10

Relative Entropy and Subfactors

Conditional entropy is, as we saw in Chap. 1, an important concept in the classical theory. In the noncommutative case it was introduced under the name of relative entropy as a tool to take care of approximation of entropy. We shall in the present chapter develop the theory of relative entropy and show its relationship to subfactors of II_1-factors. Then we shall show a formula analogous to the classical formula $h(T) = H(\xi|\xi^-)$. Finally we shall give applications to the canonical shift on the tower of relative commutants defined by an inclusion of II_1-factors and for shifts on the Jones projections.

10.1 Relative Entropy

Let M be a von Neumann algebra, φ a normal state on M, P and Q von Neumann subalgebras of M.

Definition 10.1.1. *The relative entropy of P and Q with respect to φ is*

$$H_\varphi(P|Q) = \sup \sum_i (S(\varphi_i|_P, \varphi|_P) - S(\varphi_i|_Q, \varphi|_Q)),$$

where the supremum is taken over all finite decompositions $\varphi = \sum_i \varphi_i$ of φ into a sum of positive linear functionals.

Note that if $\varphi = \tau$ is a trace, which is the case we are mainly interested in, any $\psi \leq \tau$ has the form $\psi = \tau(\cdot x)$. Hence by Theorem 2.3.1(x) the relative entropy can be written as

$$H_\tau(P|Q) = \sup \sum_i (\tau(\eta(E_Q(x_i))) - \tau(\eta(E_P(x_i)))), \tag{10.1}$$

where the supremum is taken over all finite partitions of unity $1 = \sum_i x_i$ in M, and E_P and E_Q are the τ-preserving conditional expectations on P and Q, respectively. One often suppresses τ in the notation for relative entropy.

The main properties of relative entropy are as follows.

Theorem 10.1.2. *We have:*

(i) $H_\varphi(P|Q) \geq 0$, *and if* φ *is a faithful tracial state then* $H_\varphi(P|Q) = 0$ *if and only if* $P \subset Q$;

(ii) $H_\varphi(P|R) \leq H_\varphi(P|Q) + H_\varphi(Q|R)$;

(iii) $H_\varphi(P|Q)$ *is increasing in* P *and decreasing in* Q;

(iv) *if* $Q \subset P$ *and there is a* φ*-preserving faithful normal conditional expectation* $E: P \to Q$, *then*

$$H_\varphi(P|Q) = \sup \sum_i S(\varphi_i|_P, \varphi_i \circ E),$$

where the supremum is taken over all finite decompositions $\varphi = \sum_i \varphi_i$.

Proof. Taking the trivial decomposition $\varphi = \varphi$ we see that $H_\varphi(P|Q) \geq 0$. If $P \subset Q$ then $H_\varphi(P|Q) = 0$ by monotonicity of relative entropy S, Theorem 2.3.1(vi). Let now $\varphi = \tau$ be a faithful tracial state. If P is not a subset of Q, there exists a projection $e \in P$ such that $e \notin Q$. Then by (10.1)

$$H_\tau(P|Q) \geq \tau(\eta(E_Q(e))) + \tau(\eta(E_Q(1-e))).$$

The element $E_Q(e)$ is not a projection (this follows e.g. by A.13, as if $E_Q(e)$ is a projection then $e \leq \mathrm{s}(E_Q(e)) = E_Q(e)$, where $\mathrm{s}(a)$ is the support of a self-adjoint element a, and hence $e = E_Q(e)$ by faithfulness of E_Q). Therefore $\tau(\eta(E_Q(e))) > 0$ and similarly $\tau(\eta(E_Q(1-e))) > 0$. This completes the proof of (i).

Part (ii) is immediate from the definition of $H_\varphi(P|Q)$. Part (iii) follows from monotonicity of relative entropy S, while part (iv) follows from Theorem 2.3.1(vii). □

Let us show next that relative entropy coincides with conditional entropy in the abelian case. So let $M = L^\infty(X, \mu)$, ξ and ζ be measurable partitions of X such that ξ is finite, $P = L^\infty(X/\xi)$ and $Q = L^\infty(X/\zeta)$. We want to show that

$$H_\tau(P|Q) = H_\mu(\xi|\zeta).$$

If $p_1, \ldots, p_n$ are the atoms of P, by (10.1) we have

$$H_\tau(P|Q) \geq \sum_i \tau(\eta(E_Q(p_i))).$$

The latter expression is exactly $H_\mu(\xi|\zeta)$. Thus $H_\tau(P|Q) \geq H_\mu(\xi|\zeta)$. It suffices to prove the opposite inequality for finite ζ. Indeed, if $\{\zeta_n\}_n$ is an increasing sequence of finite measurable partitions such that $\vee_n \zeta_n = \zeta$ then

$$H_\tau(P|L^\infty(X/\zeta_n)) \geq H_\tau(P|Q) \geq H_\mu(\xi|\zeta),$$

and $H_\mu(\xi|\zeta_n) \searrow H_\mu(\xi|\zeta)$ by the martingale convergence theorem. Thus if $H_\tau(P|L^\infty(X/\zeta_n)) = H_\mu(\xi|\zeta_n)$, we conclude that $H_\tau(P|Q) = H_\mu(\xi|\zeta)$.

So assume ζ is finite. Then

$$H_\tau(P|Q) \le H_\tau(P \vee Q|Q) \quad \text{and} \quad H_\mu(\xi|\zeta) = H_\mu(\xi \vee \zeta|\zeta).$$

It follows that to prove the inequality $H_\tau(P|Q) \le H_\mu(\xi|\zeta)$ we may assume that $\zeta \prec \xi$, so $Q \subset P$. In this case it is enough to consider partitions of unity in P. Then part (iv) of Theorem 10.1.2 and convexity of relative entropy of positive functionals, Corollary 2.3.2, show that it suffices to consider partitions $1 = \sum_i x_i$ such that each x_i is a scalar multiple of a minimal projection in P. Since $S(\lambda\psi, \lambda\varphi) = \lambda S(\psi, \varphi)$, we then see that this is the same as to consider one partition $1 = \sum_i p_i$, where $p_1, \dots, p_n$ are the minimal projections in P. But then clearly $H_\tau(P|Q) = H_\mu(\xi|\zeta)$.

Next we want to show that in some cases the computation of relative entropy can be reduced to the finite dimensional case. For this we need the following notion. If P, Q, R are von Neumann subalgebras of a von Neumann algebra M, and $E_P\colon M \to P$, $E_Q\colon M \to Q$, $E_R\colon M \to R$ are conditional expectations, then we say that

$$\begin{array}{ccc} P & \subset & M \\ \cup & & \cup \\ R & \subset & Q \end{array}$$

is a **commuting square** if $E_R = E_Q \circ E_P = E_P \circ E_Q$. Equivalently, $R = P \cap Q$ and $E_Q(P) \subset R$, $E_P(Q) \subset R$.

Remark that if the conditional expectations preserve a faithful normal state φ, then is suffices to check that, say, $E_R = E_Q \circ E_P$. Indeed, if $M \subset B(H)$ and the state φ is defined by a cyclic and separating vector $\xi \in H$, then $E_R(x)\xi = e_R x\xi$ for $x \in M$, where $e_R \in B(H)$ is the projection onto $\overline{R\xi}$. Then $E_R = E_Q \circ E_P$ implies that $e_R = e_Q e_P$. But then $e_R = e_P e_Q$, and so $E_R = E_P \circ E_Q$.

Proposition 10.1.3. *Let M be a von Neumann algebra with a faithful normal state φ, N a von Neumann subalgebra. Suppose $\{M_n\}_n$ and $\{N_n\}_n$ are increasing sequences of von Neumann subalgebras of M such that $N_n \subset M_n$, $M = (\cup_n M_n)''$ and $N = (\cup_n N_n)''$. Suppose there exist φ-preserving conditional expectations $E_{M_n}\colon M \to M_n$ and $E_{N_n}\colon M \to N_n$ such that*

$$\begin{array}{ccc} M_n & \subset & M_{n+1} \\ \cup & & \cup \\ N_n & \subset & N_{n+1} \end{array}$$

is a commuting square for every n. Then $H_\varphi(M|N) = \lim_n H_\varphi(M_n|N_n)$.

Proof. If $\psi \le \varphi$ then $S(\psi|_{M_n}, \varphi|_{M_n}) \nearrow S(\psi, \varphi)$ by Corollary 2.3.5, and a similar convergence holds for N_n and N. This implies that

$$H_\varphi(M|N) \le \liminf_n H_\varphi(M_n|N_n).$$

Since $E_{N_{n+1}} \circ E_{M_n} = E_{N_n}$ for all n, we have

$$E_{N_{n+k+1}} \circ E_{M_n} = E_{N_{n+k+1}} \circ E_{M_{n+k}} \circ E_{M_n} = E_{N_{n+k}} \circ E_{M_n},$$

so by induction $E_{N_{n+k}} \circ E_{M_n} = E_{N_n}$ for all $k \in \mathbb{N}$. Next note that $\{E_{N_n}\}_n$ converges in the pointwise strong operator topology to a φ-preserving conditional expectation $E_N \colon M \to N$. It follows that for any n

$$\begin{array}{ccc} M_n & \subset & M \\ \cup & & \cup \\ N_n & \subset & N \end{array}$$

is a commuting square. In other words, $E_N|_{M_n} = E_{N_n}|_{M_n}$. Hence for any positive linear functional ψ we have

$$S(\psi, \psi \circ E_N) \geq S(\psi|_{M_n}, \psi \circ E_N|_{M_n}) = S(\psi|_{M_n}, \psi \circ E_{N_n}|_{M_n}).$$

By Theorem 10.1.2(iv) it follows that $H_\varphi(M|N) \geq H_\varphi(M_n|N_n)$, which completes the proof. $\square$

From now onwards we shall only consider tracial states, and our next goal is to compute the relative entropy $H_\tau(M|N)$ when $N \subset M$ are finite dimensional (and τ is a trace on M).

Let $Z(M)$ and $Z(N)$ be the centers of M and N, respectively,

$$M = \bigoplus_{l=1}^{m} M_l, \quad N = \bigoplus_{k=1}^{n} N_k,$$

where $M_l \cong \mathrm{Mat}_{m_l}(\mathbb{C})$ and $N_k \cong \mathrm{Mat}_{n_k}(\mathbb{C})$. Let w_l and z_k be the central projections in M and N such that $M_l = M w_l$ and $N_k = N z_k$. Let a_{kl} be the multiplicity of $N_k w_l$ in M_l. Thus if t_l denotes the trace of a minimal projection in M_l and s_k that of a minimal projection in N_k,

$$\tau(w_l) = m_l t_l = \sum_k n_k a_{kl} t_l \quad \text{and} \quad \tau(z_k) = n_k s_k = \sum_l n_k a_{kl} t_l.$$

Theorem 10.1.4. *With the above notation the relative entropy $H_\tau(M|N)$ is*

$$(2H_\tau(M) - H_\tau(Z(M))) - (2H_\tau(N) - H_\tau(Z(N))) + \sum_{k,l} n_k a_{kl} t_l \log c_{kl},$$

where $c_{kl} = \min\left\{\dfrac{n_k}{a_{kl}}, 1\right\}$.

In particular, if M is abelian we have

$$H_\tau(M|N) = H_\tau(M) - H_\tau(N),$$

which we already know, and if $M = \mathrm{Mat}_m(\mathbb{C})$ and $N = \mathrm{Mat}_n(\mathbb{C})$ then

$$H_\tau(M|N) = 2\log m - 2\log n + \log\min\left\{\frac{n^2}{m}, 1\right\} = \min\left\{\log m, 2\log\frac{m}{n}\right\}.$$

A straightforward computation of the formula in the theorem shows that it is equivalent to the formula

$$\begin{aligned} H_\tau(M|N) = &-\sum_l m_l t_l \log t_l + \sum_l m_l t_l \log m_l + \sum_k n_k s_k \log s_k \\ &-\sum_k n_k s_k \log n_k + \sum_{k,l} n_k a_{kl} t_l \log c_{kl}. \end{aligned} \tag{10.2}$$

Note that if $\tau = \sum_i \lambda_i \varphi_i$ is a decomposition of τ into a convex combination of states, then

$$\sum_i S(\lambda_i\varphi_i, \tau) = S(\tau) - \sum_i \lambda_i S(\varphi_i) - \sum_i \eta(\lambda_i),$$

and we have a similar expression for $\sum_i S(\lambda_i\varphi_i|_N, \tau|_N)$. It follows that

$$H_\tau(M|N) = \sup\left\{H_\tau(M) - H_\tau(N) + \sum_i \lambda_i(S(\varphi_i|_N) - S(\varphi_i))\right\}, \tag{10.3}$$

where the supremum is taken over all decompositions $\tau = \sum_i \lambda_i\varphi_i$ of τ into a convex combination of states. This is the expression we shall use in the proof of the theorem. Note also that instead of finite convex decompositions we can use integral decompositions of τ.

Proof of Theorem 10.1.4. We first show that the left side is majorized by the right side in equation (10.2). Consider an integral decomposition $\tau = \int_X \varphi_x d\mu(x)$. By (10.3) we have to estimate

$$H_\tau(M) - H_\tau(N) + \int_X (S(\varphi_x|_N) - S(\varphi_x))d\mu(x). \tag{10.4}$$

By Theorem 2.2.2(ii) we have

$$S(\varphi_x|_N) = \sum_k S(\varphi_x|_{Nz_k}) \leq \sum_{k,l} S(\varphi_x|_{Nz_kw_l}).$$

For fixed k and l consider the state $\psi = \varphi_x(z_kw_l)^{-1}\varphi_x$ on $z_kMw_lz_k$. By Theorem 2.2.2(i) we have

$$S(\psi|_{Nz_kw_l}) - S(\psi) \leq S(\psi|_{Nz_kw_l}) \leq \log n_k.$$

On the other hand, since $z_kMw_lz_k \cong Nz_kw_l \otimes \mathrm{Mat}_{a_{kl}}(\mathbb{C})$, by Theorem 2.2.2(vi) we have

$$S(\psi|_{Nz_kw_l}) - S(\psi) \leq \log a_{kl}.$$

Using $S(\lambda\psi) = \lambda S(\psi) + \eta(\lambda)$ with $\lambda = \varphi_x(z_k w_l)$, we therefore get

$$S(\varphi_x|_{Nz_k w_l}) - S(\varphi_x|_{z_k M w_l z_k}) \le \varphi_x(z_k w_l) \log \min\{n_k, a_{kl}\}.$$

Finally, from Lemma 2.2.4 applied to the state $\varphi_x(w_l)^{-1}\varphi_x$ on Mw_l, we obtain

$$\sum_k S(\varphi_x|_{z_k M w_l z_k}) - S(\varphi_x|_{Mw_l}) \le \varphi_x(w_l) \sum_k \eta\left(\frac{\varphi_x(z_k w_l)}{\varphi_x(w_l)}\right).$$

Thus, since $\sum_l S(\varphi_x|_{Mw_l}) = S(\varphi_x)$, the integral in (10.4) is estimated from above by

$$\sum_{k,l} \int_X \varphi_x(z_k w_l) \log(\min\{n_k, a_{kl}\}) d\mu(x) + \sum_{k,l} \int_X \varphi_x(w_l)\eta\left(\frac{\varphi_x(z_k w_l)}{\varphi_x(w_l)}\right) d\mu(x).$$

Since $\int_X \varphi_x d\mu(x) = \tau$ and $\tau(z_k w_l) = n_k a_{kl} t_l$, the first summand in the above expression is equal to

$$\sum_{k,l} n_k a_{kl} t_l \log \min\{n_k, a_{kl}\}.$$

On the other hand, using concavity of η and that $\tau(w_l) = m_l t_l$, we can estimate the second summand by

$$\begin{aligned} m_l t_l \int_X \frac{\varphi_x(w_l)}{m_l t_l} \eta\left(\frac{\varphi_x(z_k w_l)}{\varphi_x(w_l)}\right) d\mu(x) &\le m_l t_l \eta\left(\int_X \frac{\varphi_x(w_l)}{m_l t_l} \frac{\varphi_x(z_k w_l)}{\varphi_x(w_l)} d\mu(x)\right) \\ &= m_l t_l \eta\left(\frac{n_k a_{kl} t_l}{m_l t_l}\right) \\ &= -t_l n_k a_{kl} \log \frac{n_k a_{kl}}{m_l}. \end{aligned}$$

Thus we estimate (10.4) by

$$\begin{aligned} &H_\tau(M) - H_\tau(N) + \sum_{k,l} n_k a_{kl} t_l \log(\min\{a_{kl}, n_k\}) - \sum_{k,l} n_k a_{kl} t_l \log \frac{n_k a_{kl}}{m_l} \\ &\quad = H_\tau(M) - H_\tau(N) + \sum_{k,l} n_k a_{kl} t_l \log c_{kl} + \sum_{k,l} n_k a_{kl} t_l \log a_{kl} \\ &\qquad - \sum_{k,l} n_k a_{kl} t_l (\log a_{kl} + \log n_k - \log m_l) \\ &\quad = H_\tau(M) - H_\tau(N) + \sum_{k,l} n_k a_{kl} t_l \log c_{kl} \\ &\qquad - \sum_k n_k s_k \log n_k + \sum_l m_l t_l \log m_l, \end{aligned}$$

where we used that $s_k = \sum_l a_{kl} t_l$ and $m_l = \sum_k n_k a_{kl}$. Since

$$H_\tau(M) = -\sum_l m_l t_l \log t_l \quad \text{and} \quad H_\tau(N) = \sum_k n_k s_k \log s_k,$$

we have shown that the left side of equation (10.2) is majorized by the right side.

In order to prove the opposite inequality choose a pure state φ_l of Mw_l such that

$$\begin{aligned}
&\varphi_l(z_k w_l) = \frac{n_k a_{kl}}{m_l},\\
&\varphi_l|_{Nz_k w_l} \text{ is a trace if } n_k < a_{kl},\\
&\varphi_l|_{(N'\bigcap M)z_k w_l} \text{ is a trace if } n_k \geq a_{kl}.
\end{aligned}$$

Explicitly φ_l can be constructed as follows. Identify $\ell_2^{m_l}$ with $\oplus_k(\ell_2^{n_k} \otimes \ell_2^{a_{kl}})$, and let $\{e_i\}_{i\in\mathbb{N}}$ denote the standard basis in ℓ_2. Put

$$\xi_l = \bigoplus_k \left(\frac{n_k a_{kl}}{\min\{n_k, a_{kl}\} m_l}\right)^{1/2} \sum_{i=1}^{\min\{n_k,a_{kl}\}} e_i \otimes e_i.$$

Then let $\varphi_l = \omega_{\xi_l}$ be the vector state defined by ξ_l.

Let K_l denote the subalgebra of $\oplus_k z_k M w_l z_k \subset M w_l$ consisting of elements $\oplus_k x_k$ such that $x_k \in N z_k w_l$ if $n_k < a_{kl}$, and $x_k \in (N' \cap M) z_k w_l$ if $n_k \geq a_{kl}$. Then φ_l is a pure state on Mw_l such that its restriction to K_l coincides with the restriction of the unique tracial state τ_l on Mw_l to K_l. Furthermore, the map $E_l\colon Mw_l \to K_l$ defined by

$$E_l(x) = \int_{U(K_l' \cap M w_l)} (\operatorname{Ad} u)(x) d\mu_l(u),$$

where μ_l is the normalized Haar measure on the unitary group $U(K_l' \cap M w_l)$ of $K_l' \cap M w_l$, is a τ_l-preserving conditional expectation. It follows that

$$\tau = \sum_l m_l t_l \int_{U(K_l' \cap M w_l)} \varphi_l \circ \operatorname{Ad} u \, d\mu_l(u).$$

Thus by (10.3)

$$H_\tau(M|N) \geq H_\tau(M) - H_\tau(N) + \sum_{k,l} m_l t_l S(\varphi_l|_{N z_k w_l}). \tag{10.5}$$

The restricition of φ_l to $z_k M w_l z_k$ is a scalar multiple of a pure state. Hence by Lemma 2.2.3(i)

$$S\left(\frac{m_l}{n_k a_{kl}} \varphi_l|_{N z_k w_l}\right) = S\left(\frac{m_l}{n_k a_{kl}} \varphi_l|_{(N' \cap M) z_k w_l}\right) = \log \min\{n_k, a_{kl}\},$$

whence

$$S(\varphi_l|_{Nz_k w_l}) = \frac{n_k a_{kl}}{m_l}(\log c_{kl} + \log a_{kl}) - \frac{n_k a_{kl}}{m_l}\log\frac{n_k a_{kl}}{m_l}.$$

Thus the right side of (10.5) equals

$$H_\tau(M) - H_\tau(N) + \sum_{k,l} n_k a_{kl} t_l (\log c_{kl} + \log a_{kl}) - \sum_{k,l} n_k a_{kl} t_l \log\frac{n_k a_{kl}}{m_l},$$

which is exactly what we need. □

10.2 Index of Subfactors

Our goal in this section is to show that relative entropy is related to index of subfactors. We shall briefly recall some basic facts of the latter theory referring the reader to [214, Chapter XIX] for more details.

Let M be a II_1-factor with trace τ. If M acts on a Hilbert space H then H can be embedded into $L^2(M)\otimes K$ for a sufficiently large Hilbert space K. Let $p \in (M \otimes \mathbb{C}1)' = M'\bar{\otimes}B(K)$ be the projection onto H. Then the **dimension of H relative to M** is defined by

$$\dim_M(H) = (\tau' \otimes \mathrm{Tr})(p),$$

where τ' is the unique tracial state on $M' \subset B(L^2(M))$ and Tr is the canonical trace on $B(K)$. The dimension can take any value in $(0, +\infty]$, and if H is countably generated in the sense that it is a direct sum of at most countably many cyclic subspaces for M, then $\dim_M(H)$ is a complete invariant of the unitary equivalence class of the representation $M \to B(H)$.

The dimension $\dim_M(H)$ is finite if and only if the commutant M' of M in $B(H)$ is a II_1-factor. In the latter case

$$\dim_M(H) = \frac{\tau([M'\xi])}{\tau'([M\xi])}, \tag{10.6}$$

where $\xi \in H$ is any nonzero vector, τ' is the unique tracial state on $M' \subset B(H)$, and $[M\xi]$ denotes the projection onto the subspace $\overline{M\xi}$.

If N is a subfactor of M, the **index** of N in M is

$$[M\colon N] = \dim_N(L^2(M)).$$

If M acts on H then $\dim_N(H) = [M\colon N]\dim_M(H)$. It follows that if $N \subset M \subset L$ then $[L\colon N] = [L\colon M][M\colon N]$. It follows also that if $[M\colon N] < \infty$ and N is represented on H then this representation extends to a representation of M on H. Indeed, decomposing H into a direct sum of cyclic subspaces we may assume that $\dim_N(H) \le 1$. Choose a representation of M on a Hilbert space H' such that $\dim_M(H') = [M\colon N]^{-1}\dim_N(H)$. Then $\dim_N(H') = \dim_N(H)$.

Hence the representations of N on H and H' are equivalent. Since the representation of N on H' extends to a representation of M, the same is true for H.

Using Proposition 10.1.3 and Theorem 10.1.4 it is easy to check that if R is the hyperfinite II_1-factor then

$$H_\tau(R \otimes \mathrm{Mat}_n(\mathbb{C}) | R \otimes 1) = 2 \log n.$$

On the other hand, $[R \otimes \mathrm{Mat}_n(\mathbb{C}) : R \otimes 1] = n^2$. So we see that in this case the relative entropy is the logarithm of the index. This is a particular case of a far more general result. For simplicity of the presentation we stick to the **irreducible** case, that is, when $N' \cap M = \mathbb{C}1$.

Theorem 10.2.1. *Let $N \subset M$ be II_1-factors such that $[M : N] < \infty$ and $N' \cap M = \mathbb{C}1$. Then*

$$H_\tau(M|N) = \log[M : N].$$

To prove this theorem we need two properties of finite index subfactors. The first one is the **Pimsner-Popa inequality** stating that

$$E_N(x) \geq [M : N]^{-1} x \quad \text{for any} \quad x \in M_+,$$

where $E_N : M \to N$ is the trace preserving conditional expectation, see [214, Theorem XIX.4.14].

Proposition 10.2.2. *If $N \subset M$ is a finite index subfactor, then*

$$H_\tau(M|N) \leq \log[M : N].$$

Proof. Let $\lambda = [M : N]^{-1}$. Then by the Pimsner-Popa inequality $\varphi \circ E_N \geq \lambda \varphi$ for any positive linear functional φ on M. Since the relative entropy S is decreasing in the second variable, Theorem 2.3.1(iii), we then have

$$S(\varphi, \varphi \circ E_N) \leq S(\varphi, \lambda \varphi) = -\varphi(1) \log \lambda.$$

By Theorem 10.1.2(iv) it follows that $H_\tau(M|N) \leq -\log \lambda = \log[M : N]$. □

The second property of subfactors which we need, is existence of special projections.

Lemma 10.2.3. *If $N \subset M$ is a finite index subfactor then there exists a projection $q \in M$ such that $E_N(q) = [M : N]^{-1} 1$. Moreover, any other projection with this property is of the form vqv^*, where v is a unitary in N.*

Proof. See [214, Theorem XIX.4.12]. The existence will also essentially be shown in Lemma 10.4.1 below. □

Proof of Theorem 10.2.1. Let $\lambda = [M : N]^{-1}$ and q be the projection from Lemma 10.2.3. The idea of the proof is to consider the state $\varphi = \lambda^{-1} \tau(\cdot\, q)$

and the decomposition $\tau = \int_{U(N)} \varphi \circ \operatorname{Ad} u \, du$, which formally holds because of $E_{N'\cap M}(q) = \lambda 1$. For II_1-factors the above decomposition does not make sense, but it is still possible to get an approximate version of it.

Let $x = q - \lambda 1$. Then $\tau(x) = 0$. Let

$$K_x = \overline{\operatorname{conv}\{vxv^*\colon v \in U(N)\}},$$

where the closure is in the weak operator topology. Then K_x is a convex w-compact set in M, and $\tau(y) = 0$ for all $y \in K_x$. The weak operator closure of a bounded convex set coincides with the closure with respect to the L^2-norm. Thus there exists a unique $y_0 \in K_x$ such that

$$\|y_0\|_2 = \inf\{\|y\|_2\colon y \in K_x\}.$$

But $vy_0v^* \in K_x$ for all $v \in U(N)$, and $\|vy_0v^*\|_2 = \|y_0\|_2$, so that $vy_0v^* = y_0$, i.e., $y_0 \in N' \cap M$. Since $N' \cap M = \mathbb{C}1$, it follows that y_0 is a scalar, so $y_0 = 0$. Hence for any fixed $\varepsilon > 0$ there are unitaries $v_1, \dots, v_n \in N$ such that

$$\left\| \frac{1}{n} \sum_{i=1}^{n} v_i q v_i^* - \lambda 1 \right\|_2 < \varepsilon^2 \lambda.$$

Let $y = (\lambda n)^{-1} \sum_i v_i q v_i^*$. Then $0 \le y \le \lambda^{-1} 1$ and

$$\|y - 1\|_2 \le \varepsilon^2.$$

Let p be the spectral projection of y corresponding to the interval $[0, 1+\varepsilon]$. Set

$$x_i = ((1+\varepsilon)\lambda n)^{-1} v_i q v_i^* \wedge p.$$

Then $\sum_i x_i \le ((1+\varepsilon)\lambda n)^{-1} \sum_i p v_i q v_i^* p = (1+\varepsilon)^{-1} y p \le 1$. Hence

$$H_\tau(M|N) \ge \sum_i \tau(\eta(E_N(x_i)) - \eta(x_i)) = ((1+\varepsilon)\lambda n)^{-1} \sum_i \tau(\eta(E_N(v_i q v_i^* \wedge p))).$$

Using operator monotonicity of log as in the proof of Theorem 2.2.2(ii), we get $\tau(\eta(e+f)) \le \tau(\eta(e)) + \tau(\eta(f))$ for any positive e and f. Hence

$$\begin{aligned} \tau(\eta(E_N(v_i q v_i^* \wedge p))) &\ge \tau(\eta(E_N(v_i q v_i^*))) - \tau(\eta(E_N(v_i q v_i^* - v_i q v_i^* \wedge p))) \\ &\ge \eta(\lambda) - \eta(\tau(v_i q v_i^* - v_i q v_i^* \wedge p)), \end{aligned}$$

where in the second inequality we used that $E_N(q) = \lambda 1$ and that $\eta \circ \tau \ge \tau \circ \eta$ by concavity of η. Thus

$$H_\tau(M|N) \ge -(1+\varepsilon)^{-1} \log \lambda - ((1+\varepsilon)\lambda n)^{-1} \sum_i \eta(\tau(v_i q v_i^* - v_i q v_i^* \wedge p)). \quad (10.7)$$

Now recall that $\tau(e \vee f) - \tau(f) = \tau(e) - \tau(e \wedge f)$ for any projections e and f. Hence

$$\tau(v_i q v_i^* - v_i q v_i^* \wedge p) \le 1 - \tau(p).$$

Since $y(1-p) \ge (1+\varepsilon)(1-p)$, we have

$$\|y-1\|_2^2 \ge \tau((y-1)^2(1-p)) \ge \varepsilon^2\tau(1-p),$$

so that $1-\tau(p) \le \varepsilon^{-2}\|y-1\|_2^2 \le \varepsilon^2$. Thus

$$\tau(v_i q v_i^* - v_i q v_i^* \wedge p) \le \varepsilon^2.$$

Since η is increasing on $[0, \mathrm{e}^{-1}]$, for ε small enough we obtain

$$\eta(\tau(v_i q v_i^* - v_i q v_i^* \wedge p)) \le \eta(\varepsilon^2).$$

From (10.7) we then get

$$H_\tau(M|N) \ge -(1+\varepsilon)^{-1}\log\lambda - ((1+\varepsilon)\lambda)^{-1}\eta(\varepsilon^2).$$

Letting $\varepsilon \to 0$ we conclude that $H_\tau(M|N) \ge -\log\lambda = \log[M\colon N]$. The opposite inequality follows from Proposition 10.2.2. □

Remark 10.2.4. What we actually used in the proof is not the irreducibility $N' \cap M = \mathbb{C}1$, but that $E_{N'\cap M}(q)$ is a scalar. Subfactors satisfying the latter condition are called **extremal**, and in fact the equality $H_\tau(M|N) = \log[M\colon N]$ holds if and only if $N \subset M$ is extremal [162, Corollary 4.5].

Note that if $[M\colon N] < 4$ then automatically $N' \cap M = \mathbb{C}1$, see [214, Corollary XIX.2.10]. Recall also that the possible values of the index is the set $\{4\cos^2\pi/n \,|\, n \ge 3\} \cup [4, +\infty]$, see [214, Theorem XIX.2.22].

Finally note that with some extra work the above theorem can be extended to arbitrary subfactors [162, Theorem 4.5]. In that case $H_\tau(M|N) = \infty$ whenever $N' \cap M$ has a diffuse part. If $N' \cap M$ is atomic, and $\{f_k\}_k$ are minimal projections in $N' \cap M$ with sum 1, then

$$H_\tau(M|N) = 2\sum_k \eta(\tau(f_k)) + \sum_k \tau(f_k)\log[f_k M f_k\colon f_k N f_k].$$

◆

10.3 Generators and Relative Entropy

In this section we shall prove an analogue of the formula $h(T) = H(\xi|\xi^-)$, see (1.3) and Theorem 1.1.4.

Before we state the main result we introduce some notation. Let (M, τ, α) be a W*-dynamical system, where τ is a faithful normal trace. Let $\{A_n\}_{n=1}^\infty$ be an increasing sequence of finite dimensional C*-subalgebras of M.

We say that $\{A_n\}_{n=1}^\infty$ is a **generating sequence** for α if

(i) $\alpha(A_n) \subset A_{n+1}$, $n \in \mathbb{N}$;

(ii) the von Neumann subalgebra generated by $\alpha^k(A_n)$, $k \in \mathbb{Z}$, $n \in \mathbb{N}$, coincides with M;

(iii) $h_\tau(\alpha) = \lim\limits_{n\to\infty} \dfrac{1}{n} H_\tau(A_n)$.

We say $\{A_n\}_n$ satisfies the **commuting square condition** if

$$\begin{array}{ccc} A_{n+1} & \subset & A_{n+2} \\ \cup & & \cup \\ \alpha(A_n) & \subset & \alpha(A_{n+1}) \end{array}$$

is a commuting square with respect to the τ-preserving conditional expectations for every $n \in \mathbb{N}$. Remark that if M is abelian and $A_n = \vee_{k=0}^{n-1}\alpha^k(A_1)$, then this condition is satisfied if and only if A_1 corresponds to the generating partition of a Markov process.

Write A_n in the form

$$A_n = \bigoplus_{l\in K_n} M_k^n,$$

where $M_k^n \cong \mathrm{Mat}_{m_k^n}(\mathbb{C})$. Let $(a_{kl}^n)_{k\in K_{n-1}, l\in K_n}$ be the inclusion matrix for $\alpha(A_{n-1}) \subset A_n$. Denote by $Z(A_n)$ the center of A_n.

Theorem 10.3.1. *Let (M,τ,α) be a W^*-dynamical system, where τ is a faithful normal trace and M is of type II_1. Suppose $\{A_n\}_{n=1}^\infty$ is a generating sequence for α satisfying the commuting square condition. Assume also that*

$$h_\tau(\alpha) < \infty \quad \text{and} \quad \sup_{n,k,l} \frac{a_{kl}^n}{m_k^{n-1}} < \infty.$$

Then the limit $\lim\limits_{n\to\infty} \dfrac{1}{n} H_\tau(Z(A_n))$ exists, and with $R = (\cup_n A_n)''$ we have

$$h_\tau(\alpha) = \frac{1}{2} H_\tau(R|\alpha(R)) + \frac{1}{2}\lim_{n\to\infty}\frac{1}{n} H_\tau(Z(A_n)).$$

If we combine this theorem with Theorem 10.2.1, we immediately obtain the following corollary.

Corollary 10.3.2. *If we in the above theorem add the assumptions that $\alpha(R)' \cap R = \mathbb{C}1$ and $\lim_n \frac{1}{n} H_\tau(Z(A_n)) = 0$, then $h_\tau(\alpha) = \frac{1}{2}\log[R\colon \alpha(R)]$.* $\square$

To prove the theorem we need some preparation.

Lemma 10.3.3. *Fix $r > 0$. Let f_n be the central projection in A_n such that $A_n f_n = \oplus_{m_k^n \le r} M_k^n$. Then $\{f_n\}_n$ is a decreasing sequence converging strongly to zero.*

Proof. First note that $R = (\cup_n A_n)''$ is of type II_1. Indeed, let z be the central projection in R corresponding to the sum of the type I_k components of R with $k \le r$, and assume $z \ne 0$. Since $\alpha(R) \subset R$, and type II and type I_k algebras can not be embedded into a type I_l algebra with $l < k$, we have $z \le \alpha(z)$. As τ is α-invariant and faithful, it follows that $z = \alpha(z)$. The maximal number of mutually equivalent orthogonal projections in Rz is not larger than r. By the definition of a generating sequence, the sequence $\{\alpha^{-n}(R)\}_{n=1}^{\infty}$ is increasing with union dense in M. Since $z = \alpha^{-n}(z)$ is central in each $\alpha^{-n}(R)$, the projection z is central in M, and the maximal number of mutually equivalent orthogonal projections in Mz is not larger than r. In other words, Mz is a sum of type I_k algebras with $k \le r$. But this contradicts our assumption on M.

Since a type I_k algebra can not be embedded into a type I_l algebra with $l < k$, it is clear that $\{f_n\}_n$ is a decreasing sequence. Let f be its strong limit. Since f_n is central in A_n, the projection f is central in R. Then $Rf = (\cup_n A_n f)''$, which by the same argument as above contradicts the fact that R is of type II_1 unless $f = 0$. $\square$

Denote by t_k^n the trace of a minimal projection in $M_k^n \subset A_n$. Since $H_\tau(A_{n-1}) = H_\tau(\alpha(A_{n-1}))$ and $H_\tau(Z(A_{n-1})) = H_\tau(Z(\alpha(A_{n-1})))$, by Theorem 10.1.4 we have

$$H_\tau(A_n|\alpha(A_{n-1})) = (2H_\tau(A_n) - H_\tau(Z(A_n))) - (2H_\tau(A_{n-1}) - H_\tau(Z(A_{n-1}))) + \sum_{k,l} m_k^{n-1} a_{kl}^n t_l^n \log c_{kl}^n, \tag{10.8}$$

where $c_{kl}^n = \min\{m_k^{n-1}/a_{kl}^n, 1\}$.

Lemma 10.3.4. *Suppose*

$$C = \sup_N \frac{1}{N} \sum_{n=2}^{N} \sum_{k,l} m_k^{n-1} a_{kl}^n t_l^n \log a_{kl}^n < \infty.$$

Then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=2}^{N} \sum_{k,l} m_k^{n-1} a_{kl}^n t_l^n \log c_{kl}^n = 0.$$

Proof. Since by assumption of the theorem there is $c > 0$ such that $c \le c_{kl}^n \le 1$ for all k, l and n, it suffices to prove that

$$\lim_N \frac{1}{N} \sum_{n=2}^{N} \sum_{k,l: a_{kl}^n > m_k^{n-1}} m_k^{n-1} a_{kl}^n t_l^n = 0.$$

Fix $r > 0$. Let f_n be as in Lemma 10.3.3. Then

$$\frac{1}{N} \sum_{n=2}^{N} \sum_{k,l: m_k^{n-1} \le r} m_k^{n-1} a_{kl}^n t_l^n = \frac{1}{N} \sum_{n=2}^{N} \sum_{k: m_k^{n-1} \le r} m_k^{n-1} t_k^{n-1} = \frac{1}{N} \sum_{n=2}^{N} \tau(f_{n-1}),$$

and the latter expression tends to zero by Lemma 10.3.3. On the other hand,

$$\frac{1}{N}\sum_{n=2}^{N}\sum_{k,l:a_{kl}^n>m_k^{n-1}>r} m_k^{n-1}a_{kl}^n t_l^n \le (\log r)^{-1}\frac{1}{N}\sum_{n=2}^{N}\sum_{k,l} m_k^{n-1}a_{kl}^n t_l^n \log a_{kl}^n$$
$$\le C(\log r)^{-1}.$$

Since we can take an arbitrary large r, this proves the lemma. □

Proof of Theorem 10.3.1. Let us check that the assumption of the previous lemma is satisfied. Since $\sum_k m_k^{n-1}a_{kl}^n = m_l^n$ and $\sum_l a_{kl}^n t_l^n = t_l^{n-1}$, we have

$$\begin{aligned}\sum_{k,l} m_k^{n-1}a_{kl}^n t_l^n \log a_{kl}^n &\le \sum_{k,l} m_k^{n-1}a_{kl}^n t_l^n \log \frac{m_l^n}{m_k^{n-1}}\\ &= \sum_l m_l^n t_l^n \log m_l^n - \sum_k m_k^{n-1}t_k^{n-1}\log m_k^{n-1}\\ &= H_\tau(A_n) - H_\tau(Z(A_n)) - H_\tau(A_{n-1}) + H_\tau(Z(A_{n-1})).\end{aligned}$$

Hence, with $A_0 = \mathbb{C}1$,

$$\frac{1}{N}\sum_{n=1}^{N}\sum_{k,l} m_k^{n-1}a_{kl}^n t_l^n \log a_{kl}^n \le \frac{1}{N}H_\tau(A_N) - \frac{1}{N}H_\tau(Z(A_N)) \le \frac{1}{N}H_\tau(A_N).$$

Since the sequence $\{\frac{1}{N}H_\tau(A_N)\}_N$ is bounded by assumption of the theorem, it follows that the assumption of Lemma 10.3.4 is satisfied. Hence if we put $C_n = \sum_{k,l} m_k^{n-1}a_{kl}^n t_l^n \log c_{kl}^n$, we get

$$\frac{1}{N}\sum_{n=1}^{N} C_n \to 0.$$

By equation (10.8) we have

$$\frac{1}{N}\sum_{n=1}^{N} H_\tau(A_n|\alpha(A_{n-1})) = \frac{2}{N}H_\tau(A_N) - \frac{1}{N}H_\tau(Z(A_N)) + \frac{1}{N}\sum_{n=1}^{N} C_n.$$

Since $\{A_n\}_n$ satisfies the commuting square condition, by Proposition 10.1.3

$$H_\tau(A_n|\alpha(A_{n-1})) \to H_\tau(R|\alpha(R)).$$

Since $\frac{1}{N}H_\tau(A_N) \to h_\tau(\alpha)$ by assumption, it follows that $\lim_N \frac{1}{N}H_\tau(Z(A_N))$ exists and

$$H_\tau(R|\alpha(R)) = 2h_\tau(\alpha) - \lim_N \frac{1}{N}H_\tau(Z(A_N)).$$

Thus

$$h_\tau(\alpha) = \frac{1}{2}H_\tau(R|\alpha(R)) + \frac{1}{2}\lim_N \frac{1}{N}H_\tau(Z(A_N)).$$

□

10.4 The Canonical Shift

In the present section we shall apply the results of the preceding sections to an interesting automorphism, called the canonical shift, defined on the tower of relative commutants associated with an inclusion $N \subset M$ of II_1-factors of finite index.

In order to introduce it we need a few more facts from index theory, again see [214, Chapter XIX] for more details. Let τ be the trace on M, $\xi \in L^2(M)$ the cyclic vector defining τ. Denote by e_N the projection onto $\overline{N\xi}$, and by $M_1 = \langle M, e_N \rangle$ the von Neumann subalgebra of $B(L^2(M))$ generated by M and e_N. Since $e_N x e_N = E_N(x)e_N = e_N E_N(x)$ for $x \in M$, we have $\{e_N\}' \cap M = N$. In other words, N' is generated by M' and e_N. So if we denote by J the modular conjugation defined by τ, $Jx\xi = x^*\xi$, so that $JMJ = M'$ and $Je_N J = e_N$, then $JM_1J = N'$. It follows that M_1 is a II_1-factor and

$$\dim_{M_1}(L^2(M)) = \dim_{N'}(L^2(M)) = \frac{1}{\dim_N(L^2(M))} = \frac{1}{[M\colon N]},$$

where the second equality follows from (10.6). Hence $[M_1\colon M] = [M\colon N]$. Furthermore, by (10.6) the unique tracial state τ' on N' has the property $\tau'(e_N) = [M\colon N]^{-1}$. The unique tracial state on M_1, which we again denote by τ, is such that $\tau(x) = \tau'(Jx^*J)$. Hence $\tau(e_N) = [M\colon N]^{-1}$. Since N is a factor, $\tau(ab) = \tau(a)\tau(b)$ for $a \in N$ and $b \in N' \cap M_1$. Hence, for $x \in M$,

$$\tau(xE_M(e_N)) = \tau(xe_N) = \tau(E_N(x)e_N) = \tau(E_N(x))\tau(e_N) = [M\colon N]^{-1}\tau(x),$$

so we also have $E_M(e_N) = [M\colon N]^{-1}1$, where $E_M\colon M_1 \to M$ is the trace preserving conditional expectation.

Thus starting with a finite index inclusion $N \subset M$ we construct a II_1-factor M_1 and a projection $e_N \in M_1$ such that $[M_1\colon M] = [M\colon N]$, $M_1 = \langle M, e_N \rangle$, $N = \{e_N\}' \cap M$ and $E_M(e_N) = [M\colon N]^{-1}1$. This is called the **basic construction**. Iterating it we get the **Jones tower** of II_1-factors

$$M_{-1} = N \subset M_0 = M \subset M_1 \subset \ldots \subset M_k \subset \ldots$$

and projections $e_k \in M_{k+1}$ such that $M_{k+1} = \langle M_k, e_k \rangle$, $k \geq 0$, $e_0 = e_N$.

There is also a downward construction. So there exists a projection e_{-1} and a subfactor $N_{-1} \subset N$ such that

$$N_{-1} \subset N \subset \langle N, e_{-1} \rangle = M$$

is the basic construction. Since $E_N(e_{-1}) = [M\colon N]^{-1}1$, by Lemma 10.2.3 the projection e_{-1} is defined up to conjugation by a unitary in N. Since $N_{-1} = \{e_{-1}\}' \cap N$, the subfactor N_{-1} is also defined up to conjugation by a unitary in N. Denoting N_{-1} by M_{-2} and iterating the downward construction we get a **tunnel**

$$\ldots \subset M_{-k} \subset \ldots \subset M_{-1} \subset M_0$$

and projections $e_{-k} \in M_{-k+1}$, $k \geq 1$.

Using that $e_k x e_k = e_k E_{M_{k-1}}(x) = E_{M_{k-1}}(x) e_k$ for $x \in M_k$ and $E_{M_k}(e_k) = [M : N]^{-1}1$ one then checks that the projections e_k, $k \in \mathbb{Z}$, satisfy the relations

$$e_k e_{k\pm 1} e_k = [M\colon N]^{-1} e_k, \quad e_k e_j = e_j e_k \;\text{ if }\; |k-j| \geq 2.$$

Since $[M_{k+1}\colon M_k] = [M\colon N] < \infty$, by the discussion at the beginning of Sect. 10.2 we can inductively construct a representation of $\cup_k M_k$ on $L^2(M)$ extending the representation of M_1. This representation is not unique, but as soon as we fix a tunnel there exists a canonical one.

Lemma 10.4.1. *With the above notation there exists a unique extension of the representation of M_1 on $L^2(M)$ to a representation of $\cup_{k=1}^{\infty} M_k$ on $L^2(M)$ such that $e_k = J e_{-k} J$ for $k \geq 1$. In this representation $JM_k'J = M_{-k}$ for every $k \in \mathbb{Z}$.*

Proof. Since $M_{k+1} = \langle M_k, e_k \rangle$, the condition $e_k = J e_{-k} J$ completely determines the representation. We therefore only have to prove its existence. We shall do this by constructing the required representation of M_k by induction on k. For $k = 1$ we have a representation by definition, and $M_1 = JM_{-1}'J$. So assume the representation is well-defined for some k, and $M_j = JM_{-j}'J$ for $j = 1, \dots, k$, where we identify M_j with its image under the representation.

Since $[M_{k+1}\colon M_k] < \infty$, the identity map $M_k \to M_k \subset B(L^2(M))$ extends to a representation $\pi\colon M_{k+1} \to B(L^2(M))$. Let $f = J\pi(e_k)J$. Denote $[M\colon N]^{-1}$ by λ. We assert that $f \in M_{-k+1}$ and $E_{M_{-k}}(f) = \lambda 1$. Indeed, since $e_k \in M_{k-1}' \cap M_{k+1}$, we have $\pi(e_k) \in M_{k-1}'$, so that $f \in JM_{k-1}'J = M_{-k+1}$. Then

$$\begin{aligned} E_{M_{-k}}(f) e_{-k+1} &= e_{-k+1} f e_{-k+1} = J e_{k-1} J J \pi(e_k) J J e_{k-1} J \\ &= J\pi(e_{k-1} e_k e_{k-1}) J = \lambda J e_{k-1} J = \lambda e_{-k+1}, \end{aligned}$$

so that $E_{M_{-k}}(f) = \lambda 1$. But then by Lemma 10.2.3 there exists a unitary $u \in M_{-k}$ such that $e_{-k} = ufu^* = J(JuJ\pi(e_k)Ju^*J)J$. Since $JuJ \in JM_{-k}J = M_k'$, we thus see that $x \mapsto JuJ\pi(x)Ju^*J$ is a representation of M_{k+1} which coincides with π on M_k and maps e_k onto $Je_{-k}J$. Therefore it is the required representation of M_{k+1}.

Since $M_{-k-1} = \{e_{-k}\}' \cap M_{-k}$, we have

$$JM_{-k-1}'J = \langle JM_{-k}'J, Je_{-k}J \rangle = \langle M_k, e_k \rangle = M_{k+1},$$

which completes the proof of the induction step. □

Since $JM_{-n}'J = M_n$, it follows that $M_{-n} \subset M_0 \subset M_n$ is the basic construction for each $n \in \mathbb{N}$. Hence, more generally, $M_{k-n} \subset M_k \subset M_{k+n}$ is the basic construction for any $k \in \mathbb{Z}$ and $n \in \mathbb{N}$.

Consider now the AF-algebra A_∞ obtained by taking the inductive limit of the algebras $M_k' \cap M_l$ as $k \to -\infty$ and $l \to +\infty$ (note that $M_k' \cap M_l$ is finite dimensional by [214, Corollary XIX.2.9]). By the above lemma we get a **mirroring** γ_0, the involutive anti-automorphism of A_∞ defined by $\gamma_0(x) = Jx^*J$ for $x \in \cup_{k,l}(M'_k \cap M_l)$. It has the properties $\gamma_0(M'_k \cap M_l) = M'_{-l} \cap M_{-k}$ and $\gamma_0(e_k) = e_{-k}$.

On the other hand, we can consider

$$\ldots \subset M_k \subset M_{k+1} \subset \ldots$$

as the tower and the tunnel associated with the inclusion $M_0 \subset M_1$, and get an anti-automorphism γ_1 of A_∞ such that $\gamma_1(M'_k \cap M_l) = M'_{-l+2} \cap M_{-k+2}$ and $\gamma_1(e_k) = e_{-k+2}$. The **canonical shift** for the inclusion $N \subset M$ is the automorphism $\Gamma = \gamma_1 \circ \gamma_0$ of A_∞. It has the properties

$$\Gamma(M'_k \cap M_l) = M'_{k+2} \cap M_{l+2} \quad \text{and} \quad \Gamma(e_k) = e_{k+2}.$$

Since each algebra M_k has a unique trace, the C*-algebra A_∞ has a canonical tracial state, which we continue to denote by τ. Our next goal is to show that Γ is τ-preserving. By [214, Proposition XIX.4.19] in the finite depth case, which we shall consider below, the trace τ is the unique tracial state on A_∞. Hence in that case it is γ_0- and γ_1-invariant. However, this is not true in general. To prove that nevertheless the trace is Γ-invariant we need to compare the constructed representations of $\cup_k M_k$ on $L^2(M_0)$ and $L^2(M_1)$. To simplify the notation we consider $\cup_k M_k$ as a subalgebra of $B(L^2(M_1))$, and denote by π the representation of $\cup_k M_k$ on $L^2(M_0)$.

Lemma 10.4.2. *If we identify $L^2(M_0)$ with $e_1 L^2(M_1)$, then the representation π satisfies*

$$\pi(x) = \lambda^{-k} e_1 \ldots e_k x e_{k+1} e_k \ldots e_1 \quad \textit{for} \quad x \in M_k, \quad k \geq 1,$$

where $\lambda = [M : N]^{-1}$.

Proof. Using the relations $e_i e_{i-1} e_i = \lambda e_i$, $i = 1, \ldots, k$, we get

$$e_{k+1} e_k \ldots e_1 e_1 \ldots e_k e_{k+1} = \lambda^k e_{k+1}.$$

Since $e_{k+1} x = x e_{k+1}$ for $x \in M_k$, it follows that the expression in the formulation of the lemma defines a homomorphism. For $x \in M_0$ we have $\pi(x) = x e_1$. Since M_k is generated by M_0 and $e_0, \ldots, e_{k-1}$, all that remains to show is that the equality in the formulation holds for $x = e_0, \ldots, e_{k-1}$. In other words,

$$\pi(e_{k-1}) = e_{k+1} e_1 \quad \text{for} \quad k \geq 2, \quad \text{and} \quad \pi(e_0) = \lambda^{-1} e_1 e_0 e_2 e_1.$$

Denote by J_1 and J_0 the modular conjugations on $L^2(M_1)$ and $L^2(M_0)$, respectively. Since $J_0 = J_1|_{e_1 L^2(M_1)}$, for $k \geq 2$ we have

$$\pi(e_{k-1}) = J_0 \pi(e_{-k+1}) J_0 = J_1 e_{-k+1} e_1 J_1 = e_{k+1} e_1,$$

proving the first identity. To prove the second note that if ξ_1 is the cyclic trace vector in $L^2(M_1)$, then

$$
\begin{aligned}
(\lambda^{-1} e_1 e_0 e_2 e_1 \xi_1, \xi_1) &= \lambda^{-1}(e_2 \xi_1, e_0 \xi_1) = \lambda^{-1}(J_1 e_0 J_1 \xi_1, e_0 \xi_1) \\
&= \lambda^{-1}(e_0 \xi_1, e_0 \xi_1) = \lambda^{-1}\tau(e_0) = 1.
\end{aligned}
$$

Since $\lambda^{-1} e_1 e_0 e_2 e_1$ is a projection, we conclude that $\xi_1 = \lambda^{-1} e_1 e_0 e_2 e_1 \xi_1$. Since $e_1 e_0 e_2 e_1 \in M'_{-1}$, it follows that the projection $[M_{-1}\xi_1]$ is majorized by $\lambda^{-1} e_1 e_0 e_2 e_1$. On the other hand, since M_1 is generated by M_0 and e_0, we have $e_0 M_1 e_0 = e_0 M_{-1}$. Hence

$$
[M_{-1}\xi_1] = [M_{-1} e_1 e_2 e_0 e_1 \xi_1] = [e_1 e_2 e_0 M_{-1} \xi_1] = [e_1 e_2 e_0 M_1 e_0 \xi_1].
$$

Since $e_2 \xi_1 = J_1 e_0 J_1 \xi_1 = e_0 \xi_1$ and $e_2 \in M'_1$, the above projection equals

$$
[e_1 e_2 e_0 M_1 e_2 \xi_1] = [e_1 e_2 e_0 M_1 \xi_1] = [e_1 e_2 e_0 L^2(M_1)] \geq \lambda^{-1} e_1 e_0 e_2 e_1.
$$

By definition of $e_0 = e_N$ we get $\pi(e_0) = [M_{-1}\xi_1] = \lambda^{-1} e_1 e_0 e_2 e_1$, completing the proof of the lemma. □

Proposition 10.4.3. *The canonical shift Γ is trace preserving.*

Proof. We use the same notation as in the proof of the previous lemma. Since $x = \gamma_0(\gamma_1(\Gamma(x)))$, $\pi \circ \gamma_0 = J_0 \pi(\cdot)^* J_0$ and $J_0 = J_1|_{e_1 L^2(M_1)}$, for any $x \in A_\infty$ we have

$$
\pi(x) = J_1 \pi(\gamma_1(\Gamma(x^*))) J_1.
$$

If $x \in M'_k \cap M_l$ with $k \leq -1$ and $l \geq 1$, then $\gamma_1(\Gamma(x)) = \gamma_0(x) \in M'_{-l} \cap M_{-k}$. Using Lemma 10.4.2 and that $J_1 e_i J_1 = e_{-i+2}$, we conclude that the above expression equals

$$
\lambda^k J_1 e_1 \dots e_{-k} J_1 \Gamma(x) J_1 e_{-k+1} e_{-k} \dots e_1 J_1 = \lambda^k e_1 e_0 \dots e_{k+2} \Gamma(x) e_{k+1} \dots e_0 e_1.
$$

Thus

$$
\lambda^k e_1 e_0 \dots e_{k+2} \Gamma(x) e_{k+1} \dots e_0 e_1 = \pi(x) = \lambda^{-l} e_1 \dots e_l x e_{l+1} \dots e_1. \tag{10.9}
$$

Applying the trace to the left side and using that $e_i e_{i+1} e_i = \lambda e_i$ and $\Gamma(x) \in M'_{k+2} \cap M_{l+2}$ we get

$$
\lambda^k \tau(\Gamma(x) e_{k+1} \dots e_0 e_1 e_0 \dots e_{k+1}) = \tau(\Gamma(x) e_{k+1}) = \lambda \tau(\Gamma(x)).
$$

Similarly applying τ to the right side of (10.9) we get $\lambda\tau(x)$, so that $\tau(\Gamma(x)) = \tau(x)$. □

Remark 10.4.4. If $x \in M'_k \cap M_l$ then multiplying (10.9) by $e_{k+1} \dots e_0$ from the left and by $e_0 \dots e_{k+1}$ from the right, we conclude that $\Gamma(x)$ is an element in $M'_{k+2} \cap M_{l+2}$ such that

$$
\Gamma(x) e_{k+1} = \lambda^{k-l} e_{k+1} \dots e_l x e_{l+1} \dots e_{k+1}.
$$

Since $M'_{k+2} \subset B(L^2(M_1))$ is a factor and $e_{k+1} \in M_{k+2}$, the homomorphism $M'_{k+2} \ni a \mapsto ae_{k+1}$ is faithful. Hence $\Gamma(x)$ is uniquely determined by the above identity. In particular, the restriction of Γ to $\overline{\cup_{n\geq 0}(N' \cap M_n)} \subset A_\infty$ is independent of the choice of the tunnel. Often it is this restriction which is called the canonical shift.

It follows also that if for some $n \in \mathbb{Z}$ we consider

$$\ldots \subset M_k \subset M_{k+1} \subset \ldots$$

as the tower and the tunnel for $M_n \subset M_{n+1}$ and define the corresponding mirrorings γ_n and γ_{n+1}, then $\Gamma = \gamma_{n+1} \circ \gamma_n$. ♦

To compute the entropy of Γ we shall use Theorem 10.3.1. We have to check that its assumptions are satisfied.

To show that $\{M' \cap M_{2n}\}_n$ is a generating sequence for Γ we prove the next general proposition.

Proposition 10.4.5. *Let (P, τ, σ) be a W^*-dynamical system, where τ is a normal tracial state. Let $\{P_n\}_n$ be an increasing sequence of finite dimensional C^*-subalgebras such that $\sigma(P_n) \subset P_{n+1}$ and P is generated by $\sigma^k(P_n)$, $k \in \mathbb{Z}$, $n \in \mathbb{N}$. Assume there exists a sequence $\{k_n\}_n$ of natural numbers such that*

(i) $\dfrac{k_n}{n} \to 1$ *as* $n \to \infty$;

(ii) *the algebras $\sigma^{mk_n}(P_n)$, $m \in \mathbb{N}$, are mutually τ-independent.*

Then $\{P_n\}_n$ is a generating sequence for σ.

Recall that we say that $A_1, \ldots, A_n$ are τ-independent if they mutually commute, and

$$\tau(a_1 \ldots a_n) = \tau(a_1) \ldots \tau(a_n)$$

for $a_i \in A_i$.

Proof of Proposition 10.4.5. Since $\sigma^k(P_l) \subset \sigma^{-[n/2]}(P_n)$ for $-[n/2] \leq k \leq [n/2] - l$, the algebra $\cup_n \sigma^{-[n/2]}(P_n)$ is dense in P. Hence, by the Kolmogorov-Sinai type theorem, Theorem 3.2.3,

$$h_\tau(\sigma) = \lim_{n\to\infty} h_\tau(P_n; \sigma).$$

Since $\sigma^i(P_k) \subset P_{k+i}$, for any $n \in \mathbb{N}$ we have

$$H_\tau(P_k, \sigma(P_k), \ldots, \sigma^n(P_k)) \leq H_\tau(P_{n+k}),$$

whence

$$h_\tau(P_k; \sigma) \leq \liminf_{n\to\infty} \frac{1}{n} H_\tau(P_n).$$

On the other hand, since the algebras $\sigma^{mk_n}(P_n)$, $m \in \mathbb{Z}$, are independent, we have

$$h_\tau(P_n;\sigma^{k_n}) = H_\tau(P_n).$$

Hence

$$h_\tau(\sigma) = \frac{1}{k_n} h_\tau(\sigma^{k_n}) \geq \frac{1}{k_n} h_\tau(P_n;\sigma^{k_n}) = \frac{1}{k_n} H_\tau(P_n).$$

Since $\dfrac{k_n}{n} \to 1$, it follows that

$$h_\tau(\sigma) \geq \limsup_{n\to\infty} \frac{1}{n} H_\tau(P_n).$$

Thus the limit $\lim_n \dfrac{1}{n} H_\tau(P_n)$ exists and equals $h_\tau(\sigma)$, that is, $\{P_n\}_n$ is a generating sequence. □

The proposition implies that $\{M' \cap M_{2n}\}_n$ is a generating sequence for the canonical shift Γ. Indeed, since $\Gamma^{mn}(M' \cap M_{2n}) = M'_{2mn} \cap M_{2(m+1)n}$, we see that $\Gamma^{mn}(M' \cap M_{2n})$ commutes with $M' \cap M_{2n}$ for any $m \geq 1$. Moreover, since M_{2n} is a factor, we have $\tau(xy) = \tau(x)\tau(y)$ for any $x \in M_{2n}$ and $y \in \cup_k(M'_{2n} \cap M_{2n+k})$. Thus the algebras $\Gamma^{mn}(M' \cap M_{2n})$, $m \in \mathbb{N}$, are independent, and we can apply the previous proposition with $k_n = n$.

Next note that the von Neumann algebra generated by $\cup_n(M' \cap M_{2n})$ in the GNS-representation of A_∞ defined by τ has type II_1, since already the von Neumann algebra generated by the projections e_n, $n \geq 1$, is a II_1-factor by [214, Theorem XIX.3.1].

The generating sequence $\{M' \cap M_{2n}\}_n$ satisfies the commuting square condition, that is,

$$\begin{array}{ccc} M' \cap M_{2n} & \subset & M' \cap M_{2n+2} \\ \cup & & \cup \\ M'_2 \cap M_{2n} & \subset & M'_2 \cap M_{2n+2} \end{array}$$

is a commuting square for any $n \in \mathbb{N}$. Indeed, the trace preserving conditional expectation $M' \cap M_{2n+2} \to M' \cap M_{2n}$ is the restriction of the conditional expectation $E_{M_{2n}}\colon M_{2n+2} \to M_{2n}$. But then we obviously have $E_{M_{2n}}(M'_2 \cap M_{2n+2}) \subset M'_2 \cap M_{2n}$.

Next we check that the entropy is finite.

Proposition 10.4.6. *We have*

$$h_\tau(\Gamma) \leq ht(\Gamma) \leq \log[M\colon N].$$

Proof. If Ω is a finite subset of $M' \cap M_k$, then $\Gamma^n(\Omega) \subset M' \cap M_{2n+k}$. It follows that

$$ht(\Omega;\Gamma) \leq \limsup_{n\to\infty} \frac{1}{n} \log \operatorname{rank} M' \cap M_{2n}.$$

By [214, Proposition XIX.2.8] if $p_1, \ldots, p_m$ are projections in $M' \cap M_{2n}$ with sum 1, then

$$\sum_i \frac{[p_i M_{2n} p_i : M p_i]}{\tau(p_i)} = [M_{2n} : M].$$

Hence $\sum_i \tau(p_i)^{-1} \leq [M_{2n} : M]$, so by the Cauchy-Schwarz inequality

$$m = \sum_{i=1}^{m} \tau(p_i)^{1/2} \tau(p_i)^{-1/2} \leq [M_{2n} : M]^{1/2}.$$

Therefore $\operatorname{rank} M' \cap M_{2n} \leq [M_{2n} : M]^{1/2} = [M : N]^n$, and the proof of the proposition is complete. □

Finally, the multiplicities for the inclusion $M' \cap M_{n-1} \subset M' \cap M_n$ are bounded by $[M : N]$ (in fact, by $[M : N]^{1/2}$). To see this let p be a minimal projection in $M' \cap M_{n-1}$. Then $q = pe_n$ is a minimal projection in $M' \cap M_{n+1}$. Indeed, since $e_n M_{n+1} e_n = M_{n-1} e_n$, we have

$$e_n (M' \cap M_{n+1}) e_n = (M e_n)' \cap e_n M_{n+1} e_n = (M e_n)' \cap M_{n-1} e_n = (M' \cap M_{n-1}) e_n,$$

so that $q(M' \cap M_{n+1})q = p(M' \cap M_{n-1}) pe_n = \mathbb{C}q$. Furthermore, if f is a projection in $M' \cap M_n$ and $f \leq p$, then $qfq = e_n f e_n = E_{M_{n-1}}(f) e_n \neq 0$. Therefore if p majorizes a minimal projection $f \in M' \cap M_n$ then the central support of f in $M' \cap M_{n+1}$ majorizes q. In other words, if $(a_{ij})_{i,j}$ and $(b_{jk})_{j,k}$ are the inclusion matrices for $M' \cap M_{n-1} \subset M' \cap M_n$ and $M' \cap M_n \subset M' \cap M_{n+1}$, respectively, i_0 corresponds p and k_0 corresponds to q, then $b_{jk_0} \neq 0$ as soon as $a_{i_0 j} \neq 0$. If $\{z_k\}_k$ is the set of minimal central projections in $M' \cap M_{n+1}$, then denoting by s_k the trace of a minimal projection majorized by z_k, we get

$$\tau(p) = \sum_{j,k} a_{i_0 j} b_{jk} s_k \geq a_{i_0 j} b_{jk_0} \tau(q) = [M : N]^{-1} a_{i_0 j} b_{jk_0} \tau(p).$$

It follows that $a_{i_0 j} \leq [M : N]$ (in fact, it is known that $a_{i_0 j} = b_{jk_0}$, so that $a_{i_0 j} \leq [M : N]^{1/2}$).

Since $\ldots \subset M_{2k-2} \subset M_{2k} \subset \ldots$ can be considered as the tower and the tunnel associated with $M_{-2n} \subset M_{-2n+2}$, we conclude that the multiplicities for the inclusion $M'_{-2n} \cap M_{-2} \subset M'_{-2n} \cap M$ are bounded by $[M : N]^2$. Since $\gamma_0(M'_2 \cap M_{2n}) = M'_{-2n} \cap M_{-2}$ and $\gamma_0(M' \cap M_{2n}) = M'_{-2n} \cap M$, it follows that the multiplicities for the inclusion $\Gamma(M' \cap M_{2n-2}) \subset M' \cap M_{2n}$ are also bounded by $[M : N]^2$.

We have thus checked that all the assumptions of Theorem 10.3.1 are satisfied, so we get our first and most general result on the entropy of Γ.

Theorem 10.4.7. *Let Γ be the canonical shift for the inclusion $N \subset M$ of II_1-factors of finite index. Let $R = \pi_\tau(\cup_{n=1}^{\infty}(M' \cap M_{2n}))''$. Then*

$$h_\tau(\Gamma) = \lim_{n\to\infty} \frac{1}{n} H_\tau(M' \cap M_{2n}) = \frac{1}{2} H_\tau(R|\Gamma(R)) + \lim_{n\to\infty} \frac{1}{2n} H_\tau(Z(M' \cap M_{2n})).$$

□

Combining this theorem and Proposition 10.4.6 we obtain the following estimates.

Corollary 10.4.8. *We have*

$$\frac{1}{2}H_\tau(R|\Gamma(R)) \le h_\tau(\Gamma) \le ht(\Gamma) \le \log[M\colon N].$$

□

In particular, if $H_\tau(R|\Gamma(R)) = 2\log[M\colon N]$, then both the topological entropy and the dynamical entropy of Γ with respect to τ are equal to $\log[M\colon N]$. By [169, Theorem 5.3.1] the condition $H_\tau(R|\Gamma(R)) = 2\log[M\colon N]$ is one of several equivalent characterizations of extremality of $N \subset M$ and strong amenability of the standard invariant of the inclusion. The simplest case when this condition is satisfied, is that of finite depth inclusions.

A finite index inclusion $N \subset M$ of II_1-factors is said to have **finite depth** if

$$\sup_n \dim Z(M' \cap M_n) < \infty.$$

Then, by the discussion prior to [214, Proposition XIX.4.19], there exists n_0 such that if G denotes the matrix of the inclusion $M' \cap M_{2n_0} \subset M' \cap M_{2n_0+1}$, then the inclusions $M' \cap M_{2n} \subset M' \cap M_{2n+1}$ and $M' \cap M_{2n+1} \subset M' \cap M_{2n+2}$ are given by the matrices G and G^t, respectively, for any $n \ge n_0$. Moreover, the matrix GG^t is primitive, and if s denotes the vector formed by the values of the trace on minimal projections in $M' \cap M_{2n_0}$, then $GG^t s = [M\colon N]s$. We shall now show that this allows us to compute the entropy without relying on the deep results of Popa [169].

Recall that a matrix $B \in \mathrm{Mat}_r(\mathbb{R})$ with nonnegative coefficients is called primitive if there exists $m \in \mathbb{N}$ such that B^m has only positive entries. By the Perron-Frobenius theorem, see e.g. [193, Section 1.1], the spectral radius β of B is an eigenvalue of B, called the Perron-Frobenius eigenvalue, and the following conditions are satisfied:

(i) there is an eigenvector ξ of B with eigenvalue β whose coordinates are all positive; similarly, there is an eigenvector ζ of B^* with eigenvalue β and positive coordinates;

(ii) if ζ is normalized such that $(\xi, \zeta) = 1$ then $\beta^{-n}B^n \to P$ as $n \to +\infty$, where $P = (\cdot, \zeta)\xi$ is the projection onto $\mathbb{R}\xi$ along $\zeta^\perp$.

These properties imply $\cap_{n\in\mathbb{N}} B^n(\mathbb{R}^r_+) = \mathbb{R}_+\xi$. Indeed, assume $\vartheta_n \in \mathbb{R}^r_+$, $n \ge 0$, are such that $\vartheta_0 = \beta^{-n}B^n\vartheta_n$. Since ζ has positive coordinates, there exists $\varepsilon > 0$ such that $\|P\vartheta\| = (\vartheta, \zeta)\|\xi\| \ge \varepsilon\|\vartheta\|$ for $\vartheta \in \mathbb{R}^r_+$. Then if $\varepsilon' < \varepsilon$ and n is such that $\|\beta^{-n}B^n - P\| \le \varepsilon'$, we get $\|\beta^{-n}B^n\vartheta\| \ge (\varepsilon - \varepsilon')\|\vartheta\|$. It follows that the sequence $\{\vartheta_n\}_n$ is bounded. But then the vectors $\beta^{-n}B^n\vartheta_n$ become arbitrarily close to $P\vartheta_n$ as $n \to \infty$. Hence $\vartheta_0 \in \mathbb{R}_+\xi$.

It follows that ξ is the unique (up to a scalar factor) eigenvector of B with nonnegative coordinates. In particular, if $B = GG^t$ corresponds to a finite depth inclusion $N \subset M$, then the Perron-Frobenius eigenvalue is $[M\colon N]$.

Proposition 10.4.9. *Let $A = \varinjlim A_n$ be an AF-algebra such that for every n the inclusion $A_n \subset A_{n+1}$ is defined by a fixed primitive matrix B. Then there exists a unique tracial state τ on A, and*

$$\lim_{n\to\infty} \frac{1}{n} H_\tau(A_n) = \log \beta,$$

where β is the Perron-Frobenius eigenvalue of B.

Proof. Let $A_n = \oplus_{k=1}^r \mathrm{Mat}_{m_k^n}(\mathbb{C})$. A tracial state τ on A is defined by a sequence of vectors $s^n \in \mathbb{R}_+^r$, $n \in \mathbb{N}$, such that $\sum_k s_k^1 m_k^1 = 1$ and $s^n = Bs^{n+1}$. Let ξ be the Perron-Frobenius eigenvector of B normalized by the condition $\sum_k \xi_k m_k^1 = 1$. Then we can take $s^n = \beta^{-n+1}\xi$ and obtain a tracial state on A. Conversely, given a tracial state τ on A, we have $s^n \in \cap_m B^m(\mathbb{R}_+^r)$. Hence by the discussion before the formulation of the proposition the vector s^n is a scalar multiple of ξ for any n, and the scalar is completely determined by the condition $\tau(1) = 1$.

Then we have

$$\begin{aligned} H_\tau(A_n) &= -\sum_{k=1}^r m_k^n s_k^n \log s_k^n = -\sum_{k=1}^r m_k^n s_k^n \log(\beta^{-n+1}\xi_k) \\ &= (n-1)\log\beta - \sum_{k=1}^r m_k^n s_k^n \log \xi_k. \end{aligned}$$

Since $\sum_k m_k^n s_k^n = 1$, dividing by n and letting $n \to \infty$ we get the result. □

Applying this proposition to $A_n = M' \cap M_{2n}$ for a finite depth inclusion, we then get

$$\lim_{n\to\infty} \frac{1}{n} H_\tau(M' \cap M_{2n}) = \log[M{:}\,N].$$

We also obviously have

$$\lim_{n\to\infty} \frac{1}{n} H_\tau(Z(M' \cap M_{2n})) = 0.$$

Thus combining Proposition 10.4.6 and Theorem 10.4.7 we obtain the following result.

Theorem 10.4.10. *Let $N \subset M$ be an inclusion of II_1-factors of finite index and finite depth. Let $R = \pi_\tau(\cup_{n=1}^\infty (M' \cap M_n))''$. Then*

$$\frac{1}{2} H_\tau(R|\Gamma(R)) = h_\tau(\Gamma) = ht(\Gamma) = \log[M{:}\,N].$$

□

Remark that by [214, Corollary XIX.4.9] any inclusion of index < 4 has finite depth.

10.5 Shifts on Temperley-Lieb Algebras

Let λ be a real number such that $\lambda^{-1} \in \{4\cos^2 \frac{\pi}{n+1} \mid n \geq 3\} \cup [4, \infty)$. Consider the universal C*-algebra A generated by a sequence of projections $\{e_k\}_{k\in\mathbb{Z}}$ such that

$$e_k e_{k\pm 1} e_k = \lambda e_k, \quad e_k e_j = e_j e_k \;\text{ if }\; |k-j| \geq 2.$$

As we saw in the previous section, a representation of this algebra arises naturally from an inclusion of II_1-factors with index λ^{-1}. We also saw that there exists a trace τ on A such that

$$\tau(we_k) = \lambda\tau(w) \tag{10.10}$$

for any $k \in \mathbb{Z}$ and any w in the algebra generated by projections e_j with $j < k$. For $n \geq 2$ denote by A_n the C*-algebra generated by $e_1, \dots, e_{n-1}$, and set $A_0 = A_1 = \mathbb{C}1$. An easy induction argument shows that

$$A_n = A_{n-1} + A_{n-1} e_{n-1} A_{n-1}, \quad n \geq 2.$$

In particular, the trace is completely determined by (10.10).

The dimension and the trace vectors m^n and s^n of the algebras A_n, $n \in \mathbb{N}$, are described as follows, see § 3 of Chapter XIX in [214] for a proof.

If $\lambda^{-1} \geq 4$ then

$$A_n \cong \bigoplus_{k=0}^{[n/2]} \mathrm{Mat}_{m_k^n}(\mathbb{C}), \quad m_k^n = \binom{n}{k} - \binom{n}{k-1}, \quad s_k^n = \lambda^k P_{n-2k}(\lambda),$$

where $\{P_n\}_{n=0}^\infty$ is the sequence of polynomials such that $P_0 = P_1 = 1$,

$$P_{n+1}(t) = P_n(t) - tP_{n-1}(t), \quad n \geq 1.$$

On the other hand, if $\lambda^{-1} = 4\cos^2 \frac{\pi}{n+1}$ with $n \geq 3$ then for any $k \geq 0$ the embeddings $A_{n+2k-1} \hookrightarrow A_{n+2k+1}$ are given by a fixed primitive matrix with the Perron-Frobenius eigenvalue λ^{-1}.

Denote by θ_λ the automorphism of A defined by $\theta_\lambda(e_k) = e_{k+1}$.

Theorem 10.5.1. *We have:*

(i) $h_\tau(\theta_\lambda) = -\dfrac{1}{2}\log\lambda$, *if* $\lambda^{-1} \leq 4$;

(ii) $h_\tau(\theta_\lambda) = \eta(\alpha) + \eta(1-\alpha)$, *where* $\alpha = \dfrac{1+\sqrt{1-4\lambda}}{2}$, *if* $\lambda^{-1} \geq 4$.

Proof. The algebras $\theta_\lambda^{nm}(A_n)$, $m \in \mathbb{Z}$, are mutually τ-independent. By Proposition 10.4.5 it follows that $\{A_n\}_n$ is a generating sequence, that is,

$$h_\tau(\theta_\lambda) = \lim_{n\to\infty} \frac{1}{n} H_\tau(A_n).$$

If $\lambda^{-1} < 4$, then by Proposition 10.4.9

$$\lim_{n\to\infty} \frac{1}{n} H_\tau(A_{2n}) = \log \lambda.$$

Hence $h_\tau(\theta_\lambda) = -\frac{1}{2}\log\lambda$. It remains to consider the case $\lambda^{-1} \ge 4$.

Assume first that $\lambda^{-1} > 4$. Set

$$\alpha = \frac{1+\sqrt{1-4\lambda}}{2} \quad \text{and} \quad \beta = 1-\alpha = \frac{1-\sqrt{1-4\lambda}}{2}.$$

Then using that $\alpha\beta = \lambda$ we check that

$$P_n(\lambda) = \frac{\alpha^{n+1}-\beta^{n+1}}{\alpha-\beta}, \quad n \ge 0.$$

Since $s_k^{2n} = \lambda^k P_{2n-2k}(\lambda) = \lambda^k \alpha^{2n-2k}(1-(\beta/\alpha)^{2n-2k+1})(1-\beta/\alpha)^{-1}$, we see that the difference

$$\log s_k^{2n} - \log(\lambda^k\alpha^{2n-2k}) = \log s_k^{2n} - ((2n-k)\log\alpha + k\log\beta)$$

is uniformly bounded. Since

$$H_\tau(A_{2n}) = -\sum_{k=0}^{n} m_k^{2n} s_k^{2n} \log s_k^{2n} \quad \text{and} \quad \sum_{k=0}^{n} m_k^{2n} s_k^{2n} = 1,$$

the limit of $H_\tau(A_{2n})/2n$ coincides with the limit of

$$-\frac{1}{2n}\sum_{k=0}^{n} m_k^{2n} s_k^{2n}((2n-k)\log\alpha + k\log\beta) = -\log\alpha + \log\left(\frac{\alpha}{\beta}\right)\sum_{k=0}^{n} m_k^{2n} s_k^{2n}\frac{k}{2n}. \tag{10.11}$$

Since $s_k^{2n} = (\alpha\beta)^k(\alpha^{2n-2k+1}-\beta^{2n-2k+1})/(\alpha-\beta)$, we have

$$\sum_{k=0}^{n} m_k^{2n} s_k^{2n}\frac{k}{2n} = \frac{\alpha}{\alpha-\beta}\sum_{k=0}^{n} m_k^{2n}\alpha^{2n-k}\beta^k\frac{k}{2n} - \frac{\beta}{\alpha-\beta}\sum_{k=0}^{n} m_k^{2n}\alpha^k\beta^{2n-k}\frac{k}{2n}.$$

As $\alpha > \beta$, $m_k^{2n} \le \binom{2n}{k}$ and $4\alpha\beta = 4\lambda < 1$,

$$\sum_{k=0}^{n} m_k^{2n}\alpha^k\beta^{2n-k}\frac{k}{2n} \le (\alpha\beta)^n\sum_{k=0}^{n}\binom{2n}{k} \le (4\alpha\beta)^n \to 0.$$

On the other hand,

$$\begin{aligned}\sum_{k=0}^{n}\binom{2n}{k}\alpha^{2n-k}\beta^k\frac{k}{2n} &= \sum_{k=1}^{n}\binom{2n-1}{k-1}\alpha^{2n-k}\beta^k \\ &= \beta\sum_{k=0}^{n-1}\binom{2n-1}{k}\alpha^{2n-1-k}\beta^k \to \beta,\end{aligned}$$

since $\sum_{k=n}^{2n-1}\binom{2n-1}{k}\alpha^{2n-1-k}\beta^k \le \alpha^{-1}(\alpha\beta)^n \sum_{k=n}^{2n-1}\binom{2n-1}{k} \to 0$. Similarly

$$\begin{aligned}\lim_{n\to\infty}\sum_{k=0}^{n}\binom{2n}{k-1}\alpha^{2n-k}\beta^k\frac{k}{2n} &= \lim_{n\to\infty}\sum_{k=1}^{n}\binom{2n}{k-1}\alpha^{2n-k}\beta^k\frac{k-1}{2n}\\ &= \lim_{n\to\infty}\frac{\beta}{\alpha}\sum_{k=0}^{n-1}\binom{2n}{k}\alpha^{2n-k}\beta^k\frac{k}{2n} = \frac{\beta^2}{\alpha}.\end{aligned}$$

Therefore, since $m_k^{2n} = \binom{2n}{k} - \binom{2n}{k-1}$,

$$\begin{aligned}\lim_{n\to\infty}\sum_{k=0}^{n} m_k^{2n} s_k^{2n}\frac{k}{2n} &= \frac{\alpha}{\alpha-\beta}\lim_{n\to\infty}\sum_{k=0}^{n} m_k^{2n}\alpha^{2n-k}\beta^k\frac{k}{2n}\\ &= \frac{\alpha}{\alpha-\beta}\left(\beta-\frac{\beta^2}{\alpha}\right) = \beta,\end{aligned}$$

so that from (10.11) we get that

$$h_\tau(\theta_\lambda) = -\log\alpha + \log\left(\frac{\alpha}{\beta}\right)\beta = \eta(\alpha)+\eta(\beta).$$

Finally consider the case $\lambda^{-1} = 4$. Then

$$P_n(\lambda) = \frac{n+1}{2^n},$$

whence $s_k^{2n} = 2^{-2n}(2n-2k+1)$, and

$$\frac{1}{2n}H_\tau(A_{2n}) = -\frac{1}{2n}\sum_{k=0}^{n} m_k^{2n}s_k^{2n}\log s_k^{2n} = \log 2 - \sum_{k=0}^{n} m_k^{2n}s_k^{2n}\frac{\log(2n-2k+1)}{2n}.$$

Letting $n\to\infty$ we get $h_\tau(\theta_{1/4}) = \log 2$. □

Denote by M the von Neumann subalgebra of $\pi_\tau(A)''$ generated by e_n, $n \le -1$, and by R the von Neumann subalgebra generated by e_n, $n \ge 1$. By [214, Theorem XIX.3.1] both M and R are the hyperfinite II_1-factors, $N = \theta_\lambda^{-1}(M) \subset M$ is a subfactor of index λ^{-1}, and $N \subset M \subset \theta_\lambda(M) = \langle M, e_0\rangle$ is the basic construction. It follows that the automorphism θ_λ^2 of $\pi_\tau(A)''$ is (the extension to the weak operator closure of) the canonical shift associated with the inclusion $N \subset M$.

It is clear that $A_n \subset M' \cap M_n$ for any $n \ge 0$. It can be shown that if $\lambda^{-1} \le 4$ then $A_n = M' \cap M_n$. For $\lambda^{-1} < 4$ the inclusion $N \subset M$ has finite depth, and from Theorem 10.4.10 we again obtain

$$h_\tau(\theta_\lambda) = \frac{1}{2}h_\tau(\theta_\lambda^2) = \frac{1}{2}\log[M\colon N] = -\frac{1}{2}\log\lambda.$$

On the other hand, if $\lambda^{-1} > 4$ then it can be shown that already $M' \cap M_1$ is different from $A_1 = \mathbb{C}1$. It is not difficult to check that the generating sequence $\{A_n\}_n$ satisfies the commuting square condition. Hence by Theorem 10.3.1

$$h_\tau(\theta_\lambda) = \frac{1}{2} H_\tau(R|\theta_\lambda(R)).$$

Therefore $H_\tau(R|\theta_\lambda(R)) = 2(\eta(\alpha) + \eta(\beta))$. This equality, as well as the entropy $h_\tau(\theta_\lambda)$, can also be obtained from the embedding $A \hookrightarrow \mathrm{Mat}_2(\mathbb{C})^{\otimes \mathbb{Z}}$,

$$e_k \mapsto (1-\lambda) e_{11}^k \otimes e_{22}^{k+1} + \lambda e_{22}^k \otimes e_{11}^{k+1} + \sqrt{\lambda(1-\lambda)} (e_{12}^k \otimes e_{21}^{k+1} + e_{21}^k \otimes e_{12}^{k+1}),$$

where $\{e_{ij}^k\}_{i,j=1}^2$ are the matrix units in the k-th factor $\mathrm{Mat}_2(\mathbb{C})$. It can be shown, see [162], that in the GNS-representation of $\mathrm{Mat}_2(\mathbb{C})^{\otimes \mathbb{Z}}$ corresponding to the product-state

$$\psi = \varphi^{\otimes \mathbb{Z}}, \quad Q_\varphi = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix},$$

the weak operator closure of A coincides with the centralizer of the state. Thus θ_λ is one of the Bernoulli shifts on the hyperfinite factor considered in Example 3.2.6(ii). Moreover, the von Neumann algebra generated by the projections e_n, $n \ge 1$, coincides with the centralizer of ψ on $\pi_\psi(\mathrm{Mat}_2(\mathbb{C})^{\otimes \mathbb{N}})''$. In particular, there exists a projection $p \in R$ which is mapped onto e_{11}^1 under the above embedding, and then

$$\tau(p) = \alpha, \quad p \in \theta_\lambda(R)' \cap R, \quad pRp = \theta_\lambda(R)p, \quad (1-p)R(1-p) = \theta_\lambda(R)(1-p).$$

It should be remarked that the projections e_n, $n \le -1$, do not generate the centralizer of ψ on $\pi_\psi(\otimes_{n \le 0} \mathrm{Mat}_2(\mathbb{C}))''$.

10.6 Notes

The notion of relative entropy for subalgebras of a von Neumann algebra with a normal tracial state was introduced by Connes and Størmer [51], who used it as a tool to prove continuity of mutual entropy. The definition was extended to arbitrary states by Connes [49]. To see that relative entropy is relevant for estimates of mutual entropy, observe that for finite dimensional C*-subalgebras $P_1, \dots, P_n$ and $Q_1, \dots, Q_n$ of M we have

$$H_\varphi(P_1, \dots, P_n) \le H_\varphi(Q_1, \dots, Q_n) + \sum_{k=1}^n H_\varphi(P_k|Q_k).$$

It was Pimsner and Popa [162] who undertook a serious study of this notion and discovered its connection with Jones' index of II$_1$-factors. Sects. 10.1 and 10.2 are almost entirely based on their paper. For some examples of computations of relative entropy see [22], [103], [233]. The results of Pimsner and Popa were extended to type III factors by Hiai [86].

Theorem 10.3.1 was proved by Størmer [210] extending results of Choda [40] and Hiai [88]. See [81] for further discussion of generating sequences.

The canonical shift was introduced by Pimsner and Popa [162] and Ocneanu [146]. One usually defines the canonical shift a bit differently. As we remarked after the proof of Lemma 10.4.1, for any $n \in \mathbb{N}$ the algebras $M_{-1} \subset M_n \subset M_{2n+1}$ form a basic construction. Pimsner and Popa [163] found an explicit formula for the corresponding Jones projection in M_{2n+1}. In other words, they found an explicit representation of M_{2n+1} on $L^2(M_n)$ extending the representation of M_{n+1}. Using this representation we can define an antiautomorphism of $M'_{-1} \cap M_{2n+1}$, which for the moment we denote by γ'_n. Then one checks that $\gamma'_{n+2} \circ \gamma'_{n+1}$ and $\gamma'_{n+1} \circ \gamma'_n$ agree on $M_{-1} \subset M_n \subset M_{2n+1}$, so there exists a well-defined endomorphism of $\overline{\cup_n (M'_{-1} \cap M_{2n+1})}$. To see that we get the same endomorphism as the one defined in the present chapter, by Remark 10.4.4 it suffices to check that $\gamma'_n = \gamma_n$ on $M'_{-1} \cap M_{2n+1}$, or even better, the representation of M_{2n+1} on $L^2(M_n)$ alluded to above coincides with the representation defined as in Lemma 10.4.1. Since these representations are determined by the images of $e_{n+1}, \dots, e_{2n}$, and $\gamma_n(e_{n+k}) = e_{n-k}$ by construction, we just have to check that $\gamma'_n(e_{n+k}) = e_{n-k}$ for $k = 1, \dots, n$. This is indeed the case, see e.g. [54]. Alternatively one can use the identity in Remark 10.4.4 and an explicit formula for the canonical shift, see e.g. [23].

The entropy of the shifts on the Temperley-Lieb algebras, Theorem 10.5.1, was computed by Pimsner and Popa [162] for all $\lambda \neq 1/4$. For the case $\lambda < 1/4$ instead of the elementary but tedious computations presented above they proved that the shifts are isomorphic to the Bernoulli shifts on the hyperfinite factor considered in Example 3.2.6(ii); see [191] for a more general result. The missing case $\lambda = 1/4$ was treated by Choda [40] and Yin [236]. The latter paper also contains the computation of the entropy for $\lambda < 1/4$, which we used.

For arbitrary subfactors the study of entropic properties of the canonical shift was initiated by Choda [41]. In particular, she obtained the inequalities from Corollary 10.4.8 and computed the entropy in the case of a finite depth subfactor, Theorem 10.4.10. The most general result, Theorem 10.4.7, was proved by Hiai [88]. For similar results in the type III case see [48], [87]. An argument similar to the proof of Proposition 10.4.5 appeared already in the paper of Connes and Størmer [51] and had been used by several authors until it was formalized by Choda [40], [41]. Proposition 10.4.9 was discovered by so many people that it should probably be considered a folklore. It was successfully applied by Choda [40], [41] to a number of models.

http://www.springer.com/978-3-540-34670-8

Dynamical Entropy in Operator Algebras
Neshveyev, S.; Størmer, E.
2006, X, 296 p., Hardcover
ISBN: 978-3-540-34670-8