1},\dots,P_p)\leq H(P,P_{k+1},\dots,P_p)$
(D) For any family of minimal projections of N, $(e_\ell)_{\ell\in I}$ such that $\sum_{\ell\in I}e_\ell=1$ we have $H(N)=\sum_{\ell\in I}\eta\tau(e_\ell)$.
(E) If $\bigvee_{\ell=1}^{k}N_\ell$ is generated by pairwise commuting von Neumann subalgebras P_ℓ of N_ℓ then $H(N_1,\dots,N_k)=H(\bigvee_{\ell=1}^{k}N_\ell)$.

Two inequalities were useful in [2]. The first is [2, eq. (12)] which states that if $x,y\in R^+$, then

$$\eta\tau(x+y)-\tau\eta(x+y)\leq(\eta\tau(x)-\tau\eta(x))+(\eta\tau(y)-\tau\eta(y))\ . \tag{5}$$

Since E_N is completely positive, it follows from Jensen's inequality for operator convex functions, cf. [3] combined with Stinespring's theorem, that

$$\eta E_N(x)\geq E_N\eta(x),\qquad x\in R^+\ . \tag{6}$$

In particular it follows that $\tau\eta E_N(x)\geq\tau\eta(x)$. Note that if $x\in R^+$ then we have

$$x\in N\quad\text{if and only if}\quad \eta E_N(x)=E_N\eta(x)\ . \tag{7}$$

Indeed, by operator concavity of η if $\eta E_N(x)=E_N\eta(x)$ then

$$\begin{aligned}E_N\eta(x)=&E_N\Big(\frac{1}{2}(\eta E_N(x)+\eta(x))\Big)\leq E_N\eta\Big(\frac{1}{2}(E_N(x)+x)\Big)\\ \leq&\eta\Big(E_N\Big(\frac{1}{2}(E_N(x)+x)\Big)\Big)=\eta E_N(x)=E_N\eta(x)\ .\end{aligned}$$

Thus, by strict operator concavity of η, Lemma 2.1, and faithfulness of E_N $E_N(x)=x$. The converse is trivial.

Lemma 4.1 *Let $M\subset N$ be finite dimensional von Neumann subalgebras of R. Then $H(M)=H(N)$ if and only if each masa in M is a masa in N.*

Proof. If $H(M)=H(N)$ and A is a masa in M, let $(e_\ell)_{\ell\in I}$ be the minimal projections in A. Suppose $e_\ell=f_\ell+g_\ell$ with f_ℓ and g_ℓ projections in N. Since $\eta(e)=0$ for each projection e it follows from (5) that

$$\begin{aligned}\eta\tau(e_\ell) &= \eta\tau(f_\ell+g_\ell)-\tau\eta(f_\ell+g_\ell)\\ &\le (\eta\tau(f_\ell)-\tau\eta(f_\ell))+(\eta\tau(g_\ell)-\tau\eta(g_\ell))\\ &= \eta\tau(f_\ell)+\eta\tau(g_\ell)\ .\end{aligned}$$

By hypothesis and (D), $\eta\tau(e_\ell)=\eta\tau(f_\ell)+\eta\tau(g_\ell)$. Now if $\eta(a+b)=\eta(a)+\eta(b)$ with $a,b\in[0,1]$ then a or b equals zero. Indeed, if $a\neq 0$ and

$$f(b)=\eta(a)+\eta(b)-\eta(a+b)$$

then $f(0)=0$, and $f'(b)=\log(a+b)-\log b>0$, proving the assertion. Thus $\tau(f_\ell)$ or $\tau(g_\ell)=0$, hence f_ℓ or $g_\ell=0$, and e_ℓ is minimal in N. Therefore A is a masa in N.

The converse is immediate from (D). □

Lemma 4.2 *Let $N_1,\ldots,N_k$ be finite dimensional von Neumann subalgebras of R and $N=\bigvee_{\ell=1}^{k} N_\ell$. Suppose $e\neq 0,1$ is a central projection in N such that $e\in N_\ell$ for each ℓ. Put $f=1-e$. Then*

$$\begin{aligned}H(N_1,\ldots,N_k) &= \tau(e)H_{\tau_e}(N_1e,\ldots,N_ke)+\tau(f)H_{\tau_f}(N_1f,\ldots,N_kf)\\ &\quad+\eta\tau(e)+\eta\tau(f)\ .\end{aligned}$$

Here $\tau_e(xe)=\tau(e)^{-1}\tau(xe)$ is normalized trace on Ne.

Proof. Since e belongs to the center of N_ℓ the conditional expectation of R_e ($=eRe$ acting on eH) is given by $E_{N_\ell e}(exe)=eE_{N_\ell}(x)$. Let $(x_{i_1\ldots i_k})\in S_k$. By the definition of $H(N_1,\ldots,N_k)$ we may assume $x_{i_1\ldots i_k}\in N$. Thus

$$\begin{aligned}&\sum_{i_1\ldots i_k}\eta\tau(ex_{i_1\ldots i_k})-\sum_\ell\sum_{i_\ell}\tau\eta\big(E_{N_\ell}ex^\ell_{i_\ell}\big)\\ &=\sum_{i_1\ldots i_k}\eta(\tau(e)\tau_e(ex_{i_1\ldots i_k}))-\sum_\ell\sum_{i_\ell}\tau(e)\tau_e\eta\big(E_{N_\ell e}ex^\ell_i\big)\\ &=\eta(\tau(e))+\tau(e)\sum_{i_1\ldots i_k}\eta\tau_e(ex_{i_1\ldots i_k})-\tau(e)\sum_\ell\sum_{i_\ell}\tau_e\eta\big(E_{N_\ell e}ex^\ell_i\big)\\ &\le\eta\tau(e)+\tau(e)H_{\tau_e}(N_1e,\ldots,N_ke)\end{aligned}$$

by definition of $H_{\tau_e}(N_1e,\ldots,N_ke)$, and similarly for f. By the inequality (5) we therefore have

$$\begin{aligned}
&\sum_{i_1\ldots i_k}\eta\tau(x_{i_1\ldots i_k})-\sum_\ell\sum_{i_\ell}\tau\eta\big(E_{N_\ell}x^\ell_{i_\ell}\big)\\
&=\sum_{i_1\ldots i_k}(\eta\tau(ex_{i_1\ldots i_k}+fx_{i_1\ldots i_k})-\tau\eta(ex_{i_1\ldots i_k}+fx_{i_1\ldots i_k})\\
&\quad+\sum_{i_1\ldots i_k}\tau\eta(x_{i_1\ldots i_k})-\sum_\ell\sum_{i_\ell}\tau\eta\big(E_{N_\ell}x^\ell_{i_\ell}\big)\\
&\quad\leq\sum_{i_1\ldots i_k}((\eta\tau(ex_{i_1\ldots i_k})-\tau\eta(ex_{i_1\ldots i_k}))+(\eta\tau(fx_{i_1\ldots i_k})-\tau\eta(fx_{i_1\ldots i_k})))\\
&\quad+\sum_{i_1\ldots i_k}\tau\eta(x_{i_1\ldots i_k})-\sum_\ell\sum_{i_\ell}\tau(e\eta(E_{N_\ell}x^\ell_{i_\ell}))-\sum_\ell\sum_{i_\ell}\tau\big(f\eta\big(E_{N_\ell}x^\ell_{i_\ell}\big)\big)\\
&=\left(\sum_{i_1\ldots i_k}\eta\tau(ex_{i_1\ldots i_k})-\sum_\ell\sum_{i_\ell}\tau\eta\big(E_{N_\ell}ex^\ell_{i_\ell}\big)\right)\\
&\quad+\left(\sum_{i_1\ldots i_k}\eta\tau(fx_{i_1\ldots i_k})-\sum_\ell\sum_{i_\ell}\tau\eta\big(E_{N_\ell}fx^\ell_{i_\ell}\big)\right)\\
&\leq\eta\tau(e)+\tau(e)H_{\tau_e}(N_1e,\ldots,N_ke)+\eta\tau(f)+\tau(f)H_{\tau_f}(N_1f,\ldots,N_kf)
\end{aligned}$$

Taking sup over all $(x_{i_1\ldots i_k})\in S_k$ we see that the left side is smaller than the right side in the formula in the lemma.

Conversely let $(y_{p_1,\ldots,p_k})\in N^+e$ with $\sum y_{p_1\ldots p_k}=e$, and $(z_{q_1,\ldots,q_k})\in N^+f$ with $\sum z_{q_1\ldots q_k}=f$. Put

$$x_{i_1\ldots i_k}=\begin{cases} y_{p_1\ldots p_k} & \text{if } i_\ell=2p_\ell \quad \forall l\\ z_{q_1\ldots q_k} & \text{if } i_\ell=2q_\ell-1 \quad \forall l\\ 0 & \text{otherwise}\end{cases}$$

Then

$$x^\ell_{i_\ell}=\begin{cases} y^\ell_{p_\ell} & \text{if } i_\ell=2p_\ell\\ z^\ell_{q_\ell} & \text{if } i_\ell=2q_\ell-1\end{cases}$$

We have

$$\begin{aligned}
&\sum_{i_1\ldots i_k}\eta\tau(x_{i_1\ldots i_k})-\sum_\ell\sum_{i_\ell}\tau\eta\big(E_{N_\ell}x^\ell_{i_\ell}\big)\\
&=\left(\sum_{p_1,\ldots,p_k}\eta\tau(y_{p_1\ldots p_k})-\sum_\ell\sum_{p_\ell}\tau\eta\big(E_{N_\ell e}y^\ell_{p_\ell}\big)\right)\\
&\qquad+\left(\sum_{q_1\ldots q_k}\eta\tau(z_{q_1\ldots q_k})+\sum_\ell\sum_{q_\ell}\tau\eta\big(E_{N_\ell f}z^\ell_{q_\ell}\big)\right)\\
&=\eta\tau(e)+\tau(e)\left(\sum_{p_1\ldots p_k}\eta\tau_e(y_{p_\ell\ldots p_k})-\sum_\ell\sum_{p_\ell}\tau_e\eta\big(E_{N_\ell e}y^\ell_{p_\ell}\big)\right)\\
&\quad+\eta\tau(f)+\tau(f)\left(\sum_{q_1\ldots q_k}\eta\tau_f(z_{q_1\ldots q_k})-\sum_\ell\sum_{q_\ell}\tau_f\eta\big(E_{N_\ell f}z^\ell_{q_\ell}\big)\right)
\end{aligned}$$

Taking sup over all $(y_{p_1\dots p_k})$ and $(z_{q_1\dots q_k})$ we find

$$H(N_1,\dots,N_k) \geq \eta\tau(e) + \tau(e)H_{\tau_e}(N_1e,\dots,N_ke) + \eta\tau(f) + \tau(f)H_{\tau_f}(N_1f,\dots,N_kf)\ ,$$

proving the lemma. □

Lemma 4.3 *Let* $N_1,\dots,N_k$ *be finite dimensional von Neumann subalgebras of* R, *and* $N = \bigvee_{\ell=1}^{k} N_\ell$. *Let* e *be a central projection in* N, *and suppose* $H(N_1,\dots,N_k) = H(N)$. *Then* $H_{\tau_e}(N_1e,\dots,N_ke) = H_{\tau_e}(Ne)$.

Proof. Let M_ℓ denote the von Neumann algebra generated by N_ℓ and e. Then $N_\ell \subset M_\ell \subset N$, so that $N = \bigvee_{\ell=1}^{k} M_\ell$, and $M_\ell e = N_\ell e$. By property (A) and (C) and Lemma 4.2 applied first to $H(M_1,\dots,M_k)$ and then to N, we have

$$\begin{aligned} H(N) &= H(N_1,\dots,N_k)\\ &\leq H(M_1,\dots,M_k)\\ &= \tau\eta(e) + \tau\eta(f) + \tau(e)H_{\tau_e}(N_1e,\dots,N_ke) + \tau(f)H_{\tau_f}(N_1f,\dots,N_kf)\\ &\leq \tau\eta(e) + \tau\eta(f) + \tau(e)H_{\tau_e}(Ne) + \tau(f)H_{\tau_f}(Nf)\\ &= H(N)\ . \end{aligned}$$

It follows that $\tau(e)H_{\tau_e}(N_1e,\dots,N_ke) = \tau(e)H_{\tau_e}(Ne)$, proving the lemma. □

Lemma 4.4 *Let* N *be a von Neumann subalgebra of* R *such that* $H(N) < \infty$. *Then* N *is finite of type I with totally atomic center.*

Proof. If the conclusion does not hold there exists a weakly closed abelian C^*-subalgebra A of N without minimal projections. Say e is the identity for A. Then for each $n \in \mathbb{N}$ there exists an orthogonal family $e_1,\dots,e_n$ of projections in A with sum e and $\tau(e_i) = \frac{1}{n}\tau(e)$. Thus if $f = 1 - e$,

$$\begin{aligned} H(N) &\geq H(A \oplus \mathbb{C}f)\\ &\geq \sum_{i=1}^{n} \eta(\tau(e)\tfrac{1}{n}) + \eta\tau(f)\\ &= \tau(e)\log n + \tfrac{1}{n}\eta\tau(e) + \eta\tau(f)\ , \end{aligned}$$

proving that $H(N) = +\infty$, contrary to assumption. □

5 The main theorem

In this section we prove our main theorem. Recall that if N is a finite dimensional von Neumann algebra then the rank of N, $\operatorname{rank} N$, is the dimension of a masa in N. Thus $\dim N \le (\operatorname{rank} N)^2$.

Theorem 5.1 *Let R be a finite von Neumann algebra with a faithful normal trace τ with $\tau(1) = 1$. Let $N_1, \ldots, N_k$ be finite dimensional von Neumann subalgebras of R and let $N = \bigvee_{\ell=1}^{k} N_\ell$. Then the following two conditions are equivalent*

(i) $H(N_1, \ldots, N_k) = H(N)$.

(ii) *There exists a masa A in N such that $A = \bigvee_{\ell=1}^{k} (A \cap N_\ell)$.*

In particular, if the above conditions hold, then N is finite dimensional and $\operatorname{rank} N \le \prod_{\ell=1}^{k} \operatorname{rank} N_\ell$.

Proof of Theorem 5.1 (part 1). The implication (ii) $\Rightarrow$(i) is well-known and follows easily from properties (A), (C), (D) and (E). Indeed,

$$H(N) \ge H(N_1, \ldots, N_k) \ge H(A \cap N_1, \ldots, A \cap N_k) = H\left(\bigvee_{\ell=1}^{k} (A \cap N_\ell)\right)$$
$$= H(A) = H(N) \ .$$

From now on we assume (i). By property (B) $H(N) \le \sum_{\ell=1}^{k} H(N_\ell) < \infty$, hence N is by Lemma 4.4 finite of type I with totally atomic center. By Lemma 4.3 assumption (i) holds for Ne and $N_1 e, \ldots, N_k e$ for each central projection e in N. Since the center of N is totally atomic the identity 1 is the sup of central projections e with $\dim Ne < \infty$. If we prove the theorem for Ne we conclude in particular that $\operatorname{rank} Ne \le \prod_{\ell=1}^{k} \operatorname{rank} N_\ell$. Since this holds for all such e it holds for N itself, i.e. N is finite dimensional with $\operatorname{rank} N \le \prod_{\ell=1}^{k} \operatorname{rank} N_\ell$. We shall therefore in the sequel assume $\dim N < \infty$. In order to prove Theorem 5.1 we need to replace $(x_{i_1 \ldots i_k}) \in S_k$ and the operators $x_{i_\ell}^{\ell}$ by functions with values in N. Since $\dim N < \infty$ we can consider N as a subset of $\mathbb{R}^r$ for some $r \in \mathbb{N}$. Let

$$c = \sup\left\{ \frac{\|x\|}{\tau(x)} : 0 \ne x \in N, 0 \le x \le 1 \right\} \ .$$

Since $\dim N < \infty$, $c < \infty$, It is clear that Corollary 2.6 holds with the cube $[0,1]^r$ replaced by $[0,c]^r$. We therefore let ψ: $[0,c]^r \to [0,1]$ be a function with the properties of Corollary 2.6. Let $(x_{i_1 \ldots i_k}) \in S_k$, and assume, as we may, that $x_{i_1 \dots i_k} \in N$. For each ℓ let the numbers i_ℓ run through the numbers $\{1, 2, \dots, n_\ell\}$. Put

$$c^\ell_{i_\ell} = \psi(\tau(x^\ell_{i_\ell})^{-1} x^\ell_{i_\ell}), \qquad i_\ell = 1, \dots, n_\ell \ .$$

Let σ_ℓ be a permutation of $\{1, \dots, n_\ell\}$ such that

$$c^\ell_{\sigma_\ell(1)} \leq c^\ell_{\sigma_\ell(2)} \leq \cdots \leq c^\ell_{\sigma_\ell(n_\ell)} \ .$$

The defining sums in the definition of $H(N_1, \dots, N_k)$ remain the same if we replace i_ℓ by $\sigma_\ell(i_\ell)$, or $(i_1, \dots, i_k)$ by $(\sigma_1(i_1), \dots, \sigma_k(i_k))$. We can therefore assume $c^\ell_1 \leq c^\ell_2 \leq \cdots \leq c^\ell_{n_\ell}$, $\ell = 1, \dots, k$. Choose numbers in $[0, 1]$ as follows:

$$\begin{gathered} 0 = a^\ell_0 < a^\ell_1 < \cdots < a^\ell_{n_\ell} = 1, \\ a^\ell_j - a^\ell_{j-1} = \tau(x^\ell_j) \ . \end{gathered}$$

Put $A^\ell_j = [a^\ell_{j-1}, a^\ell_j)$, $j < n_\ell$, $A^\ell_{n_\ell} = [a^\ell_{n_\ell - 1}, 1]$. Put

$$g^\ell(t) = \sum_{j=1}^{n_\ell} \chi_{A^\ell_j}(t) \tau(x^\ell_j)^{-1} x^\ell_j, \qquad t \in [0, 1] \ ,$$

where χ_A is the characteristic function of the set A, or in our previous notation, if we write A_{i_ℓ} for $A^\ell_{i_\ell}$,

$$g^\ell(t) = \sum_{i_\ell} \chi_{A_{i_\ell}}(t) \tau(x^\ell_{i_\ell})^{-1} x^\ell_{i_\ell}, \qquad t \in [0, 1] \ .$$

Then $\psi \circ g^\ell$ is an increasing function $[0, 1]$ into $[0, 1]$. Put

$$g(t_1, \dots, t_k) = \sum_{i_1 \dots i_k} \chi_{A_{i_1} \times \cdots \times A_{i_k}}(t_1, \dots, t_k) \left(\prod_{\ell=1}^k \tau(x^\ell_{i_\ell}) \right)^{-1} x_{i_1 \dots i_k}$$

Then $g\colon [0, 1]^k \to N^+ \subset \mathbb{R}^r$ is Borel. Note that we have

$$\begin{gathered} \int_0^1 \cdots \int_0^1 g(t_1, \dots, t_k) dt_1 \dots dt_k = 1 \\ \int_0^1 g^\ell(t) dt = 1 \ . \end{gathered}$$

To complete the proof of Theorem 5.1 we shall need the following two lemmas.

Lemma 5.2 *With the above notation we have*

$$\int_0^1 \cdots \int_0^1 \eta\tau(g(t_1,\ldots,t_k))dt_1\ldots dt_k - \sum_{\ell=1}^{k}\int_0^1 \tau\eta E_{N_\ell}(g^\ell(t_\ell))dt_\ell$$
$$= \sum_{i_1\ldots i_k}\eta\tau(x_{i_1\ldots i_k}) - \sum_{\ell=1}^{n}\sum_{i_\ell}\tau\eta E_{N_\ell}(x^\ell_{i_\ell}) \ .$$

Proof. The proof is a computation of the integrals involved.

$$\int_0^1 \cdots \int_0^1 \eta\tau(g(t_1,\ldots,t_k))dt_1\ldots dt_k$$
$$= \int_0^1 \cdots \int_0^1 \eta\tau\left(\sum_{i_1\ldots i_k}\chi_{A_{i_1}}(t_1)\ldots\chi_{A_{i_k}}(t_k)\left(\prod_{\ell=1}^{k}\tau\left(x^\ell_{i_\ell}\right)\right)^{-1}x_{i_1\ldots i_k}\right)dt_1\ldots dt_k$$
$$= \sum_{i_1\ldots i_k}\int_0^1 \cdots \int_0^1 \chi_{A_{i_1}}(t_1)\ldots\chi_{A_{i_k}}(t_k)\eta\left(\left(\prod_{\ell=1}^{k}\tau\left(x^\ell_{i_\ell}\right)\right)^{-1}\tau(x_{i_1\ldots i_k})\right)dt_1\ldots dt_k$$
$$= \sum_{i_1\ldots i_k}\prod_{\ell=1}^{k}\tau(x^\ell_{i_\ell})\left(\left(\prod_{\ell=1}^{k}\tau\left(x^\ell_{i_\ell}\right)\right)^{-1}\eta\tau(x_{i_1\ldots i_k}) + \eta\left(\left(\prod_{\ell=1}^{k}\tau\left(x^\ell_{i_\ell}\right)\right)^{-1}\right)\tau(x_{i_1\ldots i_k})\right)$$
$$= \sum_{i_1\ldots i_k}\eta\tau(x_{i_1\ldots i_k}) + \log\left(\prod_{\ell=1}^{k}\tau\left(x^\ell_{i_\ell}\right)\right)\tau(x_{i_1\ldots i_k})$$
$$= \sum_{i_1\ldots i_k}\eta\tau(x_{i_1\ldots i_k}) + \tau(x_{i_1\ldots i_k})\sum_{\ell=1}^{k}\log\tau(x^\ell_{i_\ell})$$
$$= \sum_{i_1\ldots i_k}\eta\tau(x_{i_1\ldots i_k}) + \sum_{i_1\ldots i_k}\tau(x_{i_1\ldots i_k})\log\tau(x^\ell_{i_\ell})$$
$$= \sum_{i_1\ldots i_k}\eta\tau(x_{i_1\ldots i_k}) - \sum_{\ell=1}^{k}\sum_{i_\ell}\eta\tau(x^\ell_{i_\ell}) \ .$$

Similarly we have

$$\int_0^1 \tau\eta E_{N_\ell}(g^\ell(t))dt = \int_0^1 \sum_{i_\ell}\chi_{A_{i_\ell}}(t)\tau\eta(\tau(x^\ell_{i_\ell})^{-1}E_{N_\ell}x^\ell_{i_\ell})dt$$
$$= \sum_{i_\ell}\tau(x^\ell_{i_\ell})\tau(\tau(x^\ell_{i_\ell})^{-1}\eta(E_{N_\ell}x^\ell_{i_\ell}) + \eta(\tau(x^\ell_{i_\ell})^{-1})E_{N_\ell}x^\ell_{i_\ell})$$
$$= \sum_{i_\ell}\tau\eta(E_{N_\ell}x^\ell_{i_\ell}) + \log\tau(x^\ell_{i_\ell})\tau(E_{N_\ell}x^\ell_{i_\ell})$$
$$= \sum_{i_\ell}\tau\eta(E_{N_\ell}x^\ell_{i_\ell}) - \sum_{i_\ell}\eta\tau(x^\ell_{i_\ell}) \ .$$

Subtraction of the second formula from the first yields the lemma. □

Lemma 5.3 *Let $N_1, \ldots, N_k$ be finite dimensional von Neumann subalgebras of R and $N = \bigvee_{\ell=1}^{k} N_\ell$. Suppose $\varepsilon > 0$ and $(x_{i_1 \ldots i_k}) \in S_k$ satisfy*

$$H(N) - \varepsilon < \sum_{i_1 \ldots i_k} \eta\tau(x_{i_1 \ldots i_k}) - \sum_{\ell=1}^{k} \sum_{i_\ell} \tau\eta(E_{N_\ell} x^{\ell}_{i_\ell}) \ .$$

Then we have

(i) $\sum_{\ell=1}^{k} \sum_{i_\ell} \tau\eta(x^{\ell}_{i_\ell}) < \sum_{i_1 \ldots i_k} \tau\eta(x_{i_1 \ldots i_k}) + \varepsilon$

(ii) $\sum_{\ell=1}^{k} \sum_{i_\ell} (\tau\eta(E_{N_\ell} x^{\ell}_{i_\ell}) - \tau\eta(x^{\ell}_{i_\ell})) < \varepsilon$.

Note that by Jensen's inequality (6) $E_{N_\ell}\eta(x^{\ell}_{i_\ell}) \leq \eta(E_{N_\ell} x^{\ell}_{i_\ell})$, so each term in (ii) is nonnegative.

Proof. We may assume each $x_{i_1 \ldots i_k} \in N$. Since $\tau\eta(E_{N_\ell} x_{i_\ell}) \geq \tau\eta(x^{\ell}_{i_\ell})$ we have by Corollary 3.3

$$\begin{aligned}
H(N) - \varepsilon &< \sum_{i_1 \ldots i_k} \eta\tau(x_{i_1 \ldots i_k}) - \sum_{\ell=1}^{k} \sum_{i_\ell} \tau\eta(E_{N_\ell} x^{\ell}_{i_\ell}) \\
&\leq \sum_{i_1 \ldots i_k} \eta\tau(x_{i_1 \ldots i_k}) - \sum_{\ell} \sum_{i_\ell} \tau\eta(x^{\ell}_{i_\ell}) \\
&\leq \sum_{i_1 \ldots i_k} \eta\tau(x_{i_1 \ldots i_k}) - \sum_{i_1 \ldots i_k} \tau\eta(x_{i_1 \ldots i_k}) \\
&= \sum_{i_1 \ldots i_k} \eta\tau(x_{i_1 \ldots i_k}) - \sum_{i_1 \ldots i_k} \tau\eta(E_N x_{i_1 \ldots i_k}) \\
&\leq H(N) \ ,
\end{aligned}$$

where the last inequality follows since we can consider $(x_{i_1 \ldots i_k})$ as an element in S_1 by sufficient reindexing, and

$$H(N) = \sup_{(x_i) \in S_1} \left(\sum_i \eta\tau(x_i) - \sum_i \tau\eta(E_N x_i) \right) .$$

The conclusion of the lemma is now immediate. □

Proof of Theorem 5.1. (*part* 2). We assume (i) so there is a sequence $(x^n_{i_1 \ldots i_k})_{n \in \mathbb{N}}$ in S_k such that

$$H(N) = \lim_{n \to \infty} \left(\sum_{i_1 \ldots i_k} \eta\tau\left(x^n_{i_1 \ldots i_k}\right) - \sum_{\ell=1}^{k} \sum_{i_\ell} \tau\eta\left(E_{N_\ell} x^{n\ell}_{i_\ell}\right) \right) .$$

Let g_n and g^{ℓ}_n denote the functions corresponding to $(x^n_{i_1 \ldots i_k})$ as defined before Lemma 5.2. By that lemma we have

$$(8)\ H(N) = \lim_{n\to\infty}\left(\int_0^1\cdots\int_0^1 \eta\tau(g_n(t_1,\ldots,t_k))dt_1\ldots dt_k - \sum_{\ell=1}^{k}\int_0^1 \tau\eta(E_{N_\ell}g_n^\ell(t_\ell))dt_\ell\right)$$

By Lemma 5.3 we find

$$(9)\qquad \lim_{n\to\infty}\sum_{\ell=1}^{k}\int_0^1(\tau\eta(E_{N_\ell}g_n^\ell(t_\ell)) - \tau\eta(g_n^\ell(t_\ell)))dt_\ell = 0\ .$$

By Lemma 5.3 and Corollary 3.3 together with a straightforward computation we find

$$(10)\qquad \lim_{n\to\infty}\int_0^1\int_0^1 \left\|\left[g_n^\ell(t_\ell)^{1/2}, g_n^m(t_m)^{1/2}\right]\right\|_2^2 dt_\ell dt_m = 0 \qquad \text{for } \ell \neq m$$

By Corollary 2.6 and the construction of the functions (g_n^ℓ) we can choose a sequence $(n_i)_{i\in\mathbb{N}}$ such that the subsequences $(g_{n_i}^\ell)$ converge pointwise almost everywhere. Reindexing we can therefore find functions $g^\ell\colon[0,1]\to\{x\in N\colon 0\le x\le c1\}$ such that $g_n^\ell\to g^\ell$ pointwise a.e. Since $\|g_n^\ell\|\le c$ the Lebesgue dominated convergence theorem applies. From (9) we get

$$\sum_{\ell=1}^{k}\int_0^1(\tau\eta(E_{N_\ell}g^\ell(t_\ell)) - \tau\eta(g^\ell(t_\ell)))dt_\ell = 0$$

Since the integrand is nonnegative this implies

$$\begin{aligned}\tau\eta(E_{N_\ell}g^\ell(t_\ell)) &= \tau\eta(g^\ell(t_\ell)) \text{ a.e.}\\ &= \tau E_{N_\ell}\eta(g^\ell(t_\ell)) \text{ a.e.}\end{aligned}$$

Thus by (6) and faithfulness of τ,

$$\eta(E_{N_\ell}g^\ell(t_\ell)) = E_{N_\ell}\eta(g^\ell(t_\ell))\ ,$$

hence by (7) we have $g^\ell(t_\ell)\in N_\ell$ a.e.

Similarly by (10) we find

$$\left[g^\ell(t_\ell)^{1/2}, g^m(t_m)^{1/2}\right] = 0 \quad \text{a.e. in } [0,1]^2 \text{ for } l\neq m\ ,$$

hence $[g^\ell(t_\ell), g^m(t_m)] = 0$ a.e. when $l\neq m$.

We assert that the operators $g^\ell(s)$ and $g^m(t)$ commute for almost all s and almost all t. Indeed, put

$$K_\ell = \left\{ \int_0^1 g^\ell(s)\phi(s)ds \colon \phi \in C([0,1]) \right\}$$
$$K_m = \left\{ \int_0^1 g^m(t)\psi(t)dt \colon \psi \in C([0,1]) \right\} .$$

Operators in K_ℓ commute with those in K_m because

$$\begin{aligned}
&\left(\int_0^1 g^\ell(s)\phi(s)ds \right) \left(\int_0^1 g^m(t)\psi(t)dt \right) \\
&\quad = \int_0^1 \int_0^1 g^\ell(s)g^m(t)\phi(s)\psi(t)ds\, dt \\
&\quad = \int_0^1 \int_0^1 g^m(t)g^\ell(s)\phi(s)\psi(t)ds\, dt \\
&\quad = \left(\int_0^1 g^m(t)\psi(t)dt \right) \left(\int_0^1 g^\ell(s)\phi(s)dt \right)
\end{aligned}$$

Consider N in its Hilbert-Schmidt norm. Then K_ℓ and K_m are closed subspaces. Choose $h_1, \ldots, h_p$ which span $K_\ell^\perp$. Then

$$\int_0^1 \langle g^\ell(s), h_i \rangle \phi(s)ds = 0 \quad \forall\ \phi \in C([0,1])$$

Hence $\langle g^\ell(s), h_i \rangle = 0$ a.e. Thus $g^\ell(s) \in K_\ell^{\perp\perp} = K_\ell$ a.e., and similarly $g^m(t) \in K_m$ a.e., proving that $[g^\ell(s), g^m(t)] = 0$ for almost all s and almost all t, as asserted.

From the above we can choose nullsets $V_\ell \subset [0,1]$ such that $g^\ell(s) \in N_\ell$ whenever $s \notin V_\ell$, and $[g^\ell(s), g^m(t)] = 0$ whenever $s \notin V_\ell$, $t \notin V_m$. Let B_ℓ be the von Neumann subalgebra of N_ℓ generated by $g^\ell(s)$, $s \notin V_\ell$. Then $B_\ell \subset B'_m \cap N$ whenever $l \neq m$. Since $\tau \circ \eta \circ E_{B_\ell}$ is continuous

$$\tau\eta(E_{B_\ell} g_n^\ell) \to \tau\eta(E_{B_\ell} g^\ell) = \tau\eta(g^\ell) \quad \text{pointwise a.e.}$$

Thus by the Lebesgue dominated convergence theorem

$$\int_0^1 \tau\eta(E_{B_\ell} g_n^\ell(t_\ell))dt_\ell \to \int_0^1 \tau\eta(g^\ell(t_\ell))dt_\ell, \qquad \ell = 1, \ldots, k.$$

Similarly

$$\int_0^1 \tau\eta(E_{N_\ell} g_n^\ell(t_\ell))dt_\ell \to \int_0^1 \tau\eta(g^\ell(t_\ell))dt_\ell, \qquad \ell = 1, \dots, k.$$

Thus from (8) we get

$$H(N) = \lim_{n\to\infty}\left(\int_0^1 \cdots \int_0^1 \eta\tau(g_n(t_1, \dots, t_k))dt_1 \dots dt_k - \sum_{\ell=1}^k \int_0^1 \tau\eta(E_{B_\ell} g_n^\ell(t_\ell))dt_\ell\right)$$

By Lemma 5.2 this means

$$H(N) = \lim_{n\to\infty}\left(\sum_{i_1\dots i_k} \eta\tau(x^n_{i_1\dots i_k}) - \sum_{\ell=1}^k \sum_{i_\ell} \tau\eta\left(E_{B_\ell} x^{n\ell}_{i_\ell}\right)\right)$$

By definition of $H(B_1, \dots, B_k)$ the right side of the last equation is majorized by $H(B_1, \dots, B_k)$. By property (C) we therefore have, letting $B = \bigvee_{\ell=1}^k B_\ell$,

$$H(N) \le H(B_1, \dots, B_k) \le H(B) \le H(N) \ .$$

In particular, by Lemma 4.1 each masa in B is a masa in N. Let A_ℓ be a masa in B_ℓ. Since the B_ℓ all commute with each other, the abelian algebra $A = \bigvee_{\ell=1}^k A_\ell$ is a masa in B and hence in N. Since $A = \bigvee_{\ell=1}^k (A \cap N_\ell)$ the proof is complete. □

Theorem 5.1 sheds light on two problems which have well-known classical analogues. The first says in our notation that if A and B are commuting finite dimensional abelian von Neumann subalgebras of R then $H(A, B) = H(A) + H(B)$ if and only if A and B are independent, i.e. $A \vee B \cong A \otimes B$ and $\tau(ab) = \tau(a)\tau(b)$ when $a \in A$, $b \in B$. One might believe that something similar holds in the nonabelian situation. The following example shows that an affirmative result must be quite restrictive.

Example 5.4 Let $N_1 = \left\{\begin{pmatrix} \alpha 1, & \beta 1 \\ \gamma 1, & \delta 1\end{pmatrix} \in M_4(\mathbb{C}) : 1 \text{ identity in } M_2(\mathbb{C})\right\}$, so $N_1 = \mathbb{C} \otimes M_2(\mathbb{C}) \subset M_4(\mathbb{C})$. Let

$$v = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \in M_4(\mathbb{C}) \ .$$

Let $N_2 = vN_1v^*$. If A_1 is the diagonal matrices, $\begin{pmatrix} \alpha 1 & 0 \\ 0 & \delta 1\end{pmatrix}$ in N_1, then $A_2 = vA_1v^*$ is an abelian subalgebra of $M_4(\mathbb{C})$ which together with A_1 generates the diagonal matrices A in $M_4(\mathbb{C})$. We thus have

$$H(N_1, N_2) \geq H(A_1, A_2) = H(A) = H(M_4(\mathbb{C})) \ ,$$

hence

$$\begin{aligned} H(N_1, N_2) =& H(N_1 \vee N_2) = \log 4 = \log 2 + \log 2 \\ =& H(N_1) + H(N_2) \ . \end{aligned}$$

However, N_1 and N_2 do not even commute.

The restricted theorem shows that the existence of A_1 and A_2 is generic.

Corollary 5.5 *Let $N_1, \ldots, N_k$ be finite dimensional von Neumann subalgebras of R, and let $N = \bigvee_{\ell=1}^{k} N_\ell$. Assume*

$$H(N) = H(N_1, \ldots, N_k) = \sum_{\ell=1}^{k} H(N_\ell)$$

Then there exists a masa A in N such that $A_\ell = A \cap N_\ell$ is a masa in N_ℓ, and $A_1, \ldots, A_k$ are independent.

Proof. By Theorem 5.1 there is a masa A in N such that $A = \bigvee_{\ell=1}^{k} A_\ell$, $A_\ell = A \cap N_\ell$. We thus have

$$\begin{aligned} H(N) = H(A) =& H(A_1, \ldots, A_k) \leq \sum_{\ell=1}^{k} H(A_\ell) \\ \leq& \sum_{\ell=1}^{k} H(N_\ell) = H(N) \ . \end{aligned}$$

It follows that $H(A_\ell) = H(N_\ell)$, hence by Lemma 4.1 A_ℓ is a masa in N_ℓ, and

$$H(A) = H(A_1, \ldots, A_k) = \sum_{\ell=1}^{k} H(A_\ell) \ .$$

The conclusion of the corollary now follows from the classical abelian case. □

The second problem which is true in the classical case but false in the nonabelian case is that of additivity of entropy of tensor products of automorphisms, see [8] and [10]. Our next result can be used to show that it is true when one has maximality assumptions like those in Theorem 5.1.

Corollary 5.6 *Let R_1 and R_2 be finite von Neumann algebras with faithful normal traces τ_1 and τ_2 with $\tau_i(1) = 1$. Suppose $M_1, \ldots, M_k \subset R_1$, $N_1, \ldots, N_k \subset R_2$ are finite dimensional von Neumann subalgebras. Let $M = \bigvee_{\ell=1}^{k} M_\ell$, $N = \bigvee_{\ell=1}^{k} N_\ell$. Suppose*

$$H_{\tau_1}(M_1,\ldots,M_k)=H_{\tau_1}(M), H_{\tau_2}(N_1,\ldots,N_k)=H_{\tau_2}(N)\ .$$

Then

$$H_{\tau_1\otimes\tau_2}(M_1\otimes N_1,\ldots,M_k\otimes N_k)=H_{\tau_1\otimes\tau_2}(M\otimes N)=H_{\tau_1}(M)+H_{\tau_2}(N)$$

Proof. By Theorem 5.1 there exist masas $A\subset M$ and $B\subset N$ such that

$$A=\bigvee_{\ell=1}^{k}(A\cap M_\ell),\qquad B=\bigvee_{\ell=1}^{k}(B\cap N_\ell)\ .$$

Then

$$A\otimes B=\bigvee_{\ell=1}^{k}(A\cap M_\ell)\otimes(B\cap N_\ell)=\bigvee_{\ell=1}^{k}(A\otimes B)\cap(M_\ell\otimes N_\ell)$$

Since $A\otimes B$ is a masa in $M\otimes N$ it follows that

$$\begin{aligned}H(M\otimes N)&=H(A\otimes B)\\&=H((A\otimes B)\cap(M_1\otimes N_1),\ldots,(A\otimes B)\cap(M_k\otimes N_k))\\&\le H(M_1\otimes N_1,\ldots,M_k\otimes N_k)\\&\le H(M\otimes N)\ .\end{aligned}$$

This proves the first identity in the corollary. The second follows since $H(A\otimes B)=H(A)+H(B)=H(M)+H(N)$. □

References

1. Choi, M.D., Positive linear maps on C^*-algebras, Thesis, Univ. of Toronto 1972
2. Connes, A. and Størmer, E., Entropy of automorphisms of II_1 von Neumann algebras, Acta math. **134**, 289–306 (1975)
3. Hansen, F. og Pedersen, G.K., Jensen's inequality for operators and Löwner's theorem, Math. Ann. **258**, 229–241 (1982)
4. Kadison, R. and Ringrose, J., Fundamentals of the theory of operator algebras, vol. II, Academic Press 1986
5. Kunze, R., Lp-Fourier transform on locally compact unimodular groups, Trans Amer. Math. Soc. **89**, 519–540 (1958)
6. Lieb, E., Convex trace functions and the Wigner-Yanase-Dyson conjecture, Advances in Math. **11**, 267–288 (1973)
7. Nakamura, M. and Umegaki, H., A note on the entropy for operator algebras, Proc. Japan Acad. **37**, 149–154 (1961)
8. Narnhofer, H., Størmer, E. and Thirring, W., C^*-dynamical systems for which the tensor product formula for entropy fails, Ergod. Th. and Dynam. Sys. **15** 961–968 (1995)
9. Reed, M. and Simon, B., Methods of Modern Mathematical physics, Vol. I, Academic Press, New York 1972
10. Sauvageot, J.-L., Ergodic properties of the action of a matrix in $SL(2,\mathbb{Z})$ on a noncommutative 2-Torus