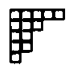
$\Omega \subset \mathbb{K}$ and construct for $q \notin \Omega$ a complete family of irreducible representations of $H_{q,n}$; this will demonstrate that $H_{q,n} \cong \mathbb{K}[\mathfrak{S}_n]$ for all n and for $q \notin \Omega$.

2.10.c – Complex representations of the symmetric group.

The conjugacy classes of the group $\mathfrak{S}_n$ are naturally indexed by partitions of n, two permutations being conjugate if and only if they have the same cycle structure. Thus there are as many irreducible representations of $\mathfrak{S}_n$ (over $\mathbb{C}$) as there are partitions. Although one cannot expect a natural correspondence between representations and partitions on the above grounds, it has long been known how to construct an irreducible representation from a partition. It is convenient to represent partitions by "Young diagrams", as amply illustrated by the following example.

Example 2.10.5. To the partition $\lambda = [\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5] = [5,3,2,2,1]$ of 13, one associates the Young diagram



The most important rule is the restriction rule: if one restricts the representation of $\mathfrak{S}_n$ corresponding to a Young diagram Y to $\mathfrak{S}_{n-1}$, it is isomorphic to the direct sum of all representations corresponding to all Young diagrams Y' obtained by removing one box from Y, all occuring with multiplicity one.

Thus the irreducible representations of $\mathfrak{S}_n$ (and hence the Bratteli diagram for $\mathbb{C}\mathfrak{S}_1 \subset \mathbb{C}\mathfrak{S}_2 \subset \mathbb{C}\mathfrak{S}_3 \subset \cdots$) are conveniently pictured by the following important diagram:

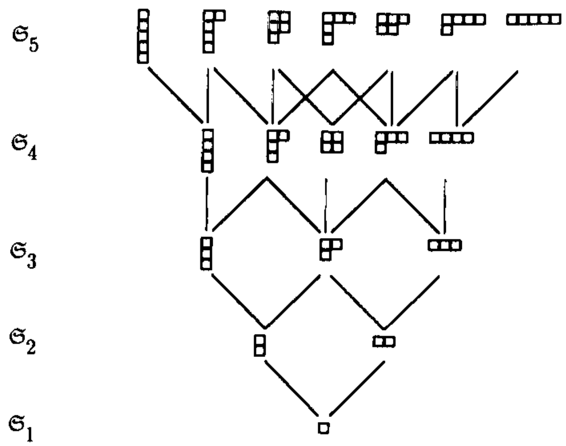


figure 2.10.6

The dimensions of the corresponding representations are given by the number of ascending paths on 2.10.6 beginning with $\square$ and ending at the Young diagram in question. The above facts will actually follow from the construction of irreducible representations for the Hecke algebras $H_{q,n}$, to which we now return.

2.10.d – Irreducible representations of $H_{q,n}$ for $q \notin \Omega$.

The material that follows is taken from Wenzl's thesis [Wen2]. The $\mathbb{K}$-algebra $H_{q,n}$ is that defined at the beginning of 2.10.b; in particular, the field $\mathbb{K}$ is arbitrary.

By our intuitive argument, we expect that figure 2.10.6 should also represent the structure of $H_{q,n}$ for all but a countable number of values of q. While this could be proved in an elegant manner due to Tits (see exercise 26 in §IV.2 of [BLie], or Lemma 85 in [Ste2]), two important pieces of information would be missing: there would be no indication of which values of q are "bad" (and that would be particularly frustrating for $\mathbb{K}$ countable!), and there would be no construction of concrete representations of $H_{q,n}$.

We shall now show how to construct an irreducible representation of $H_{q,n}$ for each partition of n, provided q is not in the set Ω defined below. It is first convenient to dispose of another presentation of $H_{q,n}$ than that of 2.10.b.

Proposition 2.10.7. *Consider a number* $q \in \mathbb{K}^*$ *and an integer* $n \geq 1$. *Assume* $q \neq -1$ *and set*

(a) $e_i = \dfrac{g_i + 1}{q+1}$ $\quad i = 1, \cdots, n-1$

These generate $H_{q,n}$ *and constitute with the relations*

(b) $e_i^2 = e_i$ $\quad i = 1,\cdots,n-1$

(c) $e_ie_{i+1}e_i - q(q+1)^{-2}e_i = e_{i+1}e_ie_{i+1} - q(q+1)^{-2}e_{i+1}$ $\quad i = 1,\cdots,n-2$

(d) $e_ie_j = e_je_i$ when $|i-j| \geq 2$ $\quad i,j = 1,\cdots,n-1$

a presentation of $H_{q,n}$.

Proof. A straightforward computation. #

Naturally this demands comparison with the definition of $A_{\beta,n}$ in Section 2.8. However we shall postpone further comments on this until the next section.

Define the subset Ω of $\mathbb{K}$ to be the union of $\{0\}$ with those q for which there exists an integer $n \geq 1$ with $\sum_{j=0}^{n} q^j = 0$. Thus $\Omega \setminus \{0\}$ is the set of non-trivial roots of 1 in characteristic 0 and the set of all roots of 1 in finite characteristic. (As already noticed in section 2.8, if $q \notin \Omega$ then $\beta = q^{-1}(q+1)^2$ is generic.) For each $d \in \mathbb{Z} \setminus \{0\}$ and $q \in \mathbb{K} \setminus \Omega$ define

$$a_d(q) = \begin{cases} \dfrac{1-q^{d+1}}{(1+q)(1-q^d)} & \text{if } q \neq 1 \\[2ex] \dfrac{d+1}{2d} & \text{if } q = 1 \end{cases}$$

(Remark: When $d > 0$, then $a_d(q) = \dfrac{1+q+\cdots+q^d}{(1+q)(1+q+\cdots+q^{d-1})}$. Note also that $\prod_{1 \leq d \leq k} a_d(q) = Q_k(q)$ where Q_k is as in Proposition 2.8.3.iv.) Suppose given a partition of n, say $\lambda = [\lambda_1,\cdots,\lambda_k]$, where we allow some of the last λ_j's to be zero. We think of λ as a Young diagram. Let V_λ be the free $\mathbb{K}$-vector space on the set of ascending paths p from $\square$ to λ on figure 2.10.6, and denote by $\{v_p\}$ its canonical basis. We define now endomorphisms $f_1,\cdots,f_{n-1}$ of V_λ.

Let $i \in \{1,\cdots,n-1\}$. For each ascending path

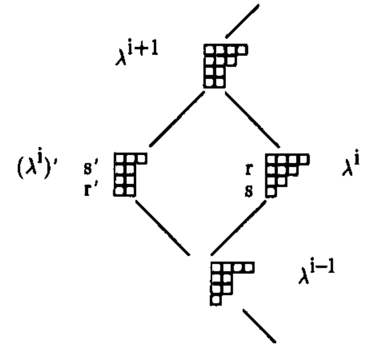
$$p = (\lambda^1 = \square, \lambda^2, \cdots, \lambda^n = \lambda)$$

we have to define f_iv_p. The partition λ^{i+1} is obtained from λ^{i-1} in one of three ways

(a) By adding two boxes to the same column of λ^{i-1}. In this case $f_iv_p = v_p$.

(b) By adding two boxes to the same row of λ^{i-1}. In this case $f_iv_p = 0$.

(c) By adding boxes in different rows and columns of λ^{i-1}; more precisely there is pair of integers (r,s) with $r \neq s$ and $\lambda_r^{i-1} \neq \lambda_s^{i-1}$ such that $\lambda_r^i = \lambda_r^{i-1} + 1$ and $\lambda_s^{i+1} = \lambda_s^{i-1} + 1$. In this case there is precisely one ascending path from $\square$ to λ which differs from p in its i^{th} vertex only; we call this path p'. For example:



Set

$$d = (\lambda_r^{i+1}-r) - (\lambda_s^{i+1}-s) = (s-r) + (\lambda_r^{i+1}-\lambda_s^{i+1})$$

and observe that $d \neq 0$. Define d' in the same way for the path p' and note that $d' = -d$. Finally, define

$$f_iv_p = a_d(q)v_p + (1-a_d(q))v_{p'}.$$

Observe that f_i leaves invariant the subspace $\mathbb{K}v_p \oplus \mathbb{K}v_{p'}$ of V_λ as well as its canonical complement; on $\mathbb{K}v_p \oplus \mathbb{K}v_{p'}$, it is described by the matrix

$$\begin{bmatrix} a_d(q) & a_d(q) \\ 1-a_d(q) & 1-a_d(q) \end{bmatrix} \qquad (2.10.8).$$

We have taken advantage of the equality $a_d(q) + a_{d'}(q) = 1$, which follows from the definition of a_d and from $d + d' = 0$.

The verification that $f_1, \cdots, f_{n-1}$ satisfy the relations (b) and (d) of Proposition 2.10.7 is trivial. They also satisfy (c), but this is more tedious to check and we refer to [Wen2]. We conclude that, for each partition λ of n there is a representation π_λ of $H_{q,n}$ in V_λ defined by $\pi_\lambda(e_i) = f_i$. A remarkably easy inductive argument shows that the π_λ's are irreducible and mutually inequivalent when λ runs over the set $\mathcal{P}_n$ of all partitions of n (for $q \in \mathbb{K} \setminus \Omega$). Indeed, these representations are absolutely irreducible, because the same argument applies to any extension of $\mathbb{K}$. By theorems of Burnside and Frobenius-Schur, this implies that $H_{q,n}$ has a quotient isomorphic to the multi-matrix algebra $\bigoplus_{\lambda \in \mathcal{P}_n} \mathrm{End}_{\mathbb{K}}(V_\lambda)$, of dimension $n!$. But we have already reported that the dimension of $H_{q,n}$ is no more than $n!$. (See the end of 2.10.a above, and §4 in [HKW].) Consequently the dimension is precisely $n!$, we have a complete set of irreducible representations of $H_{q,n}$ for $q \in \mathbb{K} \setminus \Omega$, and $H_{q,n}$ is isomorphic to $\bigoplus_{\lambda \in \mathcal{P}_n} \mathrm{End}_{\mathbb{K}}(V_\lambda)$. (In particular, setting $q = 1$, this gives for $\mathbb{K} = \mathbb{Q}$ the usual complete set of irreducible representations of the symmetric group $\mathfrak{S}_n$.)

Another trivial consequence of the construction is that the restriction of a representation π_λ of $H_{q,n}$ to $H_{q,n-1}$ is a direct sum $\oplus\, \pi_{\lambda'}$, where λ' runs over all partitions of $n-1$ obtained from the partition of λ of n by removing one box from the Young diagram. We reformulate this as follows.

<u>Theorem 2.10.9</u>. *Let $\mathbb{K}$ be a field and let $\Omega \subset \mathbb{K}$ be the union of $\{0\}$, of the non trivial roots of 1, and of 1 in case $\mathrm{char}(\mathbb{K}) \neq 0$. Consider $q \in \mathbb{K} \setminus \Omega$, an integer $n \geq 1$, and the Hecke algebra $H_{q,n}$ generated by $g_1, \cdots, g_{n-1}$ with the relations of 2.10.b, or equivalently by $e_1, \cdots, e_{n-1}$ and the relations of Proposition 2.10.7. Then*

(a) $H_{q,n}$ *is of dimension* $n!$.

(b) $H_{q,n}$ *is a multi-matrix algebra.*

(c) *The natural mapping* $H_{q,n} \to H_{q,n+1}$ *is an imbedding.*

(d) *The structure of the chain* $H_{q,1} \subset H_{q,2} \subset \cdots \subset H_{q,n} \subset \cdots$ *is given by figure* 2.10.6.

We make one further comment on Wenzl's paper, and for this we assume $\mathbb{K} = \mathbb{C}$. His exposition does not involve the matrix 2.10.8, but rather the related one

$$\begin{pmatrix} a_d & \sqrt{a_d(1-a_d)} \\ \sqrt{a_d(1-a_d)} & 1-a_d \end{pmatrix}$$

with a_d written for $a_d(q)$. It follows that $H_{q,n}$ has a C^*-algebra structure for $q \in \mathbb{R}$ and $q > 0$, for which e_i is an orthogonal projection: In fact, the main interest of [Wen2] is in the values $q = e^{\pm 2\pi i/n} \in \Omega$, for which Wenzl has constructed C^*-algebras which are quotients of the corresponding Hecke algebras.

2.11. The relationship between $\mathcal{A}_{\beta,n}$ and Hecke algebras.

It was first pointed out to V. Jones by R. Steinberg that the defining relations of $\mathcal{A}_{\beta,n}$ (see section 2.8) actually imply the Hecke relations. This is obvious from the definition of $\mathcal{A}_{\beta,n}$ and Proposition 2.10.7, but we would rather state this in terms of the generators g_i again.

<u>Proposition 2.11.1</u>. *Consider $q \in \mathbb{K} \setminus \{-1,0\}$ and $\beta = 2 + q + q^{-1} \in \mathbb{K}^*$, an integer $n \geq 1$, and the algebra $\mathcal{A}_{\beta,n}$. Set*

$$\gamma_i = (q+1)\epsilon_i - 1 \quad \textit{so that} \quad \epsilon_i = \frac{\gamma_i + 1}{q+1} \qquad i = 1, \cdots, n-1.$$

These generate $\mathcal{A}_{\beta,n}$ and constitute with

$$\begin{aligned}
&\gamma_i^2 = (q-1)\gamma_i + q && i = 1,\cdots,n-1 \\
&\gamma_i\gamma_{i+1}\gamma_i = \gamma_{i+1}\gamma_i\gamma_{i+1} && i = 1,\cdots,n-2 \\
&\gamma_i\gamma_j = \gamma_j\gamma_i \quad \textit{when } |i-j| \geq 2 && i,j = 1,\cdots,n-1 \\
&\gamma_i\gamma_{i+1}\gamma_i + \gamma_i\gamma_{i+1} + \gamma_{i+1}\gamma_i + \gamma_i + \gamma_{i+1} + 1 = 0 && i = 1,\cdots,n-2
\end{aligned}$$

a presentation of $\mathcal{A}_{\beta,n}$.

<u>Proof</u>. A straightforward computation. #

<u>Corollary 2.11.2</u>. *There is a surjective algebra morphism*

$$\psi_n : H_{q,n} \to \mathcal{A}_{\beta,n}$$

defined by $\psi_n(g_i) = \gamma_i$ for $i = 1, \cdots, n-1$. If $n = 1$ or $n = 2$ it is an isomorphism. If $n \geq 3$ its kernel I_n is the two-sided ideal of $H_{q,n}$ generated by

$$g_1g_2g_1 + g_1g_2 + g_2g_1 + g_1 + g_2 + 1.$$

Moreover the diagram

$$\begin{tikzcd} H_{q,n} \arrow[r] \arrow[d, "\psi_n"] & H_{q,n+1} \arrow[d, "\psi_{n+1}"] \\ \mathcal{A}_{\beta,n} \arrow[r] & \mathcal{A}_{\beta,n+1} \end{tikzcd}$$

commutes.

Proof. The existence of ψ_n follows from the definition of $H_{q,n}$ and Proposition 2.11.1. It is clear that ψ_1 and ψ_2 are isomorphisms, and that for $n \geq 3$ the kernel of ψ_n is generated by

$$x_i = g_ig_{i+1}g_i + g_ig_{i+1} + g_{i+1}g_i + g_i + g_{i+1} + 1$$

for $i = 1, \cdots, n-2$. As $q \neq 0$, each g_i is invertible with inverse $q^{-1}(g_i+1-q)$. By relations (b) and (c) in Proposition 2.10.3 on has

$$(g_1g_2\cdots g_{n-1})g_k(g_{n-1}^{-1}\cdots g_2^{-1}g_1^{-1}) = g_{k+1} \quad k = 1, \cdots, n-2$$

and consequently

$$(g_1g_2\cdots g_{n-1})x_i(g_{n-1}^{-1}\cdots g_2^{-1}g_1^{-1}) = x_{i+1} \quad i = 1, \cdots, n-3$$

Thus I_n is generated by x_1. The last claim is obvious. #

Thus, if q is a value for which $H_{q,n}$ is semi-simple, it must be possible to identify $\mathcal{A}_{\beta,n}$ with a certain ideal of $H_{q,n}$, given by some subset of the set of partitions of n. We shall show that this subset is precisely the set of all partitions with at most two rows; this was also explained to us by R. Steinberg.

We recall that Ω has been defined in 2.10.d, and that the relations which constitute with $g_1, \cdots, g_{n-1}$ a presentation of $H_{q,n}$ are written in Proposition 2.10.3 (see also 2.10.b).

Lemma 2.11.3. *Let* $q \in \mathbb{K} \setminus \Omega$, *let* $n \geq 3$ *be an integer, let* λ *be Young diagram with* n *boxes and with at most two rows, and let*

$$\pi_\lambda : H_{q,n} \to \mathrm{End}(V_\lambda)$$

be as in 2.10.d. *If* $x_i = g_ig_{i+1}g_i + g_ig_{i+1} + g_{i+1}g_i + g_i + g_{i+1} + 1$, *then* $\pi_\lambda(x_i) = 0$ *for* $i = 1, \cdots, n-2$.

Proof. We set first $n = 3$. As there are two partitions of 3 with at most two rows, we split the proof in two steps.

First $\lambda = [3]$, pictured as $\square\square\square$. Here, according to the definition of $\pi_\lambda(g_1) = \pi_\lambda(g_2) = -1$ and $\pi_\lambda(x_1) = 0$.

Second $\lambda = [2,1]$, pictured as $\begin{smallmatrix}\square\square\\ \square\ \end{smallmatrix}$. Here, instead of using π_λ, we may argue with any 2-dimensional irreducible representation of $H_{q,3}$, for example that defined by

$$\pi(g_1) = \begin{bmatrix} q & 0 \\ q & -1 \end{bmatrix} \quad \pi(g_2) = \begin{bmatrix} -1 & 1 \\ 0 & q \end{bmatrix}$$

(which is irreducible if q is neither 0, nor -1, nor a nontrivial cube root of 1). A routine calculation shows that $\pi(x_1) = 0$.

Assume now $n \geq 4$, and that the lemma holds for $n-1$. By the proof of 2.11.2, it is enough to check that $\pi_\lambda(x_1) = 0$.

We recall from 2.10.d that

$$\pi_\lambda \Big|_{H_{q,n-1}} = \bigoplus_{\lambda'} \pi_{\lambda'}$$

where λ' has one less box than λ. In particular λ' has at most two rows, and $\pi_{\lambda'}(x_1) = 0$ by the induction hypothesis. Consequently $\pi_\lambda(x_1) = 0$. #

Consider $q \in \mathbb{K} \setminus \Omega$ and an integer $n \geq 3$. Let $\mathcal{P}_n$ be the set of partitions of n. We know from section 2.10.d that $H_{q,n}$ is a direct sum $\bigoplus_{\lambda \in \mathcal{P}_n} I_\lambda$ of simple two sided ideals, the notation being such that, for each $\lambda_0 \in \mathcal{P}_n$, the representation π_{λ_0} of $H_{q,n}$ restricts to an isomorphism $I_{\lambda_0} \to \mathrm{End}(V_{\lambda_0})$ and maps I_λ to $\{0\}$ when $\lambda \neq \lambda_0$. We denote by $\mathcal{P}_n^2$ the subset of $\mathcal{P}_n$ of partitions with at most two rows.

Proposition 2.11.4. *Let* $q \in \mathbb{K} \setminus \Omega$, *let* $n \geq 3$ *be an integer, and let*

$$I_n = \mathrm{Ker}(\psi_n : H_{q,n} \to \mathcal{A}_{\beta,n}) = \langle g_1g_2g_1+g_1g_2+g_2g_1+g_1+g_2+1 \rangle$$

be as in Corollary 2.11.2. *Then*

$$I_n = \bigoplus_{\lambda \in \mathcal{P}_n \setminus \mathcal{P}_n^2} I_\lambda .$$

Proof. By the previous lemma for $\lambda \in \mathcal{P}_n^2$, one has $\pi_\lambda(I_n) = 0$, so that $I_\lambda \cap I_n = (0)$. Consequently $I_n \subset \bigoplus_{\lambda \in \mathcal{P}_n \setminus \mathcal{P}_n^2} I_\lambda$.

Let $\lambda = [r,s] \in \mathcal{P}_n^2$. From the definition of V_λ (see 2.10.d) one has

$$\begin{aligned} &\dim V_{[r,s]} = \dim V_{[r,s-1]} && \text{if } r = s, \\ &\dim V_{[r,s]} = \dim V_{[r,s-1]} + \dim V_{[r-1,s]} && \text{if } r > s \geq 1, \\ &\dim V_{[r,0]} = 1. \end{aligned}$$

By induction on r and s, ones deduces from this

$$\dim V_{[r,s]} = \begin{bmatrix} n \\ s \end{bmatrix} - \begin{bmatrix} n \\ s-1 \end{bmatrix} .$$

Thus

$$\sum_{\lambda \in \mathcal{P}_n^2} (\dim V_\lambda)^2 = \frac{1}{n+1}\begin{bmatrix} 2n \\ n \end{bmatrix},$$

by Lemma 2.8.2.

Suppose that $I_{\lambda_0} \not\subset I_n$ for some $\lambda_0 \in \mathcal{P}_n \setminus \mathcal{P}_n^2$, so that $I_n \subset \ker(\pi_{\lambda_0})$; we shall arrive at a contradiction. For $\lambda \in \mathcal{P}_n^2 \cup \{\lambda_0\}$, the representation π_λ of $H_{q,n}$ defines a representation $\hat{\pi}_\lambda$ of $\mathcal{A}_{\beta,n}$ in V_λ. As the π_λ's are pairwise inequivalent, so are the $\hat{\pi}_\lambda$'s, and

$$\dim \mathcal{A}_{\beta,n} \geq \frac{1}{n+1}\begin{bmatrix} 2n \\ n \end{bmatrix} + (\dim V_{\lambda_0})^2 .$$

But this contradicts Proposition 2.8.1. #

Observe that the proof above shows again the equality

$$\dim \mathcal{A}_{\beta,n} = \frac{1}{n+1}\begin{bmatrix} 2n \\ n \end{bmatrix} \quad \text{for } \beta \text{ generic}$$

of Theorem 2.8.5.

To sum up, we have shown that, for generic β and corresponding q, the algebra $\mathcal{A}_{\beta,n}$ is isomorphic to the quotient of $H_{q,n}$ by the two-sided ideal generated by $g_1g_2g_1 + g_1g_2 + g_2g_1 + g_1 + g_2 + 1$. This ideal corresponds precisely to the direct summands I_λ of $H_{q,n}$ given by all Young diagrams $\lambda \in \mathcal{P}_n$ having at least 3 rows.

CHAPTER 3

Finite von Neumann Algebras with Finite Dimensional Centers

3.1. Introduction.

In this chapter we study pairs of finite von Neumann algebras with finite dimensional centers, and the index of such pairs.

Sections 2 to 4 are purely expository, and may be taken as an encouragement to the reader having essentially no previous experience with von Neumann algebras. Sections 5 to 7 present a generalization to the present setting of some of the ideas of [Jo1] for pairs of factors. Though this chapter cannot be so self-contained as the previous ones, we have tried to minimize the technical background in operator algebras assumed from the reader.

Let us first describe Sections 2 to 4. Let M be a von Neumann algebra which is a factor of type II_1. (The definition is given in Section 3.2.) We denote by $tr : M \to \mathbb{C}$ the normalized trace on M. For every Hilbert space H on which M acts, Murray and von Neumann have defined a positive number (possibly ∞) called the coupling constant between M and its commutant; we denote this number by $\dim_M(H)$. Two representations of M by operators on two separable Hilbert spaces H and H' are equivalent if and only if $\dim_M(H) = \dim_M(H')$. Section 3.2 is an exposition of the definition and the basic properties of these coupling constants. Except for the presentation, all this material comes from the original papers by Murray and von Neumann.

In Section 3.3, we present some *geometric examples* of coupling constants arising in the theory of discrete series representations of Lie groups; they are borrowed from Atiyah–Schmid [AS]. In particular, we show:

Theorem 3.1.1. *Let G be a connected real semi–simple, non–compact Lie group without center. Let Γ be a lattice in G, and let M be the von Neumann algebra of the discrete group Γ. Then M is a II_1 factor. If $\pi : G \longrightarrow U(H)$ is an irreducible discrete series representation of G, then $\pi|_\Gamma$ extends to a representation of M on H, and*

$$\dim_M(H) = \mathrm{covol}(\Gamma)\, d_\pi,$$

where d_π is the formal dimension of π.

In Section 3.4, we consider a pair $N \subset M$ of finite factors and we recall some aspects of the original work [Jo1] on this subject. First the index of N in M is now defined to be

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

where $L^2(M)$ is the Hilbert space obtained by completion of M for the scalar product $\langle x|y\rangle = tr(x^*y)$. It was shown in [Jo4] that this definition of index agrees with the purely ring–theoretic definition of Chapter 2.

If $[M:N] < \infty$, the pair $N \subset M$ generates a tower of II_1–factors

$$1 \in M_0 = N \subset M_1 = M \subset \cdots \subset M_{k-1} \subset M_k \subset \cdots$$

by a fundamental construction which is defined as follows. The natural conditional expectation from M_k onto M_{k-1} can be seen as an orthogonal projection $e_k : L^2(M_k) \to L^2(M_{k-1})$, and M_{k+1} is the von Neumann algebra of operators on $L^2(M_k)$ generated by M_k and e_k. This M_{k+1} is again a II_1 factor. It is a particular case of Proposition 3.1.4 below that this way to define the fundamental construction agrees with that of Chapter 2. Moreover the Markov relation holds:

$$[M:N]tr_{k+1}(xe_k) = tr_k(x) \qquad \text{for all } x \in M_k,$$

where tr_k and tr_{k+1} denote the normalized traces on M_k and M_{k+1} respectively.

The sequence $(e_k)_{k\geq 1}$ of projections in $\bigcup_{k\geq 0} M_k$ satisfy the relations

$$\begin{aligned} &[M:N]\, e_ie_je_i = e_i && \text{if } |i-j| = 1 \\ &e_ie_j = e_je_i && \text{if } |i-j| \geq 2 \end{aligned}$$

and provide consequently a representation of the algebras $A_{\beta,k}$, with $\beta = [M:N]$. (See Section 2.8 and Theorem II.16.) From this follows

Theorem 3.1.2. *If $N \subset M$ is a pair of II_1–factors, either $[M:N] = 4\cos^2(\pi/q)$ for some integer $q \geq 3$ or $[M:N] \in [4,\infty]$.*

There is substantial overlap between Sections 3.2 to 3.5 and Sections I to III of Connes' report [Con].

Let us now describe Sections 5 to 7, where we consider a pair $N \subset M$ of *finite von Neumann algebras with finite dimensional centers*. There are projections $p_1,\cdots,p_m$ which are central in M and projections $q_1,\cdots,q_n$ which are central in N such that

$p_1M, \cdots, p_mM, q_1N, \cdots, q_nN$ are finite factors, and

$$N = \bigoplus_{j=1}^{n} q_jN \subset M = \bigoplus_{i=1}^{m} p_iM.$$

If $\dim_{\mathbb{C}}(M) < \infty$, this is the situation of Chapter 2. At this stage, let us assume that each of the factors p_iM, q_jN is of type II_1 (see the comment after 3.5.4).

As in Section 2.3, we define an <u>index matrix</u> $\Lambda_N^M = (\lambda_{i,j}) \in \mathrm{Mat}_{m,n}(\mathbb{R}_+ \cup \{\infty\})$ by

$$\lambda_{i,j} = \begin{cases} 0 & \text{if } p_i q_j = 0 \\ [M_{i,j} : N_{i,j}]^{1/2} & \text{if not} \end{cases}$$

where $N_{i,j} = p_iq_jN$ is a subfactor of the factor $M_{i,j} = p_iq_jMp_iq_j$. We say that N is of <u>finite index</u> in M if Λ_N^M does not have any infinite entry.

For an analysis of traces on M and N (see Section 2.4 when $\dim_{\mathbb{C}}(M)$ is finite), we define also the <u>trace matrix</u> $T_N^M = (c_{i,j}) \in \mathrm{Mat}_{m,n}(\mathbb{R}_+)$ by

$$c_{i,j} = tr_{p_iM}(p_iq_j),$$

where tr_{p_iM} denotes the normalized trace on the factor p_iM. A trace on M is described by the vector $\vec{s} \in (\mathbb{R}_+)^m$ with $s_i = tr(p_i)$, and its restriction to N by the vector $\vec{s}T_N^M \in (\mathbb{R}_+)^n$. Traces are always assumed to be positive in this chapter, so that $s_i \geq 0$ for $i = 1, \cdots, m$.

If $\dim_{\mathbb{C}}(M) < \infty$ the matrices Λ_N^M and T_N^M are simply related by $c_{i,j} = \lambda_{i,j}\nu_j\mu_i^{-1}$, with $\mu_i^2 = \dim_{\mathbb{C}}(p_iM)$ and $\nu_j^2 = \dim_{\mathbb{C}}(q_jN)$. This relation has no analogue when the p_iM's and the q_jN's are factors of type II_1:

<u>Proposition 3.1.3.</u> *Consider two irredundant matrices*

$$\Lambda = (\lambda_{i,j}) \in \mathrm{Mat}_{m,n}(\{2\cos(\pi/q)\}_{q \geq 2} \cup [2,\infty]) \quad \textit{and} \quad T = (c_{i,j}) \in \mathrm{Mat}_{m,n}(\mathbb{R}_+)$$

satisfying:

$$\lambda_{i,j} = 0 \Leftrightarrow c_{i,j} = 0 \text{ and } \sum_{1 \leq j \leq n} c_{i,j} = 1 \text{ for } i \in \{1, \cdots, m\}.$$

Then there exists a pair $N \subset M$ *as above with* $\Lambda = \Lambda_N^M$ *and* $T = T_N^M$. *Moreover,* M *and* N *may be chosen hyperfinite.*

The Skolem–Noether theorem does not hold for II_1-factors and Proposition 2.3.3 does not carry over to the present setting: the matrices Λ_N^M and T_N^M do not characterize N as a subalgebra of M.

Once a faithful trace is given on M, the fundamental construction gives a new algebra $\langle M, e_N \rangle$, just as described above in the case of factors.

<u>Proposition 3.1.4.</u> *Let* N *be of finite index in* M.

(a) *The algebras* $\langle M, e_N \rangle$ *and* $\mathrm{End}_N^r(M)$ *are isomorphic.*

(b) *The algebra* $\langle M, e_N \rangle$ *is again a finite sum of* II_1 *factors. There is a natural bijection between the minimal central idempotents of* N *and those of* $\langle M, e_N \rangle$.

A convenient isomorphism is described in Corollary 3.6.5, and the bijection of (b) appears in Proposition 3.6.1.iv.

The partial description of $N \subset M$ by Λ_N^M and T_N^M is useful because, if N is of finite index in M and if $L = \langle M, e_N \rangle$, one may compute Λ_M^L and T_M^L from Λ_N^M and T_N^M. Indeed

$$\Lambda_M^L = (\Lambda_N^M)^t \qquad T_M^L = F_N^M \tilde{T}_N^M$$

where $\tilde{T} = \tilde{T}_N^M \in \mathrm{Mat}_{n,m}(\mathbb{R}_+)$ is defined by

$$\tilde{T}_{j,i} = \begin{cases} 0 & \text{if } p_iq_j = 0 \\ \lambda_{i,j}^2 c_{i,j}^{-1} & \text{if not}, \end{cases}$$

and where F_N^M is a diagonal matrix ensuring that $\sum_{1 \leq i \leq m} (T_M^L)_{j,i} = 1$ for $j = 1, \cdots, n$.

See Propositions 3.6.6 and 3.6.8 for the details.

As in Chapter 2, a trace on M is said to be a <u>Markov trace</u> of <u>modulus</u> β for the pair $N \subset M$ if it extends to a trace tr on $\langle M, e_N \rangle$ for which

$$\beta\, tr(xe_N) = tr(x) \qquad x \in M.$$

There exists at most one such extension. As traces are positive in this chapter, β has to be a positive number. The analogues of Theorem 2.1.3 and 2.1.4 hold as follows. Recall that a pair $N \subset M$ is <u>connected</u> if the intersection $Z(M) \cap Z(N)$ of the centers is reduced to $\mathbb{C}1$.

<u>Theorem 3.1.5</u>. *Let* $M = \bigoplus_{i=1}^{m} p_i M$ *and* $N = \bigoplus_{j=1}^{n} q_j N$ *be finite direct sums of* II_1 *factors, let* N *be a subalgebra of* M *of finite index, and write* $T, \tilde{T}$ *for* $T_N^M, \tilde{T}_N^M$.

(a) *Let* $tr : M \to \mathbb{C}$ *be a trace, let* $\vec{s} \in \mathbb{R}_+^m$ *be defined by* $s_i = tr(p_i)$, *and let* $\beta \in \mathbb{R}_+^*$. *Then* tr *is a Markov trace of modulus* β *for the pair* $N \subset M$ *if and only if*

$$\vec{s}\, T\tilde{T} = \beta \vec{s}.$$

(b) *If the conditions of (a) hold, then the Markov extension* $\langle M, e_N \rangle \to \mathbb{C}$ *of* tr *is a Markov trace of modulus* β *for the pair* $M \subset \langle M, e_N \rangle$.

(c) *If* $N \subset M$ *is connected, there exists a unique normalized Markov trace on* $N \subset M$, *and its modulus* β *is the spectral radius of* $T\tilde{T}$.

Comparing Theorems 2.1.4 and 3.1.5, we may define the <u>index</u> of N in M as

$$[M:N] = \rho(T\tilde{T})$$

where ρ denotes spectral radius.

<u>Corollary 3.1.6</u>. *Theorem 3.1.2 holds for finite direct sums of* II_1 *factors.*

We note that the definition of $[M:N]$ given above is not the same as that of Chapter 2. However, P. Jollissaint has shown, in unpublished work, that the two definitions of index coincide.

If $N \subset M$ is a connected pair of finite dimensional multi-matrix algebras with $[M:N] \leq 4$, we have shown in Theorems 2.1.1 and 1.1.2 that the corresponding graph is a Coxeter graph of one of the types A,D,E. The chief result of Section 3.7 is that connected pairs $N \subset M$ of finite direct sums of II_1-factors with $[M:N] \leq 4$ give rise to <u>all</u> possible Coxeter graphs associated with finite and affine groups.

<u>Theorem 3.1.7</u>. *Let* $N \subset M$ *be a connected pair of finite direct sums of* II_1*-factors. Assume that* N *is of finite index in* M *and let* $\Lambda = \Lambda_N^M$ *be the inclusion matrix.*

(a) *If* $[M:N] < 4$, *then* Λ *is the matrix associated (in Theorem 1.1.3) to a bicoloration of one of the following Coxeter graphs:*

$$A_\ell\ (\ell \geq 2),\ B_\ell\ (\ell \geq 2),\ D_\ell\ (\ell \geq 4),\ E_\ell\ (\ell = 6,7,8),$$
$$F_4,\ G_2,\ H_\ell\ (\ell = 3,4),\ I_2(p)\ (p = 5\ \textit{or}\ p \geq 7).$$

Moreover $[M:N] = \|\Lambda\|^2 = 4\cos^2(\pi/h)$, *where* h *is the Coxeter number. (See tables 1.4.5 and 1.4.7.)*

(b) *If* $[M:N] = 4$, then Λ_N^M *corresponds to one of the graphs:*

$$A_\ell^{(1)}\ (\ell \text{ odd}, \ell \geq 1),\ B_\ell^{(1)}\ (\ell \geq 2),\ C_\ell^{(1)}\ (\ell \geq 3),$$
$$D_\ell^{(1)}\ (\ell \geq 4),\ E_\ell^{(1)}\ (\ell = 6,7,8),\ F_4^{(1)},\ G_2^{(1)}.$$

so that $[M:N] = \|\Lambda\|^2$. *(See tables 1.4.6 and 1.4.7.).*

The index range described by Theorem 3.1.2 appears also in the remarkable family of <u>Hecke groups</u>, which are discrete subgroups of $PSL(2,\mathbb{R})$ generated by two parabolic transformations. We have included an Appendix III on these groups. Its purpose is to expose the spectacular comparison with Theorem 3.1.2 as well as to illustrate Section 3.3.

3.2. The coupling constant: definition.

Let H be a (complex) Hilbert space. We denote by $B(H)$ the $*$-algebra of all bounded operators on H, with x^* the adjoint of the operator $x \in B(H)$. Besides the topology associated to the norm

$$\|x\| = \operatorname{Sup}\{\|x\xi\| : \xi \in H \text{ and } \|\xi\| \leq 1\}$$

the algebra has also the <u>ultraweak topology</u> or w-topology which is defined by the semi-norms

$$x \mapsto \Big|\sum_{j=1}^{\infty} \langle x\xi_j | \eta_j \rangle\Big| \qquad \begin{array}{l} \xi_1, \xi_2, \cdots \in H \text{ with } \sum_{j=1}^{\infty} \|\xi_j\|^2 < \infty \\ \eta_1, \eta_2, \cdots \in H \text{ with } \sum_{j=1}^{\infty} \|\eta_j\|^2 < \infty. \end{array}$$

Whenever necessary, we assume H to be separable.

A von Neumann algebra acting on H, or a <u>von Neumann subalgebra</u> of $B(H)$, is a w-closed $*$-subalgebra of $B(H)$ which contains the identity. If M_j is a von Neumann subalgebra of $B(H_j)$ for $j = 1,2$ and if $\varphi : M_1 \to M_2$ is a $*$-isomorphism, it is known that φ is continuous with respect to the w–topology on both M_1 and M_2 (corollary 5.13 in [SZ] or section I.4.3 in [DvN]). A von Neumann algebra is a $*$–algebra M which is $*$–isomorphic to a von Neumann subalgebra of $B(H)$ for some H; by the result just recalled, such an algebra has a well–defined w–topology.

A factor is a von Neumann algebra M with center Z_M reduced to the scalar multiples of the identity. Von Neumann algebras are known to be principal in the sense that any w–closed two–sided ideal is generated by a central projection (see section I.3.4 in [DvN]). Thus a von Neumann algebra M is a factor if and only if any two–sided ideal $J \neq 0$ in M is w–dense. There is not any continuity problem for representations of a factor M in the following sense: any $*$–homomorphism $M \to B(H)$ is w–continuous. (See theorem V.5.1 in [Tak]; the separability of H is crucial here.)

A II_1 factor is an infinite dimensional factor M which admits a normalized finite trace $tr : M \to \mathbb{C}$ such that

(i) $tr(\mathbf{1}) = 1$

(ii) $tr(xy) = tr(yx)$ $x,y \in M$

(iii) $tr(x^*x) \geq 0$ $x \in M$.

It is known that, on a II_1–factor, such a trace is unique in two senses. First, in the usual sense for operator algebras: tr is the unique linear form satisfying (i), (ii) and (iii); see [DvN], n^{os} I.6.4 and III.2.7; moreoever one has $tr(x^*x) > 0$ for $x \neq 0$. But also secondly, in the naive sense: tr is the unique linear form satisfying (i) and (ii), by [FH]. The existence of II_1–factors which may act on separable Hilbert spaces is one of the basic discoveries in the first paper by Murray and von Neumann [MvN I].

A finite factor is a von Neumann algebra which is either a II_1–factor, or isomorphic to $B(H)$ for some H of finite dimension. Such a factor is simple as a complex algebra by [DvN], III.5.2. Here is a characterization of finite factors; for more on this, see [KvN].

Proposition 3.2.1. *Let M be a C^*–algebra with unit and with center reduced to the scalar multiples of $\mathbf{1}$. Let $tr : M \to \mathbb{C}$ be a faithful normalized trace (namely a linear form satisfying (i), (ii), (iii) above and $tr(x^*x) > 0$ for $x \neq 0$). Assume that the unit ball of M is complete with respect to the metric $d(x,y) = \|x-y\|_2$, where $\|x\|_2 = (tr(x^*x))^{1/2}$. Then M is a finite factor.*

Proof. Let $H = L^2(M,tr)$ be the Hilbert space obtained by completion of M with respect to the scalar product defined by $<x|y> = tr(x^*y)$ for $x,y \in M$. Let $\pi : M \to B(H)$ be the $*$–representation of M on H, with $\pi(x)$ being the extension to H of the left multiplication by x on M. Then π is injective because tr is faithful. Let $\pi(M)''$ denote the double commutant of M in $B(H)$, which is, by von Neumann's bicommutant theorem, the w–closure of $\pi(M)$ in $B(H)$.

To show that M is a von Neumann algebra, it is enough to show that the inclusion of M in $\pi(M)''$ is surjective. Let $a \in \pi(M)''$ with $\|a\| = 1$. By Kaplansky's density theorem, there is a net (x_α) in M with $\|x_\alpha\| \leq 1$ for all α such that $\pi(x_\alpha)$ converges strongly to a; that is, $\pi(x_\alpha)\xi$ converges to $a\xi$ for all $\xi \in H$. Taking $\xi = 1$, this means that $(\pi(x_\alpha))$ is a Cauchy net for the $\|\cdot\|_2$–distance, so converges with respect to this distance to some element $\pi(x_0)$ by the assumed completeness of the ball of M. One can check that the strong topology and the $\|\cdot\|_2$–topology coincide on the unit ball of $\pi(M)''$, so $a = \pi(x_0) \in \pi(M)$. #

Let M be a finite factor acting on some Hilbert space H. We are going to define the coupling constant $\dim_M(H)$ which is a measure of the size of H as an M–module, the definition being made so that the standard M–module $L^2(M) = L^2(M,tr)$ has size 1. Before comparing other M–modules to that one, we recall the following facts.

Lemma 3.2.2. (a) *Let $J : L^2(M) \to L^2(M)$ be the conjugate linear isometry which extends $\begin{cases} M \to M \\ x \mapsto x^* \end{cases}$. Then JMJ is the commutant $End_M(L^2(M))$ of M in $B(L^2(M))$.*

(b) *Let K be a Hilbert space and let M act on $L^2(M) \otimes K$ by the diagonal action $x(\eta \otimes \theta) = (x\eta) \otimes \theta$. Then $JMJ \otimes B(K)$ is the commutant of M in $B(L^2(M) \otimes K)$.*

(c) *Assume that the space K of (b) is infinite dimensional. For any M–module H, there exists an isometry*

$$u : H \to L^2(M) \otimes K$$

which is M–linear, namely which intertwines the actions of M.

Proof. (a) Let $x,y,z \in M$. By definition of J

$$JxJy = (xy^*)^* = yx^* = yJx.$$

Applying this twice we get

$$JxJyz = yzJx = yJxJz,$$

and setting $z = \mathbf{1}$,

$$(JxJ)y = y(JxJ).$$

Thus $JMJ \subset M'$ where $M' = End_M(L^2(M))$.

Let moreover $a \in M'$. By definition of the adjoint

$$\langle y^*x^* | a \rangle = \langle x^* | ya \rangle = \langle x^* | ay \rangle = \langle a^*x^* | y \rangle = \langle x^*a^* | y \rangle = \langle a^* | xy \rangle.$$

Now one has $\langle J\eta|\theta\rangle = \overline{\langle\eta|J\theta\rangle}$ for all $\eta,\theta \in L^2(M)$, and consequently

$$\overline{\langle xy|Ja\rangle} = \langle y^*x^*|a\rangle = \overline{\langle xy|a^*\rangle}$$

so that $Ja = a^*$. Thus the first computation shows also that $JM'J \subset M''$ and, taking adjoints, $M' \subset JM''J$.

By von Neumann's bicommutant theorem, one has $M' = JMJ$.

(b) Let $x \in B(L^2(M) \otimes K)$. Choose an orthonormal basis $(\epsilon_i)_{i\in I}$ of K, and represent x by a matrix $(x_{i,j})_{i,j\in I}$ over $B(L^2(M))$. If x commutes to the action of M, this matrix must have entries in $\mathrm{End}_M(L^2(M))$, and thus $x \in \mathrm{End}_M(L^2(M)) \otimes B(K)$. Conversely any bounded matrix $(x_{i,j})$ with entries in $\mathrm{End}_M(L^2(M))$ commutes with the diagonal action of M.

(c) Consider $H \oplus (L^2(M) \otimes K)$ as an M-module for the diagonal action $x(\zeta \oplus (\eta \otimes \theta)) = x\zeta \oplus (x\eta \otimes \theta)$. Then $0 \oplus 1$ is an infinite projection in the commutant of M. By the Murray- von Neumann comparison theory for projections, there exists a partial isometry $\tilde{u}$ in the commutant $\mathrm{End}_M(H \oplus (L^2(M) \otimes K))$ from $1 \oplus 0$ to a subprojection of $0 \oplus 1$. One may view $\tilde{u}$ as an isometry

$$u : H \to L^2(M) \otimes K$$

which intertwines the actions. #

As there will be many traces with various normalizations in the sequel, we introduce the following convention. If M is a finite factor, tr_M will denote its <u>normalized trace</u>. So if Tr is any other trace on M, then $Tr = Tr(1)tr_M$, a formula which we will use often. Occasionally, we will have to consider a trace Tr on a factor P which is not finite (for example $B(H)$ or $M \otimes B(H)$, with H of infinite dimension). Let P_+ denote the positive cone of P, consisting of those element of the form z^*z with $z \in P$. Then a trace Tr is a map $P_+ \to [0,\infty]$ such that

(i) $Tr(x+y) = Tr(x) + Tr(y) \quad x,y \in P_+$

(ii) $Tr(\lambda x) = \lambda Tr(x) \quad \lambda \in \mathbb{R}_+, x \in P_+$ (with $0\cdot\infty = 0$)

(iii) $Tr(uxu^*) = Tr(x) \quad x \in P_+$, u a unitary in P.

Given a finite factor M acting in a Hilbert space H as in Lemma 3.2.2, we define now the <u>natural trace</u> $Tr_{M'}$ on its commutant. It is crucial for what follows that $Tr_{M'}$ is *not* necessarily normalized.

First, if $H = L^2(M)$ as in (a), we define $Tr_{M'}(JxJ) = tr_M(x)$ for all $x \in M$; in this case, $Tr_{M'}$ is normalized. Secondly, if $H = L^2(M) \otimes K$ as in (b), consider an

orthonormal basis $(\epsilon_i)_{i\in I}$ of K; then any element x in the commutant $\mathrm{End}_M(L^2(M) \otimes K)$ is represented by a matrix $(Jx_{i,j}J)_{i,j\in I}$; when x is moreover positive, then the diagonal elements $x_{i,i}$ are also positive, and we define

$$Tr_{M'}(x) = \sum_{i\in I} tr_M(x_{i,i}) \in [0,\infty].$$

For example,

$$Tr_{M'}(JxJ \otimes p) = tr_M(x)\dim_{\mathbb{C}}(pK)$$

if $x \in M_+$ and if $p \in B(K)$ is a projection.

Let $\mathcal{F}(K)$ denote the finite-rank operators on K. If $x \in JMJ \otimes \mathcal{F}(K) \subset \mathrm{End}_M(L^2(M) \otimes K)$, that is if all but finitely many of the matrix entries $x_{i,j}$ are zero, but x is not necessarily positive, then $Tr_{M'}(x)$ is well-defined by the same formula. Furthermore, $x \longmapsto Tr_{M'}(x)$ is a positive trace on the *-algebra $JMJ \otimes \mathcal{F}(K)$.

Third, for H arbitrary and for u as in (c) of Lemma 3.2.2, we define

$$Tr_{M'}(x) = Tr_{M'}(uxu^*)$$

for $x \in \mathrm{End}_M(H)_+$, and thus $uxu^* \in \mathrm{End}_M(L^2(M) \otimes K)_+$. If u_1,u_2 are two possible choices for u, then $u_1^*u_1 = u_2^*u_2 = id_H$ and $u_2xu_2^* = u_2u_1^*u_1xu_2^*$ for $x \in M$; as $Tr_{M'}$ is a trace,

$$Tr_{M'}(u_2xu_2^*) = Tr_{M'}(u_1xu_2^*u_2u_1^*) = Tr_{M'}(u_1xu_1^*)$$

and $Tr_{M'}(x)$ does not depend on the choice of u.

The word "natural" is justified by the following property (which again shows the independence just observed).

<u>Lemma 3.2.3</u>. *Let H_1,H_2 be two M-modules; let $a : H_1 \to H_2$ and $b : H_2 \to H_1$ be two M-linear bounded operators. Denote by T_j the natural trace defined on $\mathrm{End}_M(H_j)$ as above, for $j = 1,2$. Then*

$$T_2(ab) = T_1(ba).$$

Proof. Let $u_j : H_j \to L^2(M) \otimes K$ be an M-linear isometry. Then

$$\begin{aligned} T_2(ab) &= Tr_{M'}(u_2abu_2^*) = Tr_{M'}(u_2au_1^*u_1bu_2^*) \\ &= Tr_{M'}(u_1bu_2^*u_2au_1^*) = T_1(ba). \quad \# \end{aligned}$$

Definition. Let M be a finite factor and let H be a M-module. The coupling constant $\dim_M(H)$ is defined to be $Tr_{M'}(id_H)$, where the natural trace $Tr : End_M(H)_+ \to [0,\infty]$ is defined as above. If u is as in 3.2.2.c, one has also $\dim_M(H) = Tr_{M'}(uu^*)$ by 3.2.3.

Proposition 3.2.4. *Let M be a finite factor and let $H, H', H_1, H_2, \cdots$ be M-modules which are separable as Hilbert spaces. Then*

(a) $\dim_M(H) = \dim_M(H')$ *if and only if H and H' are isomorphic as M-modules,*

(b) $\dim_M(\underset{i}{\oplus} H_i) = \sum_i \dim_M(H_i)$,

(c) $\dim_M(L^2(M)) = 1$,

(d) $\dim_M(H) < \infty$ *if and only if the factor $End_M(H)$ is finite.*

Proof. Claim (a) follows from the comparison theorem for projections in the factor $End_M(L^2(M) \otimes K)$, claim (b) from the σ-additivity of the trace $Tr_{M'}$ on the same factor, and (c) is obvious.

In all cases, $End_M(H)$ is a semi-finite factor, and thus admits a non-zero trace which is unique up to a multiplicative constant. Claim (d) holds because $End_M(H)$ is finite if and only if it has a finite trace. #

In the next proposition, we continue with properties of $\dim_M$. The deep result is (f). We now describe the main step, the proof of which is in [MvN I] and [MvN IV] (see Theorem X in both papers). Again, let M be a finite factor and let H be a M-module; let tr be the normalized trace on M and let Tr' be the natural trace on $End_M(H)$. Choose $\xi \in H$ with $\xi \neq 0$. Denote by e_ξ the orthogonal projection of H onto the closure of the cyclic module $End_M(H)\xi$, and by e'_ξ that onto $\overline{M\xi}$; observe that $e_\xi \in M$ and $e'_\xi \in End_M(H)$. The basic (and difficult) fact is that the ratio

$$\tilde{c}_M = tr(e_\xi)/Tr'(e'_\xi)$$

is independent of ξ. (When M and H are finite dimensional, this basic fact reduces to Proposition 2.2.7.) Murray and von Neumann define the coupling constant of M and $End_M(H)$ to be

$$c_M = tr(e_\xi)/tr'(e'_\xi) = (Tr'(\mathbf{1}))\tilde{c}_M \in \mathbb{R}_+^*$$

if $End_M(H)$ is finite, with tr' the normalized trace on $End_M(H)$ and Tr' the natural trace. In case $End_M(H)$ is infinite, they define $c_M = +\infty$.

The M-module H gives rise to other modules as follows. Let $e \in B(H)$ be a projection $(e \neq 0)$, with range denoted by eH. If $e \in End_M(H)$ then eH is naturally a M-module (a submodule of H); if moreover $End_M(H)$ is finite, the value $D(e)$ of the normalized trace of $End_M(H)$ on e is called the dimension of e. On the other hand, if $e \in M$, then eH is a eMe-module; the algebra eMe is a finite factor (because it is simple, a fact easy to check) which is called the reduction of M by e. Following common practice, we also write M_e for eMe.

Proposition 3.2.5. *Let M be a finite factor and let H be a M-module. Assume that the factor $End_M(H)$ is finite (namely that $\dim_M(H) < \infty$). Then*

(e) $\dim_M(eH) = D(e)\dim_M(H)$ *for any non-zero projection* $e \in End_M(H)$.

(f) $\dim_M(H) = c_M$, *the coupling constant of Murray and von Neumann.*

(g) $\dim_M(H)\dim_{End_M(H)}(H) = 1$,

(h) $\dim_{eMe}(eH) = \frac{1}{D(e)}\dim_M(H)$ *for any non-zero projection* $e \in M$, *where* $D(e) = tr(e)$.

(i) *If L is a finite dimensional Hilbert space, then* $\dim_M(H \otimes L) = \dim_M(H)\dim_{\mathbb{C}}(L)$.

Proof. For (e), one may view e as an M-linear isometry from eH to H. Then if $u : H \to L^2(M) \otimes K$ is an M-linear isometry, we have by definition of $\dim_M(\cdot)$ and by Lemma 3.2.3

$$\begin{aligned} \dim_M(eH) &= Tr_{End_M(L^2(M)\otimes K)}(ueu^*) \\ &= Tr_{End_M(H)}(e) \end{aligned}$$

$$= \mathrm{Tr}_{\mathrm{End}_M(H)}(\mathrm{id}_H)\mathrm{tr}_{M'}(e)$$
$$= \dim_H(H)D(e),$$

where each Tr_* denotes a natural trace.

Next we show how (f) reduces to the result of Murray and von Neumann recalled above. Replacing H by an isomophic submodule of $L^2(M) \otimes K$, we can assume $H \subset L^2(M) \otimes K$. Let $p \in \mathrm{End}_M(L^2(M) \otimes K)$ denote the orthogonal projection from $L^2(M) \otimes K$ onto H. Then by definition

(3.2.5.1) $$\mathrm{Tr}_{\mathrm{End}_M(L^2(M)\otimes K)}(p) = \dim_M(H)$$

Let $\xi \in H$ with $\xi \neq 0$ and let $\eta \in L^2(M) \otimes K$ with $\eta \neq 0$. As earlier, denote by $e_\xi \in M$ and $e'_\xi \in \mathrm{End}_M(H)$ the projections of H onto $\overline{\mathrm{End}_M(H)\xi}$ and $\overline{M\xi}$. Likewise denote by $f_\eta \in M$ and $f'_\eta \in \mathrm{End}_M(L^2(M) \otimes K)$ the projections of $L^2(M) \otimes K$ onto $\overline{\mathrm{End}_M(L^2(M)\otimes K)\eta}$ and $\overline{M\eta}$. With respect to the orthogonal decomposition $L^2(M) \otimes K = H \oplus H^\perp$, the algebra M acts by operators of the form $\begin{bmatrix} * & 0 \\ 0 & * \end{bmatrix}$, the algebra $\mathrm{End}_M(L^2(M) \otimes K)$ is of the form $\begin{bmatrix} \mathrm{End}_M(H) & * \\ * & * \end{bmatrix}$ and the space $\overline{\mathrm{End}_M(L^2(M) \otimes K)\xi}$ is of the form $\begin{bmatrix} \overline{\mathrm{End}_M(H)\xi} \\ * \end{bmatrix}$. It follows that $pf_\xi = e_\xi p$, or in matrix form that $f_\xi = \begin{bmatrix} e_\xi & 0 \\ 0 & * \end{bmatrix}$, so that it is the same element in M which acts as f_ξ on $L^2(M) \otimes K$ and as e_ξ on H. Consequently

(3.2.5.2) $$\mathrm{tr}_M(f_\xi) = \mathrm{tr}_M(e_\xi).$$

Observe also that, more simply

(3.2.5.3) $$pf'_\xi p = f'_\xi p = e'_\xi$$

because $\overline{M\xi} \subset H$.

To compute $\tilde{c}_2 = \mathrm{tr}_M(f_\eta) \div \mathrm{Tr}_{\mathrm{End}_M(L^2(M)\otimes K)}(f'_\eta)$, we may choose $\eta = \mathbf{1} \otimes \kappa$ with $\mathbf{1} \in M \subset L^2(M)$ and $\kappa \neq 0$ in K. Then f_η is the identity on $L^2(M) \otimes K$ and f'_η is the projection onto $L^2(M) \otimes \mathbb{C}\kappa$. Consequently

$$\mathrm{tr}_M(f_\eta) = 1 = \mathrm{Tr}_{\mathrm{End}_M(L^2(M)\otimes K)}(f'_\eta)$$

and $\tilde{c}_2 = 1$. But $\tilde{c}_2$ can also be computed using $\xi \in H$, so one has

(3.2.5.4) $$\mathrm{tr}_M(f_\xi) = \mathrm{Tr}_{\mathrm{End}_M(L^2(M)\otimes K)}(f'_\xi).$$

The coupling constant of Murray and von Neumann for M and $\mathrm{End}_M(H)$ is

$$c_M = \mathrm{tr}_M(e_\xi)/\mathrm{tr}_{\mathrm{End}_M(H)}(e'_\xi),$$

since we are assuming that $\mathrm{End}_M(H)$ is finite. By uniqueness of the normalized trace on $\mathrm{End}_M(H)$, one has

(3.2.5.5) $$\mathrm{tr}_{\mathrm{End}_M(H)}(pxp) = \mathrm{Tr}_{\mathrm{End}_M(L^2(M)\otimes K)}(x) \div \mathrm{Tr}_{\mathrm{End}_M(L^2(M)\otimes K)}(p)$$

for any $x \in \mathrm{End}_M(L^2(M)\otimes K)$. Putting together (3.2.5.1) to (3.2.5.5) one obtains

$$\begin{aligned} c_M &= \mathrm{tr}_M(e_\xi)/\mathrm{tr}_{\mathrm{End}_M(H)}(e'_\xi) \\ &= \{\mathrm{tr}_M(e_\xi)/\mathrm{Tr}_{\mathrm{End}_M(L^2(M)\otimes K)}(f'_\xi)\}\ \mathrm{Tr}_{\mathrm{End}_M(L^2(M)\otimes K)}(p) \\ &= \{\mathrm{tr}_M(f_\xi)/\mathrm{Tr}_{\mathrm{End}_M(L^2(M)\otimes K)}(f'_\xi)\}\ \dim_M(H) \\ &= \dim_M(H), \end{aligned}$$

and claim (f) is proved.

Claim (g) now follows trivially from (f). As for (h), using (e) and (g) as well as $\mathrm{End}_{eMe}(eH) = e(\mathrm{End}_M(H))e$, we have

$$\begin{aligned} \{\dim_{eMe}(eH)\}^{-1} &= \dim_{\mathrm{End}_{eMe}(eH)}(eH) = D(e)\dim_{\mathrm{End}_M(H)}(H) \\ &= D(e)\{\dim_M(H)\}^{-1}. \end{aligned}$$

Point (i) follows easily from the definition of $\dim_M(\cdot)$.

This ends the proof of Proposition 3.2.5. #

If $M = \mathrm{Mat}_\mu(\mathbb{C})$ for some integer $\mu \geq 1$, then $\dim_M(H) = \mu^{-2}\dim_{\mathbb{C}}(H)$ is of the form $\frac{d}{\mu}$ with d an integer as in Proposition 2.2.7. This follows for example from claims (b) and (c) of Proposition 3.2.4. The object of the next section is to describe examples involving factors of type II_1.

3.3. The coupling constant: examples.

The situation for which the coupling constant is computed in this section is of the following kind: G is a non-compact semi-simple connected real Lie group which has the same rank as its maximal compact subgroups, $\pi : G \to U(H)$ is an irreducible representation of G in the discrete series, and $M = W^*(\Gamma)$ is the von Neumann algebra of an appropriate discrete subgroup Γ of G. Then H is naturally an M-module. Theorem 3.3.2 below is a computation of $\dim_M(H)$, due to Atiyah-Schmidt [AS,(3.3)]. First we discuss some background; the knowledgeable reader should jump to Theorem 3.3.2.

3.3.a. Discrete series.

Let G be a locally compact group. We assume that G is unimodular, we choose a Haar measure dg on G, and we denote by $\lambda_G : G \to U(L^2(G,dg))$ the left regular representation of G.

For an irreducible unitary representation $\pi : G \to U(H)$ of G, the following properties are equivalent:

(i) π is a subrepresentation of λ_G; more precisely, there exists a projection p in the commutant of $\lambda_G(G)$ such that the restriction of λ_G to the range of p is equivalent to π;

(ii) There exist $\xi,\eta \in H - \{0\}$ such that $g \mapsto \langle\pi(g)\xi|\eta\rangle$ is in $L^2(G,dg)$;

(iii) For all $\xi,\eta \in H$ the function $g \mapsto \langle\pi(g)\xi|\eta\rangle$ is in $L^2(G,dg)$.

If these hold, π is said to belong to the (unitary) discrete series. On may then attach to π a real number $d_\pi > 0$, called its formal dimension, such that Schur's orthogonality relations formally hold. In particular, for any $\pi : G \to U(H)$ in the discrete series

$$\int_G \langle\pi(g)\xi|\eta\rangle\langle\eta'|\pi(g)\xi'\rangle dg = d_\pi^{-1}\langle\xi|\xi'\rangle\langle\eta'|\eta\rangle$$

for $\xi,\xi',\eta,\eta' \in H$.

The formal dimension d_π depends on π *and* on the choice of the Haar measure for G; if $d'g = k\,dg$ for some constant $k > 0$, the two corresponding formal dimensions of π are related by $d'_\pi = k^{-1}d_\pi$. If G is compact and if $\int_G dg = 1$, then d_π is the dimension of H in the naive sense. For all this, see section 16 in [Rbt] or Chapter 14 in [DC*].

Given an arbitrary (unimodular) group G, its discrete series may be empty. This happens for G infinite abelian, or infinite discrete, or $G = SL(2,\mathbb{C})$, or $G = SL(n,\mathbb{R})$ with $n \geq 3$, to quote but a few examples. When G is a semi-simple connected real Lie group with maximal compact subgroup K, then G has discrete series representations if and only if G and K have the same rank. In particular $SL(2,\mathbb{R})$ has a discrete series, as well as $SO(n,1)^0$ for n even.

3.3.b. Factors defined by icc groups.

On a discrete group Γ, we consider always the counting measure; the space of square summable functions from Γ to $\mathbb{C}$ is denoted by $\ell^2(\Gamma)$. The von Neumann algebra $W^*(\Gamma)$ of Γ is the (ultra)weak closure of the linear span of $\lambda_\Gamma(\Gamma)$ in $B(\ell^2(\Gamma))$; by von Neumann's theorem, it is also the bicommutant of $\lambda_\Gamma(\Gamma)$ in $B(\ell^2(\Gamma))$, and $W^*(\Gamma)$ is thus also denoted by $\lambda_\Gamma(\Gamma)''$.

Let $\delta_e \in \ell^2(\Gamma)$ be the function which takes the value 1 at the identity e of Γ and 0 elsewhere. It is easy to check that $x \longmapsto x(\delta_e)$ is a linear injection of $W^*(\Gamma)$ in $\ell^2(\Gamma)$, and that the map $tr(x) = \langle x(\delta_e)|\delta_e\rangle$ is a normalized finite faithful trace on $W^*(\Gamma)$; see the end of 4.2 in [Sak]. It follows that the von Neumann algebra $W^*(\Gamma)$ is finite, and that the Hilbert space $L^2(W^*(\Gamma),tr)$ defined before Lemma 3.2.2 is canonically isomorphic to $\ell^2(\Gamma)$.

Moreover $W^*(\Gamma)$ is a factor (and thus a factor of type II_1) if and only if Γ is an infinite conjugacy class group, or for short an icc group (Lemma 4.2.18 in [Sak]). The following lemma exhibits a rich class of icc groups. Before this, we recall that the quotient G/Γ of a unimodular locally compact group G by a discrete subgroup Γ has always a G-invariant measure, which is unique up to a scalar factor; by definition, Γ is a lattice in G if the meaure of G/Γ is finite.

Lemma 3.3.1. *A lattice* Γ *in a connected semi-simple real Lie group* G *without center and without a compact factor is an icc group, and* $W^*(\Gamma)$ *is consequently a* II_1*-factor.*

Proof. The main point is Borel's density theorem, which we quote without proof (see [Bor] or [Zim]): Γ is Zariski-dense in G.

Consider $h \in \Gamma$ and its conjugacy class C_h in Γ. The map $\begin{cases} \Gamma \to C_h \\ \gamma \mapsto \gamma h \gamma^{-1} \end{cases}$ extends by continuity to the Zariski closure $\begin{cases} G \to \overline{C_h} \\ g \mapsto ghg^{-1} \end{cases}$. If C_h is finite, then $\overline{C_h} = C_h$ and $\{g \in G \mid gh = hg\}$ is a closed subgroup of finite index in G. But the algebraic group corresponding to G is Zariski- connected, and it follows that $\{g \in G \mid gh = hg\} = G$. Thus h is central in G, so that $h = e$. This shows that Γ is an icc group. #

A final remark about this: let $\Gamma_1 \subset G_1$ and $\Gamma_2 \subset G_2$ be two examples of the situation in the previous lemma. Assume moreover that G_1 and G_2 have real rank at least two. It is a conjecture, due to A. Connes and "beyond Mostow and Margulis", that $W^*(\Gamma_1)$ is isomorphic to $W^*(\Gamma_2)$ if and only if Γ_1 and Γ_2 are isomorphic.

3.3.c. $W^*(\Gamma)$-modules associated to subrepresentations of λ_G.

Let G be a unimodular Lie group with Haar measure dg and let Γ be a discrete subgroup of G. In the present context, it is convenient to define a fundamental domain for Γ in G to be a subset D of G which is measurable and satisfies

$\gamma_1 D \cap \gamma_2 D$ has null measure for $\gamma_1, \gamma_2 \in \Gamma$ with $\gamma_1 \neq \gamma_2$ and

$G \setminus \bigcup_{\gamma \in \Gamma} \gamma D$ has null measure.

Such a D always exists. Indeed, as $G \to \Gamma\backslash G$ is a topological covering, it has a Borel section, and the image of such a Borel section is a convenient D. The measure of D does not depend on D itself and is called the covolume of Γ. (If dg is defined via a differential form Ω of maximal degree on G, there is a unique form ω on $\Gamma\backslash G$ which pulls back to Ω, and the covolume of Γ is $\int_{\Gamma\backslash G} \omega$.)

Of course, $\mathrm{covol}(\Gamma)$ does depend on the choice of the Haar measure on G. If $d'g = k\,dg$ for some constant $k > 0$, the two corresponding covolumes of Γ are related by $\mathrm{covol}'(\Gamma) = k\,\mathrm{covol}(\Gamma)$.

Given $\Gamma \subset G$ and D as above, there is an isomorphism from $L^2(G,dg)$ onto $\ell^2(\Gamma) \otimes L^2(D,dg)$ which maps φ to $\sum_{\gamma\in\Gamma} \delta_\gamma \otimes \varphi_\gamma$, where $\delta_\gamma \in \ell^2(\Gamma)$ is the characteristic function of $\{\gamma\}$ in Γ, and where $\varphi_\gamma(g) = \varphi(\gamma g)$ for $\gamma \in \Gamma, g \in D$. It follows from the definitions of λ_G and λ_Γ that the restriction $\lambda_G|_\Gamma$ to Γ of the left regular representation of G is the tensor product of λ_Γ with the trivial representation of Γ on $L^2(D,dg)$. Hence the von Neumann algebra $\lambda_G(\Gamma)'$ is isomorphic to $W^*(\Gamma) \otimes \mathbb{C} \cong W^*(\Gamma)$.

More generally, let $p \in B(L^2(G,dg))$ be a projection which commutes with $\lambda_G(\Gamma)$. Denote by H_p the range of p, by $\pi_p : \Gamma \to U(H_p)$ the corresponding subrepresentation of $\lambda_G|_\Gamma$, and by $\pi_p(\Gamma)''$ the von Neumann algebra generated by $\pi_p(\Gamma)$ in $B(H_p)$. Then the W^*-morphism $\begin{cases} \lambda_G(\Gamma)' \to \pi_p(\Gamma)' \\ x \mapsto px \end{cases}$ is obviously surjective. If Γ is moreover an icc group, then $\lambda_G(\Gamma)' \cong W^*(\Gamma)$ is a factor of type II_1 and is in particular a simple ring, so that the map $\lambda_G(\Gamma)' \to \pi_p(\Gamma)'$ is an isomorphism.

We shall particularize below to the case in which the projection p commutes with all of $\lambda_G(G)$, and defines an irreducible representation of G in the discrete series.

3.3.d The formula $\dim_M(H) = \mathrm{covol}(\Gamma)d_\pi$.

Now the relevant background has been established, and we demonstrate the main result of this section.

Theorem 3.3.2. *Let G be a connected semi-simple real Lie group with Haar measure dg, let Γ be a discrete subgroup in Γ, let M denote $W^*(\Gamma)$ and let $\pi : G \to U(H)$ be an irreducible representation in the discrete series. Assume that Γ is an icc group. Then* $\dim_M(H) = \mathrm{covol}(\Gamma)d_\pi$.

Observations. (1) Lemma 3.3.1 says that Γ is automatically an icc group in case it is a lattice in a connected simple noncompact Lie group without center.

(2) Both $\mathrm{covol}(\Gamma)$ and d_π depend on dg, but these dependences cancel out in the product.

Proof. From the discussion in 3.3.c, we may assume that H is included as an M-module in $L^2(G,dg)$. This inclusion, say u, satisfies $u^*u = \mathrm{id}_H$ and $uu^* = p$, where p is the orthogonal projection from $L^2(G,dg)$ onto H. Also, $L^2(G,dg)$ may be identified with $L^2(M) \otimes K$, where $L^2(M)$ is the canonical M-module, and where K is the trivial M-module $L^2(D,dg)$ associated to some fundamental domain D of Γ in G. Thus we have

$$\dim_M(H) = \mathrm{Tr}_{M'}(p);$$

in this proof, M' denotes the commutant of M in $L^2(G,dg)$ or in $L^2(M) \otimes K$, and $\mathrm{Tr}_{M'}$ is the natural trace on M'.

By Lemma 3.2.2.b, this commutant M' is generated by finite sums of the form $x = \sum_{\gamma\in\Gamma} \rho_\gamma \otimes a_\gamma$. For each $\gamma \in \Gamma$, the symbol ρ_γ stands for $J\lambda_\Gamma(\gamma)J \in \mathrm{End}_M(L^2(M))$ and a_γ is a finite rank operator in $B(K)$. Let $(\epsilon_n)_{n\in\mathbb{N}}$ be an orthonormal basis of K. Let $\overline{\epsilon_m} \otimes \epsilon_n$ denote the operator $\xi \mapsto \langle \epsilon_m | \xi \rangle \epsilon_n$ on K. One may write $a_\gamma = \sum_{m,n\in\mathbb{N}} a_{\gamma,m,n} \overline{\epsilon_m} \otimes \epsilon_n$, where the $a_{\gamma,m,n}$ are complex numbers. By definition of $\mathrm{Tr}_{M'}$ one has

$$\mathrm{Tr}_{M'}(\rho_\gamma \otimes a_\gamma) = \mathrm{tr}_M(\lambda_\gamma) \sum_{m\in\mathbb{N}} a_{\gamma,m,m} = \begin{cases} 0 & \text{if } \gamma \neq e \\ T_K(a_e) & \text{if } \gamma = e \end{cases}$$

where tr_M is the normalized trace on M and where T_K is the trace on $B(K)$ normalized by $T_K(\overline{\epsilon_m} \otimes \epsilon_m) = 1$ for all $m \in \mathbb{N}$. With x as above, one has consequently

$$\mathrm{Tr}_{M'}(x) = T_K(a_e).$$

Let $q : L^2(G,dg) \to K$ be the orthogonal projection given by restricting functions from G to D, and let T denote the trace on $B(L^2(G,dg))$ taking value 1 on projections of rank one. Then $T_K(y) = T(qyq)$ for $y \in B(K)_+$ or $y \in \mathcal{T}(K)$. In particular, for x of the form $x = \sum_{\gamma\in\Gamma} \rho_\gamma \otimes a_\gamma$ we have

(3.3.2.1) $$\mathrm{Tr}_{M'}(x) = T_K(a_e) = T(qa_eq) = T(qxq).$$

Finally any $x \in M'_+$ is the strong limit of an increasing net of operators of the form $\sum_{\gamma\in\Gamma} \rho_\gamma \otimes a_\gamma$; as the traces are normal, the formula (3.3.2.1) holds for all $x \in M'_+$, and in particular

$$\dim_M(H) = \mathrm{Tr}_{M'}(p) = T(qpq).$$

The right-hand term is explicitly given by

$$T(qpq) = \sum_{n\in\mathbb{N}} \langle qpq\epsilon_n | \epsilon_n \rangle = \sum_{n\in\mathbb{N}} \|p\epsilon_n\|^2.$$

Recall that δ_γ denotes the characteristic function of $\{\gamma\}$ in Γ, and that ϵ_n, which is a function on D, is also naturally a function on G (vanishing outside D). Thus the

orthonormal basis $\{\delta_\gamma \otimes \epsilon_n\}_{\gamma\in\Gamma, n\in\mathbb{N}}$ of $\ell^2(\Gamma) \otimes K$ is more conveniently viewed as the basis $\{\lambda_G(\gamma)\epsilon_n\}_{\gamma\in\Gamma,n\in\mathbb{N}}$ of $L^2(G,dg)$. Let η be a unit vector in $L^2(G,dg)$; assume that $\eta \in H$, namely that $p\eta = \eta$. For any $g \in G$ one has (writing λ instead of λ_G)

$$1 = \|\lambda(g)\eta\|^2 = \sum_{\gamma\in\Gamma} \sum_{n\in\mathbb{N}} |\langle \lambda(g)\eta | \lambda(\gamma)\epsilon_n \rangle|^2.$$

Consequently, as p commutes with $\lambda(G)$:

$$\begin{aligned} \mathrm{covol}(\Gamma) &= \int_D dg = \sum_{n\in\mathbb{N}} \sum_{\gamma\in\Gamma} \int_D |\langle \lambda(\gamma^{-1}g)p\eta | \epsilon_n \rangle|^2 dg \\ &= \sum_{n\in\mathbb{N}} \int_G |\langle p\lambda(g)\eta | \epsilon_n \rangle|^2 dg \\ &= \sum_{n\in\mathbb{N}} \int_G |\langle \lambda(g)\eta | p\epsilon_n \rangle|^2 dg. \end{aligned}$$

By Schur's relations

$$\mathrm{covol}(\Gamma) = \sum_{n\in\mathbb{N}} \frac{1}{d_\pi} \|\eta\|^2 \|p\epsilon_n\|^2 = \frac{1}{d_\pi} \dim_M(H)$$

and the proof is complete. #

Corollary 3.3.3. *In the situation of the previous theorem, Γ is a lattice if and only if the commutant of M in $B(H)$ is a finite factor.*

Proof. The last condition holds if and only if $\mathrm{covol}(\Gamma)$ is finite. #

We now particularize G to the group $PSL(2,\mathbb{R})$. For each integer $k \geq 2$, let H_k be the space of holomorphic functions on the Poincaré half-plane $\mathcal{P}$ which are square-summable for the measure $y^{k-2}dxdy$. (The open unit disc Δ in the complex plane with the corresponding measure is equally good). As G acts on $\mathcal{P}$ by fractional linear transformations, there is a natural unitary representation π_k of G in H_k. It is a standard result that H_k is an infinite dimensional Hilbert space and that π_k is an irreducible discrete series representation. (These π_k consitute the holomorphic discrete series, and the full discrete series contains a second "half", the anti-holomorphic part.)

Define the Haar measure dg on G as follows: let $\mathbf{T} = SO(2)/\{\pm 1\}$ be the maximal compact subgroup of G, such that $\begin{cases} G \to \mathcal{P} \\ g \mapsto g(i) \end{cases}$ induces a diffeomorphism $G/\mathbf{T} \cong \mathcal{P}$; then

$d\mu(z) = y^{-2}dxdy$ is a G-invariant measure on $\mathcal{P}$; if φ is a continuous function $G \to \mathbb{C}$ with compact support, set

$$\int_G \varphi(g)dg = \int_{\mathcal{P}} d\mu(z) \int_{\mathbb{T}} dt\,\varphi(gt) \qquad z = g(i)$$

where dt is the Haar measure on $\mathbb{T}$ of total measure 1.

Then the virtual dimension of π_k is known to be given by $d_k = \frac{k-1}{4\pi}$; see theorem 17.8 in [Rbt]. (Warning: under the Cayley transform $\begin{cases} \mathcal{P} \to \Delta \\ z=x+iy \mapsto \frac{z-i}{z+i} = u+iv \end{cases}$ the measure $\frac{dxdy}{y^2}$ corresponds to $\frac{4dudv}{(1-u^2-v^2)^2}$, which is 4 times that in [Rbt]! The measure chosen in the present section is that which is defined by the Riemannian structure for which $\mathcal{P}$ has constant curvature -1; the computation may be found, for example, in Section 5.10 of [Car].)

Now consider an integer $q \geq 3$, set $\lambda = 2\cos(\pi/q)$ and let Γ_λ be the Hecke subgroup of $PSL(2,\mathbb{R})$ generated by the classes modulo ± 1 of the matrices

$$\begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Then $\operatorname{covol}(\Gamma_\lambda) = \pi(1 - \frac{2}{q})$ by the Gauss-Bonet formula, because Γ_λ has a triangular fundamental domain with angles $0, \frac{\pi}{q}, \frac{\pi}{q}$ (see Appendix III).

Altogether, we have shown:

Example 3.3.4. Given integers $q \geq 3$ and $k \geq 2$, consider the II_1-factor $M = W^*(\Gamma_\lambda)$ defined by the Hecke group Γ_λ with $\lambda = 2\cos(\pi/q)$ and the Hilbert space H_k of the holomorphic discrete series of $PSL(2,\mathbb{R})$. Then H_k is a M-module of coupling constant

$$\dim_M(H_k) = \frac{k-1}{4}(1 - \frac{2}{q}).$$

3.3.e. A digression on the Peterson inner product.

We particularize further, and set $q = 3$ in example 3.3.4. That is, we consider $\Gamma = PSL(2,\mathbb{Z})$ as a discrete subgroup of $PSL(2,\mathbb{R})$.

Given an integer $p \geq 1$, recall that a cusp form of weight p is (in this situation) a holomorphic function $f : \mathcal{P} \to \mathbb{C}$ on the Poincaré half-plane satisfying two conditions. The first one is an invariance:

$$f(z) = (cz+d)^{-p} f\left[\frac{az+b}{cz+d}\right], \quad z \in \mathcal{P}, \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL(2,\mathbb{Z}).$$

The second one is a growth condition: observe that $f(z) = f(z+1)$, so that $\tilde{f}$ can be defined on the punctured unit disc Δ^* by $\tilde{f}(e^{2i\pi z}) = f(z)$; the second defining condition is that the Laurent expansion of $\tilde{f}$ in Δ^* is of the form $\tilde{f}(w) = \sum_{n \geq 1} a_n w^n$ for $w \in \Delta^*$.

It is a result of Hecke that a cusp form f of weight p satisfies $|f(x+iy)| \leq By^{-p/2}$ for all $x + iy \in \mathcal{P}$ and for some constant B; see page I.24 in [Ogg].

Let $M = W^*(\Gamma)$. Consider an integer $k \geq 2$ and the M-module H_k of example 3.3.4. Given a cusp form f of weight p, the growth condition implies that f induces a multiplication operator $A_f : H_k \to H_{k+p}$, defined by $(A_f\varphi)(z) = f(z)\varphi(z)$, which is bounded (in fact $\|A_f\| \leq B$ with B as above). The invariance condition implies that A_f is M-linear. Consequently, given two cusp forms f,g of weight p, the operator $A_f^* A_g : H_k \to H_k$ is in the commutant $\operatorname{End}_M(H_k)$. Let T_k denote the natural trace on $\operatorname{End}_M(H_k)$. Then the space of cusp forms of weight p has a natural hermitian form

$$\langle f|g \rangle_k = T_k(A_f^* A_g).$$

A computation in the same spirit as that presented in the proof of Theorem 3.3.2 shows that

$$\langle f|g \rangle_k = \frac{k-1}{\pi} \int_D \overline{f(z)} g(z) y^{p-2} dxdy$$

with D a fundamental domain for Γ in $\mathcal{P}$. Up to a constant factor $\frac{k-1}{\pi}$, this is known as the Peterson scalar product for cusp forms.

This suggests a natural project, which could be interesting for the study of cusp forms: evaluate L^q-norms defined by

$$\|f\|_{k,q} = \{T_k((A_f^* A_f)^{q/2})\}^{1/q}.$$

The equality $T_k(A_f^* A_g) = T_{k+p}(A_g A_f^*)$ should be useful.

3.4. Index for subfactors of II_1 factors.

There were two main motivations for the introduction in [Jo1] of the concept of index for subfactors. The first was that, if $\Gamma_1 < \Gamma_2$ are two icc discrete groups, the II_1 factor

$N = \lambda(\Gamma_1)'$ acts in an obvious way on $\ell^2(\Gamma_2)$ and $\dim_N(\ell^2(\Gamma_2)) = [\Gamma_2:\Gamma_1]$. Furthermore $\ell^2(\Gamma_2)$ is the same as $L^2(M)$ where M is $\lambda(\Gamma_2)'$. This suggested the following definition:

Definition 3.4.1. The index of a subfactor N of a finite factor M is

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

This was the original definition of index; it was shown in [Jo4] that this definition agrees with the ring-theoretic one which we have given in Chapter 2, when M and N are finite factors. The index can also be computed as $[M:N] = \dim_N(H)/\dim_M(H)$, where H is any M-module of finite dimension over M; see Proposition 3.4.6.

The second motivation was a result of M. Goldman [Gol], who showed that, if $N \subset M$ are II_1 factors (always with the same identity $\mathbf{1}$) then, if $\dim_N(L^2(M)) = 2$, there is a crossed product decomposition $M = N \rtimes \mathbb{Z}/2\mathbb{Z}$. Consequently if one defines $[M:N]$ as above, Goldman's result is seen to be a beautiful analogue of the fact that a subgroup of index 2 of a group is normal.

It would also have been nice to have been able to call a subfactor $N \subset M$, normal when its (unitary) normalizer generates M. But unfortunately standard terminology reserves "normal" for subfactors N such that $(N' \cap M)' \cap M = N$, and the term regular is used for subfactors with the normalizer property described above. We take this opportunity to introduce some more terminology.

Definition 3.4.2. If $N \subset M$ are factors we say that N is irreducible if $N' \cap M = \mathbb{C}$.

It is not hard to see that a regular irreducible subfactor has integer index (or ∞ which we shall treat as an integer) — see [Jo7]. A more refined analysis based on [Jo6] shows that all regular subfactors have integer index. On the other hand $\dim_M(H)$ can be any positive real number so the question naturally arose:

(a) What are the possible values of $[M:N]$?
(b) What are the possible values of $[M:N]$ for an irreducible pair $N \subset M$?

Question (a) was settled completely in [Jo1] for $M = R$, the hyperfinite II_1 factor. Question (b) remains open even for $M = R$, and question (a) is open for arbitrary II_1 factors M. We summarize the most important known results as follows:

Theorem 3.4.3. *Let* N *be a subfactor of a* II_1 *factor* M.

(i) *Either* $[M:N] = 4\cos^2\pi/q$ *for some integer* $q \geq 3$, *or* $[M:N] \geq 4$.
(ii) *If* $[M:N] < 4$, *then* N *is automatically irreducible in* M.
(iii) *There exist subfactors of the hyperfinite* II_1 *factor* R *with any of the index values allowed by* (i).
(iv) *There are examples of factors* M *for which the set of all possible values* $[M:N]$ *is countable.*

Remarks: Statements (i) to (iii) are from [Jo1]. We prove (i) below. A generalization to finite direct sums of II_1 factors is shown in Corollary 3.7.6. A second proof of (i) occurs in Corollary 4.6.6, as a byproduct of the analysis of "derived towers".

Statement (ii) is proved as Corollary 3.6.2(c).

We will verify (iii) by giving several constructions of subfactors of R. The first construction, in this section, works for all allowed index values. Another construction, valid for the index values $4\cos^2\pi/q$ is given in Theorem 4.4.2. A third construction, in Section 4.5, produces irreducible pairs; the index values $4\cos^2\pi/q$ are obtained once more, as well as sporadic values greater than 4. In Section 4.7.d, we give examples of non-conjugate irreducible subfactors of R of index 4. We would also like to mention the work of Wenzl [Wen2], in which a family of irreducible subfactors of R of index greater than 4 is produced by a construction involving the Hecke algebras $H_\infty(q)$ for q a primitive root of unity.

Statement (iv) is from [PP2], and will not be proved here.

For arbitrary II_1 factors, the question of existence of subfactors of index $4\cos^2\pi/q$ remains open, more precisely we know of no example of a full II_1 factor M having a subfactor of index $4\cos^2\pi/q$, $q \neq 3,4,6$. (A II_1 factor is called "full" if the group of inner automorphisms is closed in the topology of pointwise strong convergence in the whole automorphism group — an example of such a factor is $\lambda(PSL(2,\mathbb{Z}))'$.)

Proof of 3.4.3 (i). As for finite dimensional algebras (2.6.2) there is always a (unique) faithful trace-preserving conditional expectation from M onto N, which, viewed as an operator on $L^2(M)$ is the orthogonal projection e_N onto $L^2(N)$. The fundamental construction again yields a II_1-factor

$$\mathrm{End}_N^r(L^2(M,tr)) = \mathrm{End}_N^r(M) = \langle M, e_N \rangle.$$

(See Theorem 3.4.6 below for the first equality.)

We claim that the normalized trace of $\langle M, e_N \rangle$ has the Markov property

$$[M:N]\mathrm{tr}(e_N x) = \mathrm{tr}(x) \qquad \text{for all } x \in M.$$

Indeed, the linear form defined on N by $x \mapsto \mathrm{tr}(e_N x)$ is a trace (3.6.1.iii). As $1 = \mathrm{tr}(e_N)\,[M:N]$ by Proposition 3.2.5.e applied to the N-module $L^2(M)$, the property is valid for $x \in N$, by uniqueness of the normalized trace on N. But then for $x \in M$, we have $[M:N]\mathrm{tr}(e_N x) = [M:N]\mathrm{tr}(e_N x e_N) = [M:N]\mathrm{tr}(e_N E_N(x)) = \mathrm{tr}(E_N(x)) = \mathrm{tr}(x)$, using 3.6.1.i .

Now the tower construction of Chapter 2 works and yields an increasing sequence of II_1-factors

$$M_0 = N \subset M_1 = M \subset \cdots \subset M_k \subset M_{k+1} \subset \cdots,$$

and a sequence of self-adjoint projections $(e_i)_{i\geq 1}$ satisfying

$$\text{(3.4.3.1)} \qquad \begin{aligned} &\beta\, e_i e_{i\pm 1} e_i = e_i \\ &e_i e_j = e_j e_i \quad \text{if } |i-j| \geq 2, \end{aligned}$$

with $\beta = [M:N]$. Claim (i) now follows from Theorem II.16.

An alternative proof using the trace goes as follows: The trace tr on $\bigcup_k M_k$ has the Markov property

$$\text{(3.4.3.2)} \qquad \beta\,\mathrm{tr}(we_j) = \mathrm{tr}(w) \quad \text{for } j \geq 1 \text{ and } w \in \mathrm{alg}\,\{\mathbf{1}, e_1, \cdots, e_{j-1}\}.$$

where $\beta = [M:N]$. Now suppose that $\beta < 4$ but $\beta \notin \{4\cos^2\pi/q : q \geq 3\}$. Using 2.8.5 and 2.8.7 (note that the number β is generic) as well as the relations 3.4.3.1 and 3.4.3.2, we obtain for each $k \geq 1$ a trace preserving isomorphism of the algebra $B_{\beta,k}$ of Section 2.8 onto the algebra $C_k = \{\mathbf{1}, e_1, \cdots, e_{k-1}\}''$. By 2.8.4(vii), for each k the trace of the minimal central projection Q_0^k (necessarily a self-adjoint projection in C_k) is $P_k(\beta^{-1})$. But by 2.8.3(iii), if $4\cos^2(\pi/k) < \beta < 4\cos^2(\pi/k+1)$, then $P_k(\beta^{-1}) < 0$, contradicting the positivity of the trace. It follows that if $\beta < 4$, then $\beta \in \{4\cos^2\pi/q : q \geq 3\}$. #

<u>Proof of 3.4.3(iii).</u> Fix $\beta \in \mathbb{R}$ with $\beta = 4\cos^2\pi/q$ for some integer $q \geq 3$, or $\beta \geq 4$. Consider a sequence of self-adjoint projections $(e_i)_{i\geq 1}$ on a Hilbert space, together with a faithful normal tracial state tr on $R = \{\mathbf{1}, e_1, e_2, \cdots\}''$ satisfying the relations 3.4.3.1 as well as the Markov property 3.4.3.2.

First we must recall how such a sequence of projections and such a trace can be constructed. In 2.8.4 (in case $\beta \geq 4$) and in Section 2.9 (in case $\beta = 4\cos^2\pi/q$ for some q) we have constructed an increasing sequence of finite dimensional C^*-algebras $(B_{\beta,k})_{k\geq 1}$, with $B_{\beta,k}$ generated by its identity and self-adjoint projections $e_1, \cdots e_{k-1}$ satisfying the relations 3.4.3.1, and a positive faithful normalized trace tr on $B_{\beta,k}$ satisfying the relation 3.4.3.2 for $1 \leq j \leq k$. Since tr is faithful, the trace representation π_{tr} is faithful as well, and we can take R to be $\pi_{tr}(\bigcup B_{\beta,k})''$.

A simpler procedure is available when β is the square of the norm of a non-negative integer valued matrix (i.e. $\sqrt{\beta} \in \mathcal{N}(\mathbb{N})$). In this case there is a connected pair of finite dimensional C^*-algebras $B \subset A$ with $[A:B] = \beta$, and the tower construction for this pair yields a sequence of projections $(e_i)_{i\geq 1}$ satisfying 3.4.3.1, and a positive faithful trace on $\mathrm{alg}\,\{\mathbf{1}, e_1, \cdots\}$ satisfying 3.4.3.2. Cf. 2.7.5 and the discussion at the end of Appendix IIa.

<u>Lemma 3.4.4.</u> [Jo1] *With the notation above,* R *is the hyperfinite* II_1 *factor.*

<u>Proof.</u> It is clear that R is a finite, hyperfinite von Neumann algebra. We claim that if z is in the center of R, then

$$\mathrm{tr}(zx) = \mathrm{tr}(z)\mathrm{tr}(x) \quad \text{for all } x \in R.$$

It will follow from this and the faithfulness of tr that $z = \mathrm{tr}(z)\mathbf{1}$, so R is a factor.

For each k, let $C_k = \mathrm{alg}\,\{\mathbf{1}, e_1, \cdots e_{k-1}\}$. By 2.9.6(e) (in case $\beta < 4$) or by 2.8.7(a) and 2.8.5(b) (in case $\beta \geq 4$), the map $\epsilon_j \mapsto e_j$ (on the generators $\{\epsilon_j\}$ of $\mathcal{A}_{\beta,k}$) induces a trace preserving isomorphism of $B_{\beta,k}$ onto C_k. It then follows from 2.9.6(g) (for $\beta < 4$) or from 2.8.5(f) (for $\beta \geq 4$) that $e_j \mapsto e_{k-j}$ extends to an inner automorphism of C_k, and hence to an inner automorphism α_k of R.

Note that tr has the multiplicative property $\mathrm{tr}(y_1 y_2) = \mathrm{tr}(y_1)\mathrm{tr}(y_2)$ whenever $y_1 \in C_s$ and $y_2 \in \mathrm{alg}\,\{\mathbf{1}, e_s, \cdots e_{s+m}\}$. (One can verify this directly or use the isomorphism $C_m \cong B_{\beta,m}$ together with 2.8.5(e) or 2.9.6(f).)

It will suffice to verify the relation $\mathrm{tr}(zx) = \mathrm{tr}(z)\mathrm{tr}(x)$ when $x \in C_k$ for some k. Let $\epsilon > 0$, and choose $y \in C_\ell$ for some ℓ, such that $\|z-y\|_2 < \epsilon$. Then $\mathrm{tr}(y\alpha_{k+\ell}(x)) = \mathrm{tr}(y)\mathrm{tr}(x)$, since $\alpha_{k+\ell}(x) \in \mathrm{alg}\,\{\mathbf{1}, e_{\ell+1}, \cdots e_{\ell+k-1}\}$. Consequently,

$$\begin{aligned} &|\mathrm{tr}(zx) - \mathrm{tr}(z)\mathrm{tr}(x)| \\ &= |\mathrm{tr}(z\alpha_{k+\ell}(x)) - \mathrm{tr}(z)\mathrm{tr}(x)| \qquad \text{(since } \alpha_{k+\ell} \text{ is inner)} \\ &\leq |\mathrm{tr}((z-y)\alpha_{k+\ell}(x))| + |\mathrm{tr}(y\alpha_{k+\ell}(x)) - \mathrm{tr}(z)\mathrm{tr}(x)| \end{aligned}$$

$$= |\mathrm{tr}((z-y)\alpha_{k+\ell}(x))| + |(\mathrm{tr}(y)-\mathrm{tr}(z))\mathrm{tr}(x)|$$
$$\le 2\,\epsilon\,\|x\|_2.$$

Since ϵ is arbitrary, this finishes the proof. #

Lemma 3.4.5. [Jo1] Set $R_\beta = \{\mathbf{1}, e_2, e_3, \cdots\}''$. Then $[R:R_\beta] = \beta$.

Proof. We know (by 2.8.5 and 2.8.7 or by 2.9.6) that for each $k \ge 2$, the relation $\beta\,\mathrm{tr}(e_1x) = \mathrm{tr}(x)$ holds when $x \in \mathrm{alg}\,\{\mathbf{1}, e_2, \cdots e_k\}$, and, taking limits, we have the same relation also for $x \in R_\beta$. Therefore $E_{R_\beta}(e_1) = \beta^{-1}\mathbf{1}$.

Similarly $E_N(e_2) = \beta^{-1}\mathbf{1}$, where $N = \{\mathbf{1}, e_3, e_4, \cdots\}''$. For $k \ge 3$, any $x \in \mathrm{alg}\,\{\mathbf{1}, e_2, \cdots e_k\}$ is of the form $x = a + \sum_i b_i e_2 c_i$, with $a, b_i, c_i \in \mathrm{alg}\,\{\mathbf{1}, e_3, \cdots e_k\}$. Consequently, $E_N(x) = a + \beta^{-1}\sum_i b_i c_i$ and $e_1xe_1 = E_N(x)e_1$. Taking limits again, we have

$$(*) \qquad e_1xe_1 = E_N(x)e_1 \quad \text{for all } x \in R_\beta.$$

One next verifies that $xe_1 = \beta\,E_{R_\beta}(xe_1)e_1$ for all $x \in R$, by first checking this for $x \in \mathrm{alg}\,\{\mathbf{1}, e_1, \cdots e_k\}$ (that is, for x of the form $x = a + \sum_i b_i e_1 c_i$, with $a, b_i, c_i \in \mathrm{alg}\,\{\mathbf{1}, e_2, \cdots e_k\}$) and then by taking limits. Consequently $Re_1 = R_\beta e_1$, and $Re_1R = R_\beta e_1 R_\beta$. Observe also that $R = Re_1R$, because finite factors are algebraically simple ([DvN], Cor. III.5.3).

Let e be the orthogonal projection of $L^2(R)$ onto $L^2(R_\beta)$. One has $exe = E_{R_\beta}(x)e$ for all $x \in R$, by 3.6.1.i. below, so that in particular, $ee_1e = \beta^{-1}e$. We claim that also $e_1ee_1 = \beta^{-1}e_1$. Since $R = R_\beta e_1 R_\beta$, it suffices to check this equality on vectors $xe_1y\Omega$, where $x, y \in R_\beta$ and Ω is the trace vector for R. But

$$\begin{aligned}
e_1ee_1(xe_1y\Omega) &= e_1ee_1E_N(x)y\Omega && \text{(by } (*))\\
&= e_1E_{R_\beta}(e_1E_N(x)y)\Omega && \text{(by definitions of } e \text{ and } E_{R_\beta})\\
&= e_1E_{R_\beta}(e_1)E_N(x)y\Omega && \text{(by } R_\beta\text{-linearity of } E_{R_\beta})\\
&= \beta^{-1}\,e_1E_N(x)y\Omega &&\\
&= \beta^{-1}\,e_1(xe_1y\Omega) && \text{(by } (*)).
\end{aligned}$$

It follows from the relations $ee_1e = \beta^{-1}e$ and $e_1ee_1 = \beta^{-1}e_1$ that e and e_1 are equivalent projections in $\langle R,e\rangle$. Since e is finite in $\langle R,e\rangle$ by 3.6.1(v), the projection e_1 is finite in $\langle R,e\rangle$. But $\mathbf{1}$ is the sum of finitely many projections each equivalent in R to a subprojection of e_1, so $\langle R,e\rangle$ is finite. Hence $[R:R_\beta] = \mathrm{tr}(e)^{-1} = \mathrm{tr}(e_1)^{-1} = \beta$.

This completes the proof of the lemma, and also of 3.4.3(iii). #

It is tempting to guess that the pair $R \supset R_\beta$ is irreducible, also for $\beta > 4$, since on a purely algebraic level it is easy to see that there is no element of the algebra generated by $\{e_1, e_2, \cdots\}$ which commutes with $\{e_2, e_3, \cdots\}$. V. Jones confesses to spending considerable effort to prove this, but it turned out that R_β has non-trivial relative commutant in R when $\beta > 4$. A laborious proof of this non-obvious fact was given in [Jo1] and a simpler proof in [PP1]; we will give a proof due to Popa in 4.7.5. The difficulty is that one cannot write down an explicit form for an element in $R_\beta' \cap R$ without invoking a beautiful representation of $\{e_1, e_2, \cdots\}''$ discovered by Pimsner and Popa.

We have seen that one way to obtain a sequence of projections $(e_i)_{i\ge1}$ satisfying the relations 3.4.3.1 is to form the tower from an indecomposable pair $B \subset A$ of finite dimensional C^*-algebras. Then, as we have observed in Chapter 2, the restrictions on index are related to restrictions on the type of inclusions $B \subset A$ which yield a modulus $\beta < 4$. This is where the Coxeter graphs of types A, D, and E enter the picture. But to create the sequence $(e_i)_{i\ge1}$ one can also use a pair $N \subset M$ of finite direct sums of II_1-factors. In the following sections we will see how, if one allows this extra freedom, the remaining Coxeter graphs appear!

We finish this section by recording one useful fact on index of subfactors from [Jo1].

Proposition 3.4.6. *Let* $N \subset M$ *be finite factors and let* H *be any* M*-module such that* $\dim_M(H)$ *is finite. Then* $[M:N] = \dfrac{\dim_N(H)}{\dim_M(H)}$. *(In particular,* $\dim_N(H) \ge \dim_M(H)$.*)*

Proof. If H_1 and H_2 are any two M-modules such that $\dim_M(H_i)$ is finite for $i = 1,2$, then there is a finite dimensional Hilbert space K and an M-invariant projection q such that $H_1 \cong q(H_2 \otimes K)$ as M-modules. Then $\dim_N(H_1) \le \dim_N(H_2 \otimes K) = \dim_N(H_2)\dim_{\mathbb{C}}(K)$, by 3.2.5(i), so $\dim_N(H_1)$ is finite if and only if $\dim_N(H_2)$ is. In particular, $[M:N]$ is finite if and only if $\dim_N(H)$ is.

Assuming that $[M:N]$ is finite and choosing an M-module isomorphism $H \cong q(L^2(M) \otimes K)$, as above, we have

$$\begin{aligned}\dim_N(H) &= \dim_N(q(L^2(M) \otimes K)) \\ &= tr_{N'}(q) \dim_N(L^2(M) \otimes K) \qquad \text{(by 3.2.5(e))} \\ &= tr_{M'}(q) \dim_{\mathbb{C}}(K) \dim_N(L^2(M)) \quad \text{(by 3.2.5(i))},\end{aligned}$$

while $\dim_M(H) = tr_{M'}(q) \dim_{\mathbb{C}}(K)$. #

3.5. Inclusions of finite von Neumann algebras with finite dimensional centers.

We saw in Chapter 2 that a unital inclusion $B \subset A$ of finite dimensional C^*-algebras can be specified by the inclusion matrix $\Lambda \in Mat_{fin}(\mathbb{N})$ and a vector $\vec{\nu} \in \mathbb{N}^n$ for some n, specifying the algebra B up to isomorphism. It is impossible to specify an inclusion so precisely in the II_1-case since, for example, it is possible to find infinitely many non-conjugate subfactors of index 4 in R, even irreducible ones, as we shall see in Chapter 4. What we will do is specify enough information to be able to calculate all the needed coupling constants, which will enable us to find the Markov traces as in Section 2.7.

The situation will differ in two ways from the finite dimensional case. The first is that there are no minimal projections around, so integers do not appear in this way. The second is that the subfactors can have indices different from squares of integers. This extra freedom allows the appearance of new Coxeter graphs.

First some notation. Let $M = \bigoplus_{i=1}^{m} M_i$ be a direct sum of finite factors with corresponding minimal central projections $p_1,\cdots,p_m$. Since the trace on a finite factor is unique up to a scalar multiple, a trace on M is completely specified by a row vector $\overleftarrow{s} = (s_1,\cdots,s_m)$, with $s_i = tr(p_i)$. (*Warning:* This is *not* the same vector which was used in Chapter 2 to specify a trace on a direct sum of finite dimensional factors; there we used the vector whose i^{th} component is the trace of a minimal projection in M_i.) A trace is positive (i.e., trace $(a^*a) \geq 0$) if and only $\overleftarrow{s}$ has non-negative components. *We adopt the convention that "trace" means "positive trace".* A trace is faithful (i.e., trace $(a^*a) = 0$ implies $a = 0$) if none of the components of $\overleftarrow{s}$ are zero, and normalized if $\sum_{i=1}^{m} s_i = 1$. A trace is automatically normal; i.e., if $\{f_i\}$ is a family of mutually orthogonal projections, then trace $(\bigvee_{i=1}^{\infty} f_i) = \sum_{i=1}^{\infty} \text{trace}(f_i)$.

Recall that if P is a finite factor, tr_P denotes its unique normalized trace, and if Tr is any other trace on P, then $Tr = Tr(\mathbf{1})tr_P$.

Let $N = \bigoplus_{j=1}^{n} N_j$ be another direct sum of finite factors, contained in M and having the same identity. Let $q_1,\cdots,q_n$ be the minimal central projections of N.

<u>Definition 3.5.1.</u> If $N \subset M$ are as above, we define the m-by-n matrix $T_N^M = (c_{i,j})$ by

$$c_{i,j} = tr_{p_iM}(p_iq_j).$$

<u>Proposition 3.5.2.</u>

(i) *The matrix* T_N^M *is row-stochastic; i.e.,* $c_{i,j} \geq 0$ *and* $\sum_j c_{i,j} = 1$ *for all* i.

(ii) *If* $\overleftarrow{s}$ *specifies a trace on* M, *then* $\overleftarrow{s}\, T_N^M$ *specifies its restriction to* N.

(iii) *If* $N \subset M \subset L$ *are finite direct sums of finite factors, then* $T_N^L = T_M^L T_N^M$.

<u>Proof.</u> (i) $\sum_j c_{i,j} = \sum_j tr_{p_iM}(p_iq_j) = tr_{p_iM}(p_i) = 1$, since $\sum_j q_j = \mathbf{1}$.

(ii) As $\sum_i p_i = \mathbf{1}$,

$$\text{trace}(q_j) = \sum_i \text{trace}(p_iq_j) = \sum_i \text{trace}(p_i) tr_{p_iM}(p_iq_j) = \sum_i s_i c_{i,j}.$$

(iii) Let $\{r_k\}$ denote the minimal central projections of L, so that T_N^L is the matrix whose (k,j) entry is $tr_{r_kL}(r_kq_j)$. Since $q_j = \sum_i p_iq_j$, one has

$$(T_N^L)_{k,j} = \sum_i tr_{r_kL}(r_kp_iq_j).$$

But in the finite factor $A = r_kL$, if $e \leq f$ are two projections, then $tr_A(e) = tr_A(f) tr_{fAf}(e)$. Thus

$$(T_N^L)_{k,j} = \sum_i tr_{r_kL}(r_kp_i) tr_{r_kp_iLr_kp_i}(r_kp_iq_j).$$

If $r_k p_i \neq 0$, then $x \mapsto tr_{r_k p_i L r_k p_i}(r_k x)$ is a trace on $p_i M$ whose value at p_j is 1, so in fact $tr_{r_k p_i L r_k p_i}(r_k x) = tr_{p_i M}(x)$. Hence

$$(T_N^L)_{k,j} = \sum_i tr_{r_k L}(r_k p_i) tr_{p_i M}(p_i q_j)$$
$$= \sum_i (T_M^L)_{k,i} (T_N^M)_{i,j},$$

as desired. #

A second piece of data needed is the matrix of indices of the "partial embeddings". Note that $N_{i,j} = Np_iq_j = \{p_iq_jx : x \in N\}$ is a finite factor, a subfactor of $M_{i,j} = p_iq_jMp_iq_j$.

Definition 3.5.3. (i) With notation as above, define an m-by-n matrix Λ_N^M with entries

$$\lambda_{i,j} = [M_{i,j} : N_{i,j}]^{1/2}.$$

(We note that this expression is the same as in the finite dimensional case. Observe that in the finite dimensional case Λ_N^M determines T_N^M, namely

$$(T_N^M)_{i,j} = \lambda_{i,j}\nu_j/\mu_i$$

where $p_i M \cong Mat_{\mu_i}(\mathbb{C})$, and $q_j N \cong Mat_{\nu_j}(\mathbb{C})$.)

(ii) The inclusion $N \subset M$ is called connected if $Z(M) \cap Z(N) = \mathbb{C}\mathbf{1}$. This is true if and only if Λ_N^M is indecomposable.

(iii) A representation π of M on a Hilbert space H is called a finite representation of the pair $N \subset M$ if $\pi(N)'$ is a finite von Neumann algebra.

(iv) We say that N is of finite index in M if $N \subset M$ admits a finite faithful representation.

(Note that parts (ii), (iii), and (iv) make sense for arbitrary pairs of finite von Neumann algebras - not necessarily with finite dimensional centers.)

Lemma 3.5.4. *Suppose* $N \subset M$ *are finite direct sums of finite factors. The following are equivalent:*

(i) *N is of finite index in M.*

(ii) *The matrix* Λ_N^M *has only finite entries.*

(iii) *For any faithful trace* tr *on* M, *the regular representation of* M *on* $L^2(M,tr)$ *is a finite representation of the pair* $N \subset M$.

(iv) *For any faithful representation* $\{\pi, \mathcal{H}\}$ *of* M *such that* $\pi(M)'$ *is finite, the algebra* $\pi(N)'$ *is also finite.*

Proof. (iv) ⇒ (iii) ⇒ (i) is evident.

(i) ⇒ (ii). If π is a faithful finite representation of the pair $N \subset M$ on H, then the commutant of $\pi(N_{i,j})$ on $\pi(p_iq_j)H$ is $\pi(p_iq_j)\pi(N)'\pi(p_iq_j)$, which is finite. It follows that $\dim_{N_{i,j}}(\pi(p_iq_j)H) < \infty$ (Proposition 3.2.4.d), and

$$[M_{i,j} : N_{i,j}] = \dim_{N_{i,j}}(L^2(M_{i,j})) = \frac{\dim_{N_{i,j}}(\pi(p_iq_j)H)}{\dim_{M_{i,j}}(\pi(p_iq_j)H)}$$

(by 3.4.6.), which is finite.

(ii) ⇒ (iv). Consider a faithful M-module H for which M' is finite. Since $\mathbf{1} = \sum_{i,j} p_iq_j$, to show that N' is finite, it suffices to show that each p_iq_j is a finite projection in N' (because a sum of finite projections is finite.) If $p_iq_j \neq 0$, then $p_iq_jN'p_iq_j$ is the commutant of $N_{i,j}$ on p_iq_jH. By 3.4.6 and 3.2.5.h,

$$\dim_{N_{i,j}}(p_iq_jH) = \lambda_{i,j}^2 \dim_{M_{i,j}}(p_iq_jH)$$
$$= \lambda_{i,j}^2 tr_{p_iM}(p_iq_j)^{-1} \dim_{p_iM}(p_iH).$$

Since M' is finite on H, so is $p_iM' = (p_iM)'$ on p_iH, so by 3.2.4.d, $\dim_{p_iM}(p_iH) < \infty$. Hence also $\dim_{N_{i,j}}(p_iq_jH) < \infty$, and by 3.2.4.d again, $(N_{i,j})'$ is finite. #

Observe that the analogue for Λ of Proposition 3.5.3.iii does not hold. For example, let R be the hyperfinite II_1 factor, let p be a non-trivial projection in R, let φ be an isomorphism from R_p to $R_{\mathbf{1}-p}$, and set

$$N = \{y \in R: y = x + \varphi(x) \text{ for some } x \in R_p\}, \text{ and}$$
$$M = R_p \oplus R_{\mathbf{1}-p}.$$

Then

$$\Lambda_M^R \Lambda_N^M = (1\ 1)\begin{bmatrix}1\\1\end{bmatrix} = 2,$$

and
$$\Lambda_N^R = [R:N]^{1/2} = \{\frac{1}{tr(p)} + \frac{1}{tr(1-p)}\}^{1/2}$$

by Corollary 2.2.5 of [Jo1] or 4.7.2. These are not equal, unless $tr(p) = 1/2$.

Of course, if $N \subset M \subset L$ is a triple of finite factors, then $[L:N] = [L:M]\,[M:N]$ by Proposition 3.4.6.

If N and M are as in 3.5.4, and the inclusion $N \subset M$ is connected, then all factors of N and M are of type II_1, or $\dim_{\mathbb{C}}(M) < \infty$. It is also known that all factors of N and M share (or do not share) the property of being hyperfinite (Lemma 2.1.8 in [Jo1]) or the property T (see [Ana] and [PP2]).

If π is a finite faithful representation of the pair $N \subset M$ on H, then the centers of $\pi(M)'$ and $\pi(N)'$ are the same as those of M and N respectively, and the rows and columns of $\Lambda_{\pi(M)'}^{\pi(N)'}$ are naturally indexed by the columns and rows of Λ_N^M. The generalization of Proposition 2.3.5 to this setting is the following.

Lemma 3.5.5. *Let $N \subset M$ be a pair of finite direct sums of finite factors, as above, as suppose π is a faithful finite representation of the pair. Then*
$$\Lambda_{\pi(M)'}^{\pi(N)'} = (\Lambda_N^M)^t.$$

Proof. If M and N are factors, the equality holds because $[\pi(N)':\pi(M)'] = [M:N]$ by Propositions 3.4.6 and 3.2.5.g. To extend the equality to the general case, one proceeds exactly as in the finite dimensional case (Proposition 2.3.5), with Proposition 2.2.5b being replaced by [DvN], Proposition 1 of §I.2, which says: if Q is a von Neumann algebra on H and p is a projection in Q or in Q', then $\mathrm{End}_{pQp}(pH))$ equals $p\mathrm{End}_Q(H)p$. #

Also note that $\pi(M)'$ is of finite index in $\pi(N)'$ by Lemmas 3.5.4. and 3.5.5.

Proposition 3.5.6. *Given an irredundant m-by-n matrix Λ over $\{0\} \cup \{2\cos\pi/q : q \geq 3\} \cup [2,\infty]$, and an m-by-n row stochastic matrix T having the same pattern of zero entries as Λ, there exists a pair $N \subset M$ (both hyperfinite) with $\Lambda_N^M = \Lambda$ and $T_N^M = T$.*

Proof. Take M to be the direct sum of m copies of R, the unique hyperfinite II_1 factor, denoted R_i. In each R_i, choose a partition of unity $\{q_{i,j} : 1 \leq j \leq n\}$ with $tr(q_{i,j}) = (T)_{i,j}$. If $(T)_{i,j}$ is non-zero choose a II_1 subfactor $P_{i,j}$ of $R_{i,j} = q_{i,j}R_iq_{i,j}$ with $[R_{i,j} : P_{i,j}]^{1/2} = (\Lambda)_{i,j}$ (possible by [Jo1], Theorem 4.3.2). For each i and j such that $(T)_{i,j} \neq 0$, choose an isomorphism $\theta_{i,j} : R \to P_{i,j}$ (possible since all the factors are II_1 and hyperfinite). Set $q_j = \sum_i q_{i,j}$, put $N_j = \{\sum_i \theta_{i,j}(x) : x \in R\}$, and $N = \bigoplus_{j=1}^n N_j$. Then $q_jN = N_j$, and N is the required subalgebra. #

3.6. The fundamental construction.

The discussion of the fundamental construction in Chapter 2 was purely ring theoretic. In the von Neumann algebra framework, where the preferred modules are Hilbert spaces, it is natural to make a construction which, apparently, depends on the choice of a trace on M. We begin by showing that in fact the ring theoretic construction is exactly the same.

First we recall some notions from [Jo1] which work for arbitrary finite von Neumann algebras exactly as for factors. Let $N \subset M$ be finite von Neumann algebras with the same identity. Given a faithful normalized trace on M, there is a unique faithful normal conditional expectation $E_N : M \to N$ determined by $tr(xy) = tr(E_N(x)y)$ for $x \in M$ and $y \in N$. In fact E_N is the restriction to M of the orthogonal projection $e_N : L^2(M,tr) \to L^2(N,tr)$. We denote by $\langle M,e_N\rangle$ the von Neumann algebra on $L^2(M,tr)$ generated by M and e_N.

We let J denote the conjugate linear isometry of $L^2(M,tr)$ extending the map $x \mapsto x^*$ on M.

Proposition 3.6.1.

(i) $e_Nxe_N = E_N(x)e_N$ *for* $x \in M$

(ii) $Je_NJ = e_N$

(iii) *For* $x \in M$, x *commutes with* e_N *if and only if* $x \in N$.

(iv) $\langle M,e_N\rangle = JN'J$

(v) *The map* $\psi \begin{cases} N \longrightarrow \langle M, e_N\rangle \\ y \longmapsto y\,e_N \end{cases}$ *is an injective morphism onto* $e_N\langle M,e_N\rangle e_N$.

(vi) *The central support of* e_N *in* $\langle M,e_N\rangle$ *is* $\mathbf{1}$.

(vii) *The space* Me_NM, *which denotes the linear span of* $\{x'e_Nx'' : x',x'' \in M\}$, *is a strongly dense *-subalgebra of* $\langle M,e_N\rangle$.

Proof. (cf. [Jo1]).

(i) It suffices to check that $E_N(xE_N(y)) = E_N(x)E_N(y)$, but this follows from the N-linearity of E_N.

(ii) Follows from $E_N(x^*) = E_N(x)^*$.

(iii) Note that x commutes with e_N if and only if left multiplication by x commutes with E_N. This is clearly so for $x \in N$. On the other hand, if $x \in M$ and x commutes with E_N, then $x = xE_N(\mathbf{1}) = E_N(x) \in N$.

(iv) By (iii) $N = M \cap \{e_N\}'$, so $N' = (M' \cup \{e_N\})'' = \langle M', e_N \rangle$. But $JM'J = M$ and $Je_NJ = e_N$, so $JN'J = \langle M, e_N \rangle$.

(v) By (i), the indicated map is an epimorphism. Let Ω denote the canonical trace vector in $L^2(M,tr)$. If $ye_N = 0$, then $ye_N\Omega = y\Omega = 0$ and $y = 0$ because Ω is separating, so ψ is an isomorphism.

(vi) Let z be the central support of e_N in N'. Then $z \in N \cap N'$ and $\psi(z-\mathbf{1}) = ze_N - e_N = 0$, by definition of a central support, so $z = \mathbf{1}$ by (v). Now (vi) follows from (iv) and (ii).

(vii) First note that by (i), the set

$$X = \{x_0 + \sum_{i=1}^{n} x_i e_N y_i : n \in \mathbb{N}, x_i, y_i \in M\}$$

is a $*$-subalgebra of $\langle M, e_N \rangle$ containing M and e_N, so the strong closure of X is $\langle M, e_N \rangle$. If

$$Y = \{\sum x_i e_N y_i : x_i, y_i \in M\},$$

then Y is a two sided ideal in X, so by the Kaplansky density theorem and the joint strong continuity of muliplication on the unit ball, the strong closure $\overline{Y}$ of Y is a two sided ideal in $\langle M, e_N \rangle$. But $\overline{Y}$ contains the central support of e_N, which is $\mathbf{1}$ by point (vi), so $\overline{Y} = \langle M, e_N \rangle$. #

We now specialize to the case where N and M are direct sums of finitely many II_1 factors with minimal central projections $\{q_j; j = 1,\cdots,n\}$ and $\{p_i; i = 1,\cdots,m\}$ respectively. By the equality (iv) above, $\langle M, e_N \rangle$ is also a finite direct sum of II_1 factors, with minimal central projections $\{Jq_jJ: j=1,\cdots,n\}$.

Lemma 3.6.2.

(a) *If* $N \subset M$ *are type* II_1 *von Neumann algebras with finite dimensional centers and* N *is of finite index in* M, *then* $\dim_{\mathbb{C}}(N' \cap M) < \infty$.

(b) *If* $N \subset M$ *are* II_1 *factors, then* $\dim_{\mathbb{C}}(N' \cap M) \le [M:N]$.

(c) *If* $N \subset M$ *are* II_1 *factors with* $[M:N] < 4$, *then* $N' \cap M = \mathbb{C}\mathbf{1}$.

Proof. We first consider the case that N and M are factors. Let $H = L^2(M)$ and write $Tr_{N'}$ for the natural trace on $End_N(H)$. If f is a projection in $N' \cap M$, then

$$\begin{aligned} Tr_{N'}(f) &= \dim_N(fH) && \text{(by definition of } \dim_N) \\ &\ge \dim_{fMf}(fH) && \text{(by 3.4.6)} \\ &= tr_M(f)^{-1} && \text{(by 3.2.5(h))} \\ &\ge 1 \end{aligned}$$

Suppose $N' \cap M$ contains k mutually orthogonal projections $f_1,\cdots,f_k$ with $\sum f_i = \mathbf{1}$. Then

$$[M:N] = Tr_{N'}(\mathbf{1}) = \sum_i Tr_{N'}(f_i)$$
$$\ge \sum_i tr_M(f_i)^{-1} \ge k^2.$$

In particular, if $N' \cap M \ne \mathbb{C}\mathbf{1}$. then $[M:N] \ge 4$, and if $N' \cap M$ is infinite dimensional, then $[M:N] = \infty$. Suppose $[M:N] < \infty$, and let $f_1,\cdots,f_k$ be a *maximal* family of mutually orthogonal projections in $N' \cap M$; then $[M:N] \ge k^2 \ge \dim_{\mathbb{C}}(N' \cap M)$. This proves all the assertions in the case of factors.

Now return to the situation where N and M are finite direct sums of finite factors. The projections p_iq_j are central projections in $N' \cap M$ and $p_iq_j(N' \cap M) = N'_{p_iq_j} \cap M_{p_iq_j}$. So if $\dim_{\mathbb{C}}(N' \cap M) = \infty$ there must be a pair (i,j) for which $\dim_{\mathbb{C}}(N'_{p_iq_j} \cap M_{p_iq_j}) = \infty$. But this contradicts the observation just made for the case of factors, and completes the proof of (a). #

The next results (3.6.3–3.6.5) depend on ideas of Pimsner and Popa [PP1].

Lemma 3.6.3. *Let* $N \subset M$ *be finite direct sums of type* II_1 *factors with* N *of finite index in* M, *and let* tr *be a faithful trace on* M. *If* $x \in \langle M, e_N \rangle$, *there is a unique* $y \in M$ *for which* $xe_N = ye_N$.

Proof. Regard $N \subset M$ represented on $L^2(M)$.

Let us first check uniqueness. Suppose $y, y' \in M$ with $xe_N = ye_N = y'e_N$. If Ω is the trace vector in $L^2(M)$, then

$$(y-y')\Omega = (y-y')e_N\Omega = 0,$$

so $y' = y$ because Ω is separating.

To prove existence, we have to show that $\langle M,e_N\rangle e_N = Me_N$ and we proceed as follows.

As N' is finite, $\langle M,e_N\rangle$ is finite by 3.6.1.iv, and there exists a faithful normal conditional expectation F from $\langle M,e_N\rangle$ onto M (see Proposition II.5 for the proof of this latter fact). We claim that $F(e_N)$ is invertible in M. Since F is an M–M–bimodule map, $F(e_N)$ belongs to $N' \cap M$, which is finite dimensional by Lemma 3.6.2. Consequently, to show that the self-adjoint element $F(e_N)$ is invertible, it is enough to check that $xF(e_N)x \neq 0$ for any positive element $x \neq 0$ in $N' \cap M$. But if

$$0 = xF(e_N)x = F(xe_Nx),$$

then $xe_Nx = 0$, since F is faithful. And $xe_Nx = (e_Nx)^*(e_Nx)$, so $e_Nx = 0$. Hence

$$0 = e_Nxe_N = E_N(x)e_N,$$

which implies $x = 0$ by 3.6.1.v and the faithfulness of E_N. This proves the claim that $F(e_N)$ is invertible.

Now we may obtain a formula for xe_N. Suppose first that x is in Me_NM, namely that x is a finite sum $\sum a_je_Nb_j$ with $a_j, b_j \in M$. Then $F(xe_N) = \sum a_jE_N(b_j)F(e_N)$ and

$$F(xe_N)F(e_N)^{-1}e_N = xe_N.$$

This formula holds for any $x \in \langle M,e_N\rangle$ because both sides are strongly continuous in x and because Me_NM is strongly dense in $\langle M,e_N\rangle$ by Proposition 3.6.1.vii. Thus $xe_N \in Me_N$ for any $x \in \langle M,e_N\rangle$. #

Theorem 3.6.4. *Let* $N \subset M$ *be type* II_1 *von Neumann algebras with finite dimensional centers and let* tr *be a faithful normal trace on* M *for which* N' *is finite on* $L^2(M,tr)$. *Then*

(i) *As a right module over* N, *the algebra* M *is projective of finite type.*

(ii) *The conditional expectation* $E_N : M \to N$ *is very faithful (in the sense of Section 2.6).*

(iii) $\langle M,e_N\rangle = Me_NM := \{\sum_{j=1}^{n} a_je_Nb_j : n \geq 1, a_j,b_j \in M\}$.

(iv) *If* $\alpha : M \to M$ *is a right* N*–module map, then* α *extends uniquely to an element of* $\langle M,e_N\rangle = JN'J$ *on* $L^2(M,tr)$.

(v) *If* $x \in JN'J$ *then* $x(M) \subset M$, *where* M *is viewed as a dense subspace of* $L^2(M,tr)$.

Proof. (i) Any strongly closed right ideal in N is projective of finite type, and in fact of the form pN with p a projection in N. (See [Tak], II.3.12.) We are going to show that M is isomorphic, as a right N–module, to a finite direct sum of such ideals. In the course of doing so we exhibit a basis $\{v_i : 1 \leq i \leq n\}$ of M over N with the following properties:

(a) $E_N(v_i^*v_j) = 0$ if $i \neq j$.

(b) $f_i := E_N(v_i^*v_i)$ is a projection in N, $v_if_i = v_i$, and $E_N(v_i^*x) = f_iE_N(v_i^*x)$, for $1 \leq i \leq n$ and $x \in M$.

(c) Every x in M has a unique expansion

$$x = \sum_{i=1}^{n} v_iy_i, \text{ with } y_i \in N.$$

In fact $v_iy_i = v_iE_N(v_i^*x)$.

Since the central support of e_N in $\langle M,e_N\rangle$ is $\mathbf{1}$ and since $\langle M,e_N\rangle$ is finite with finite dimensional center by 3.6.1(iv), there exists a finite set $w_1,\cdots,w_n$ of partial isometries in $\langle M,e_N\rangle$ with $w_j^*w_j \leq e_N$ and $\sum w_jw_j^* = \mathbf{1}$; in particular the w_j have mutually orthogonal range projections. (See [Tak], V.1.34.) As $w_je_N = w_j$, there are, by 3.6.3, elements $v_1,\cdots,v_n \in M$ with $w_j = v_je_N$ for all j. We verify that the v_i have the properties (a)–(c). For $i \neq j$

$$0 = w_i^*w_j = e_Nv_i^*v_je_N = E_N(v_i^*v_j)e_N,$$

so $E_N(v_i^*v_j) = 0$ by 3.6.1(v). Similarly, since $w_i^*w_i$ is a projection in $\langle M,e_N\rangle$ and $w_i^*w_i = E_N(v_i^*v_i)e_N$, 3.6.1(v) implies that $f_i := E_N(v_i^*v_i)$ is a projection in N.

Furthermore

$$v_i f_i e_N = v_i e_N v_i^* v_i e_N = w_i w_i^* w_i$$
$$= w_i = v_i e_N,$$

so that $v_i f_i = v_i$, by the uniqueness statement of 3.6.3. Therefore, since $f_i \in N$,

$$f_i E_N(v_i^* x) = E_N(f_i v_i^* x) = E_N(v_i^* x) \quad \text{for } x \in M.$$

For any $x \in M$,

$$x e_N = \sum_j w_j w_j^* x e_N = \sum_j v_j e_N v_j^* x e_N$$
$$= \sum_j v_j E_N(v_j^* x) e_N,$$

and hence $x = \sum_j v_j E_N(v_j^* x)$, by 3.6.3. To show uniqueness of the expansion, suppose that $x = \sum v_i y_i$ with $y_i \in N$. Then

$$v_i E_N(v_i^* x) = v_i E_N(v_i^* \sum_j v_j y_j)$$
$$= v_i E_N(v_i^* v_i) y_i$$
$$= v_i f_i y_i = v_i y_i,$$

using N-linearity of E_N and properties (a) and (b) of $\{v_i\}$. We will refer to a family $\{v_i\}$ having properties (a)-(c) as a Pimsner-Popa basis of M over N; see [PP1].

Now consider the N-linear map

$$\Psi \begin{cases} M \to \bigoplus_{1 \le j \le n} f_j N \\ x \mapsto (E_N(v_j^* x)). \end{cases}$$

It follows from the expansion $x = \sum v_j E_N(v_j^* x)$ that Ψ is injective. On the other hand, if $(y_j) \in \bigoplus_j f_j N$ and $x = \sum v_j y_j$ then by the uniqueness of the expansion, $v_j y_j = v_j E_N(v_j^* x)$ for all j. Multiplying both sides on the left by v_j^* and applying E_N gives $f_j y_j = f_j E_N(v_j^* x)$; since both y_j and $E_N(v_j^* x)$ are in $f_j N$, that is $y_j = E_N(v_j^* x)$. Thus $(y_j) = \Psi(x)$ and Ψ is surjective.

(ii) Let $\alpha : M \to N$ be a right N-linear map and set $a = \sum_j \alpha(v_j) v_j^*$. Recall from Section 2.6 that $E_N^b(a) : M \to N$ is defined by $E_N^b(a)(x) = E_N(ax)$ for $x \in M$. We have

$$\begin{aligned} \alpha(x) &= \alpha(\sum v_j E_N(v_j^* x)) \\ &= \sum \alpha(v_j) E_N(v_j^* x) && \text{by N-linearity of } \alpha \\ &= E_N(\sum \alpha(v_j) v_j^* x) && \text{by N-linearity of } E_N \\ &= E_N^b(a)(x), \end{aligned}$$

so that $\alpha = E_N^b(a)$.

(iii) It follows from 3.6.3 that $Me_N M$ is a two-sided ideal in $\langle M, e_N\rangle$. But $Me_N M$ contains $\sum_j v_j e_N v_j^* = \sum_j w_j w_j^* = \mathbf{1}$, so $Me_N M = \langle M, e_N\rangle$.

(iv) If $\alpha : M \to M$ is right N-linear, then for $x \in M$, $\alpha(x) = \alpha \sum_j v_j E_N(v_j^* x) = \sum_j \alpha(v_j) E_N(v_j^* x)$; thus $\alpha = \sum_j \lambda(\alpha(v_j)) \circ E_N \circ \lambda(v_j^*)$, where $\lambda(y)$ denotes left multiplication by y. The unique $\|\cdot\|_2$-continuous extension of α to $L^2(M,tr)$ is $\sum_j \alpha(v_j) e_N v_j^* \in \langle M, e_N\rangle$.

(v) Any $x \in \langle M, e_N\rangle$ is of the form $\sum_j a_j e_N b_j$ by claim (iii). If $y \in M$ then

$$x(y) = \sum a_j E_N(b_j y) \in M. \qquad \#$$

Corollary 3.6.5. *Let* $N \subset M$ *be a pair of von Neumann algebras of type* II_1 *having finite dimensional centers, and suppose that* N *is of finite index in* M. *Let* tr *be any faithful normal trace on* M *and define* e_N *and* E_N *via* tr. *Then*

$$M \otimes_N M \cong \langle M, e_N\rangle \quad \textit{as N-bimodules, and}$$
$$\mathrm{End}_N^r(M) \cong \langle M, e_N\rangle \quad \textit{as } \mathbb{C}\textit{-algebras.}$$

Proof. Since $N \subset M$ has finite index, $\langle M, e_N\rangle$ is finite. The isomorphism $\mathrm{End}_N^r(M) \cong \langle M, e_N\rangle$ follows from 3.6.4(iv) or (v); the correspondence is defined by $\sum \lambda(a_j) E_N \lambda(b_j) \mapsto \sum a_j e_N b_j$.

The isomorphism $M \otimes_N M \cong \mathrm{End}_N^r(M)$ extending the map $a \otimes_N b \mapsto \lambda(a) E_N \lambda(b)$ on elementary tensors follows from 3.6.4(i) and (ii) and 2.6.3. One can also verify directly the isomorphism $M \otimes_N M \cong \langle M, e_N\rangle$ by using a Pimsner–Popa basis. #

The next proposition determines one part of the spatial data for the inclusion $M \subset \langle M, e_N\rangle$.

Proposition 3.6.6. *Let* $N \subset M$ *be finite direct sums of finite factors such that* N *is of finite index in* M, *and let* tr *be any faithful trace on* M. *Then* $\Lambda_M^{\langle M,e_N\rangle} = (\Lambda_N^M)^t$.

Proof. This follows from 3.5.4, 3.5.5, and the formulas $JN'J = \langle M, e_N\rangle$, $JM'J = M$. #

To describe $M \subset \langle M, e_N\rangle$ more precisely, we also have to compute the matrix of traces $T_M^{\langle M,e_N\rangle}$. This is the part of the theory which differs most from the finite dimensional case presented in Chapter 2.

Before proceeding, we summarize our notation: $N \subset M$ is a pair of finite von Neuman algebras with finite dimensional centers, with N of finite index in M; the minimal central projections in M and N are respectively $\{p_i : 1 \le i \le m\}$ and $\{q_j : 1 \le j \le n\}$. A trace tr on M is specified by the row vector $\overleftarrow{s}$, $s_i = \mathrm{tr}(p_i)$. Let $H = L^2(M,\mathrm{tr})$. Set

$$\begin{aligned} N_{i,j} &= p_iq_jNp_iq_j, & M_{i,j} &= p_iq_jMp_iq_j,\\ N'_{i,j} &= p_iq_jN'p_iq_j & &= \mathrm{End}_{N_{i,j}}(p_iq_jH),\\ M'_{i,j} &= p_iq_jM'p_iq_j & &= \mathrm{End}_{M_{i,j}}(p_iq_jH), \end{aligned}$$

when $p_iq_j \ne 0$. We have the trace matrix T_N^M with entries $c_{i,j} = \mathrm{tr}_{p_iM}(p_iq_j)$, and the index matrix Λ_N^M with entries

$$\begin{aligned} \lambda_{i,j} &= 0 & & p_iq_j = 0,\\ \lambda_{i,j} &= [M_{i,j}:N_{i,j}]^{1/2} & & \text{if } p_iq_j \ne 0. \end{aligned}$$

Our present goal is to compute the entries of $T_M^{\langle M,e_N\rangle} = T_{N'}^{M'}$, namely

$$d_{j,i} = \mathrm{tr}_{q_jN'}(q_jp_i).$$

Lemma 3.6.7. *If* $p_iq_j \ne 0$, *then*

(i) $\dim_{N'_{i,j}}(p_iq_jH) = \dfrac{c_{i,j}}{\lambda_{i,j}^2}$, *and*

(ii) $d_{j,i}\dim_{N'_{i,j}}(p_iq_jH) = \dim_{q_jN'}(q_jH)$.

Proof. By 3.4.6,

$$\lambda_{i,j}^2 = \frac{\dim_{N_{i,j}}(p_iq_jH)}{\dim_{M_{i,j}}(p_iq_jH)},$$

and by 3.2.5(h),

$$\dim_{M_{i,j}}(p_iq_jH) = \mathrm{tr}_{p_iM}(q_jp_i)^{-1}\dim_{p_iM}(p_iH).$$

But since M is in standard form on H, so is p_iM on p_iH, and $\dim_{p_iM}(p_iH) = 1$. Combining these observations,

$$\begin{aligned} \lambda_{i,j}^2 &= \dim_{N_{i,j}}(p_iq_jH)\,\mathrm{tr}_{p_iM}(p_iq_j)\\ &= \dim_{N_{i,j}}(p_iq_jH)\,c_{i,j}\\ &= \dim_{N'_{i,j}}(p_iq_jH)^{-1}c_{i,j}, \end{aligned}$$

by 3.2.5.g. Hence (i).

(ii) This reads

$$\dim_{N'_{i,j}}(p_iq_jH) = [\mathrm{tr}_{q_jN'}(p_iq_j)]^{-1}\dim_{q_jN'}(q_jH),$$

which follows from 3.2.5(h). #

Notation: For each j, let

$$\varphi_j = \left[\sum_{i=1}^{m}{}' \frac{\lambda_{i,j}^2}{c_{i,j}}\right]^{-1},$$

the sum being over those i such that $p_iq_j \ne 0$, and let F be the diagonal matrix $F = \mathrm{diag}(\varphi_1,\cdots,\varphi_n)$. Furthermore, let $\tilde{T}$ be the n-by-m matrix

$$(\tilde{T})_{j,i} = 0 \quad \text{if } p_iq_j = 0.$$
$$(\tilde{T})_{j,i} = \frac{\lambda_{i,j}^2}{c_{i,j}} \quad \text{if } p_iq_j \neq 0.$$

Proposition 3.6.8. $T_M^{\langle M,e_N\rangle} = F\tilde{T}$.

Proof. Combining 3.6.7(i) and (ii) we get

$$\text{(3.6.8.a)} \qquad d_{j,i} = \frac{\dim_{q_jN'}(q_jH)}{\dim_{N'_{i,j}}(q_jp_iH)} = \frac{\lambda_{i,j}^2}{c_{i,j}} \dim_{q_jN'}(q_jH),$$

if $p_iq_j \neq 0$, and $d_{j,i} = 0$ otherwise. To eliminate $\dim_{q_jN'}(q_jH)$ we use the fact that $T_M^{\langle M,e_N\rangle}$ is row stochastic,

$$1 = \sum_{i=1}^{m} d_{j,i} = \left[\sum_{i=1}^{m}{}' \frac{\lambda_{i,j}^2}{c_{i,j}}\right] \dim_{q_jN'}(q_jH),$$

so

$$\text{(3.6.8.b)} \qquad \dim_{q_jN'}(q_jH) = \varphi_j.$$

Putting this back in (3.6.8.a) gives $d_{j,i} = \varphi_j \frac{\lambda_{i,j}^2}{c_{i,j}}$ if $p_iq_j \neq 0$ and $d_{j,i} = 0$ otherwise, as desired. #

Remark. Let us check what that formula $T_M^{\langle M,e_N\rangle} = F\tilde{T}$ means for finite dimensional algebras. Suppose that $p_iM \cong \mathrm{Mat}_{\mu_i}(\mathbb{C})$ and $q_jN \cong \mathrm{Mat}_{\nu_j}(\mathbb{C})$. As noted before, the inclusion matrix $\Lambda = \Lambda_N^M$ determines the trace matrix $T = T_N^M$ via

$$c_{i,j} = \lambda_{i,j} \frac{\nu_j}{\mu_i},$$

since q_jp_i is the sum of $\lambda_{i,j}\nu_j$ orthogonal minimal projections in p_iM. Setting $\tilde{\mu} = \mathrm{diag}(\mu_1,\cdots,\mu_m)$ and $\tilde{\nu} = \mathrm{diag}(\nu_1,\cdots,\nu_n)$, this can be written

$$T = \tilde{\mu}^{-1}\Lambda\tilde{\nu}.$$

When $p_iq_j \neq 0$, we have $(\tilde{T})_{j,i} = \frac{\lambda_{i,j}^2}{c_{i,j}} = \lambda_{i,j}\frac{\mu_i}{\nu_j}$; and when $p_iq_j = 0$, $(\tilde{T})_{j,i} = 0 = \lambda_{i,j}\frac{\mu_i}{\nu_j}$. Thus

$$\tilde{T} = \tilde{\nu}^{-1}\Lambda^t\tilde{\mu}.$$

Set $L = \langle M,e_N\rangle$, $L = \bigoplus_{j=1}^{n} L_j$; then $L_j \cong \mathrm{Mat}_{\kappa_j}(\mathbb{C})$ where $\kappa_j = (\Lambda^t\vec{\mu})_j = \sum_i \mu_i\lambda_{i,j}$. Note that

$$\varphi_j = \Big(\sum_i \tilde{T}_{j,i}\Big)^{-1} = \Big(\frac{1}{\nu_j}\sum_i \lambda_{i,j}\mu_i\Big)^{-1} = \frac{\nu_j}{\kappa_j}.$$

Thus

$$(T_M^{\langle M,e_N\rangle})_{j,i} = \varphi_j(\tilde{T})_{j,i} = \frac{\nu_j}{\kappa_j}\frac{\mu_i}{\nu_j}\lambda_{i,j} = \lambda_{i,j}\frac{\mu_i}{\kappa_j},$$

which is in accord with the relation observed above between the inclusion matrix and the index matrix.

We now return to the analysis of the general case.

As the minimal central projection in $\langle M,e_N\rangle = JN'J$ are precisely $\{Jq_jJ : 1 \leq j \leq n\}$, any trace Tr on $\langle M,e_N\rangle$ is specified by a row vector $\overleftarrow{r}$, with $r_j = \mathrm{Tr}(Jq_jJ)$. It will turn out to be useful to calculate the quantities $\mathrm{Tr}(e_NJq_jJ)$. Recall that $Je_N = e_NJ$. Also observe that

$$\text{(3.6.9)} \qquad e_NJq_jJ = e_Nq_j.$$

In fact, let Ω denote the trace vector in $H = L^2(M,\mathrm{tr})$, i.e. the identity $\mathbf{1}$ of M reqarded as an element of H. The linear space $\{x\Omega : x \in M\}$ is dense in H and we have

$$\begin{aligned} e_NJq_jJ(x\Omega) &= e_NJq_jx^*\Omega = e_Nxq_j\Omega \\ &= E_N(xq_j)\Omega = E_N(x)q_j\Omega \\ &= q_jE_N(x)\Omega = E_N(q_jx)\Omega \\ &= e_Nq_j(x\Omega). \end{aligned}$$

Lemma 3.6.10. *Let* Tr *be any trace on* $\langle M,e_N\rangle$ *and let* $r_j = \mathrm{Tr}(Jq_jJ)$. *Then*

(i) $\mathrm{tr}_{q_jN'}(q_je_N) = \varphi_j = \left[\sum_{i=1}^{m}{}' \frac{\lambda_{i,j}^2}{c_{i,j}}\right]^{-1}$.

(ii) $\mathrm{Tr}(e_Nq_j) = \mathrm{Tr}(e_NJq_jJ) = r_j\varphi_j$.

Proof. (i) Since N is in standard form on e_NH, so is q_jN and its commutant $q_je_NN'q_je_N$ on q_je_NH; hence

$$\begin{aligned} 1 &= \dim_{q_je_NN'q_je_N}(q_je_NH) \\ &= [\mathrm{tr}_{q_jN'}(q_je_N)]^{-1}\dim_{q_jN'}(q_jH) \qquad \text{(by 3.2.5(h).)} \\ &= [\mathrm{tr}_{q_jN'}(q_je_N)]^{-1}\varphi_j \qquad \text{(by 3.6.8(b).)} \end{aligned}$$

(ii) Since the map $x \mapsto \mathrm{Tr}(Jx^*J)$ is a trace on the factor q_jN' we have $\mathrm{Tr}(Jx^*J) = \mathrm{Tr}(Jq_jJ)\mathrm{tr}_{q_jN'}(x)$, and in particular, using 3.6.9,

$$\mathrm{Tr}(e_NJq_jJ) = \mathrm{Tr}(Je_Nq_jJ) = \mathrm{Tr}(Jq_jJ)\mathrm{tr}_{q_jN'}(e_Nq_j) = r_j\varphi_j,$$

by part (i). #

3.7. Markov traces on $\mathrm{End}_N(M)$, a generalization of index.

Definition 3.7.1. Let $N \subset M$ be finite von Neumann algebras with N of finite index in M. We say that a faithful trace tr on M is a Markov trace of modulus β for the pair $N \subset M$ if it extends to a trace, also called tr, on $\langle M,e_N\rangle$ for which

(3.7.2) $$\beta\,\mathrm{tr}(xe_N) = \mathrm{tr}(x) \quad \text{for } x \in M.$$

The extension of tr to $\langle M,e_N\rangle$ is uniquely determined by (3.7.2). Also it suffices for (3.7.2) to hold for $x \in N$, since then for $x \in M$

$$\begin{aligned} \mathrm{tr}(xe_N) &= \mathrm{tr}(e_Nxe_N) = \mathrm{tr}(E_N(x)e_N) \\ &= \tfrac{1}{\beta}\mathrm{tr}(E_N(x)) = \tfrac{1}{\beta}\mathrm{tr}(x). \end{aligned}$$

Cf. Lemma 2.7.1.

We restrict our attention to pairs of finite direct sums of finite factors and continue to use the notation of the previous section.

Theorem 3.7.3. *A trace on* M *specified by the vector* $\overleftarrow{s}$, $s_i = \mathrm{tr}(p_i)$ *is a Markov trace of modulus* β *if and only if*

$$\overleftarrow{s}\,T_N^M\vec{T} = \beta\overleftarrow{s}.$$

Proof. ($\Rightarrow$). Suppose Tr is a trace on $\langle M,e_N\rangle$ extending the given trace on M and satisfying the Markov property (3.7.2). Let $\overleftarrow{r}$ be the row vector, $r_j = \mathrm{Tr}(Jq_jJ)$. By the Markov property we have

$$\beta\,\mathrm{Tr}(e_Nq_j) = \mathrm{tr}(q_j) = t_j,$$

where $\overleftarrow{t} = \overleftarrow{s}\,T_N^M$ is the vector specifying $\mathrm{tr}|_N$. Putting this together with 3.6.10(ii) gives

$$t_j = \beta\,\mathrm{Tr}(e_Nq_j) = \beta\, r_j\varphi_j, \text{ or}$$

(3.7.3.1) $$\overleftarrow{t} = \beta\overleftarrow{r}\,F.$$

Hence

$$\begin{aligned} \beta\overleftarrow{s} &= \beta\overleftarrow{r}\,T_M^{\langle M,e_N\rangle} = \beta\overleftarrow{r}\,F\vec{T} \quad \text{(by 3.6.8)} \\ &= \overleftarrow{t}\,\vec{T} \qquad \text{(by 3.7.3.1)} \\ &= \overleftarrow{s}\,T_N^M\vec{T}. \end{aligned}$$

($\Leftarrow$) Given a trace tr on M satisfying $\overleftarrow{s}\,T_N^M\vec{T} = \beta\overleftarrow{s}$, define $\overleftarrow{r} = \beta^{-1}\overleftarrow{s}\,T_N^MF^{-1}$ (motivated by 3.7.3.1), and define a trace Tr on $\langle M,e_N\rangle$ by $\mathrm{Tr}(Jq_jJ) = r_j$. Then

$$\overleftarrow{r}\,T_M^{\langle M,e_N\rangle} = (\beta^{-1}\overleftarrow{s}\,T_N^MF^{-1})(F\vec{T}) = \overleftarrow{s},$$

so Tr extends tr on M (3.5.2(ii)).

It remains to show the Markov property, $\mathrm{Tr}(xe_N) = \beta^{-1}\mathrm{tr}(x)$ for $x \in N$, and by linearity it is enough to check this for $x \in Nq_j$. Now $x \mapsto \mathrm{Tr}(xe_N)$ is a trace on the factor Nq_j, so $\mathrm{Tr}(xe_N) = \mathrm{Tr}(q_je_N)\mathrm{tr}_{Nq_j}(x)$; hence it suffices to show that $\mathrm{Tr}(q_je_N) = \beta^{-1}\mathrm{tr}(q_j) = \beta^{-1}t_j$. But by 3.6.10(ii)

$$\begin{aligned}\mathrm{Tr}(q_j e_N) &= r_j \varphi_j = (\overleftarrow{r}F)_j\\ &= (\beta^{-1}\overleftarrow{s}\,T_N^M F^{-1}F)_j\\ &= \beta^{-1}t_j,\end{aligned}$$

as desired. #

Corollary 3.7.4. *Suppose* $N \subset M$ *are finite direct sums of finite factors, with* N *of finite index in* M. *Set* $T = T_N^M$.

(i) *If* $N \subset M$ *is a connected inclusion, then there is a unique normalized Markov trace on* $N \subset M$; *it is faithful and has modulus equal to the spectral radius of* $T\tilde{T}$.

(ii) *If* tr *is a Markov trace of modulus* β *on* $N \subset M$, *then the unique extension of the trace to* $\langle M,e_N\rangle$ *satisfying (3.7.2) is a Markov trace of modulus* β *(for* $M \subset \langle M,e_N\rangle$*).*

Proof. (i) Since $N \subset M$ is connected, T is indecomposable and $T\tilde{T}$ is irreducible by a straightforward generalization of Lemma 1.3.2.b. Therefore by Perron- Frobenius theory, $T\tilde{T}$ has a unique non-negative eigenvector $\overleftarrow{s}$ with $\sum_i s_i = 1$. Furthermore $s_i > 0$ and the corresponding eigenvalue is the spectral radius of $T\tilde{T}$

(ii) If $\overleftarrow{s}$ is the vector specifying the Markov trace on M, then the extension of the trace to $\langle M,e_N\rangle$ satisfying the Markov condition (3.7.2) is specified by the vector $\overleftarrow{r} = \beta^{-1}\overleftarrow{s}\,TF^{-1}$. Let R denote the matrix $T_M^{\langle M,e_N\rangle} = F\tilde{T}$, with entries

$$R_{j,i} = \begin{cases} \dfrac{\lambda_{i,j}^2}{c_{i,j}}\varphi_j & \text{if } p_i q_j \neq 0\\ 0 & \text{if } p_i q_j = 0.\end{cases}$$

Since $\Lambda_M^{\langle M,e_N\rangle} = (\Lambda_N^M)^t$, the matrix $\tilde{R}$ (which is to R as $\tilde{T}$ is to T) has entries

$$\tilde{R}_{i,j} = \begin{cases} \lambda_{i,j}^2/R_{j,i} = c_{i,j}\varphi_j^{-1} & \text{if } p_i q_j \neq 0\\ 0 & \text{if } p_i q_j = 0.\end{cases}$$

That is $\tilde{R} = TF^{-1}$. But then

$$\begin{aligned}\overleftarrow{r}\,R\tilde{R} &= (\beta^{-1}\overleftarrow{s}\,TF^{-1})(F\tilde{T})(TF^{-1})\\ &= \beta^{-1}\overleftarrow{s}\,T\tilde{T}TF^{-1}\\ &= \overleftarrow{s}\,TF^{-1} \qquad \text{(by 3.7.3)}\\ &= \beta\,\overleftarrow{r}.\end{aligned}$$

Hence $\overleftarrow{r}$ defines a Markov trace on $\langle M,e_N\rangle$ by Theorem 3.7.3. #

Remark. Before going on, let us see how the analysis above agrees with that in Chapter 2 for finite dimensional algebras. Assume that $Mp_i \cong \mathrm{Mat}_{\mu_i}(\mathbb{C})$ and $Nq_j \cong \mathrm{Mat}_{\nu_j}(\mathbb{C})$. We noted in the remark following 3.6.8 that

$$T = \tilde{\mu}^{-1}\Lambda\tilde{\nu} \quad\text{and}\quad \tilde{T} = \tilde{\nu}^{-1}\Lambda^t\tilde{\mu},$$

where $\tilde{\mu} = \mathrm{diag}(\mu_1,\cdots,\mu_m)$ and $\tilde{\nu} = \mathrm{diag}(\nu_1,\cdots,\nu_n)$. Thus

$$T\tilde{T} = \tilde{\mu}^{-1}\Lambda\Lambda^t\tilde{\mu}.$$

In this chapter we have been specifying a trace tr on M by the vector $\overleftarrow{s}$ with $s_i = \mathrm{tr}(p_i)$, while in Chapter 2 we specified the trace by $\overleftarrow{s}'$, where s_i' is the trace of a minimal projection in Mp_i. The vector $\overleftarrow{s}$ and $\overleftarrow{s}'$ are related by $\overleftarrow{s} = \overleftarrow{s}'\tilde{\mu}$. The condition given in Chapter 2 for tr to be a Markov trace of modulus β is $\overleftarrow{s}'\Lambda\Lambda^t = \beta\overleftarrow{s}'$. But this is equivalent to

$$\begin{aligned}\overleftarrow{s}(T\tilde{T}) &= (\overleftarrow{s}'\tilde{\mu})(\tilde{\mu}^{-1}\Lambda\Lambda^t\tilde{\mu})\\ &= \overleftarrow{s}'\Lambda\Lambda^t\tilde{\mu} = \beta\overleftarrow{s}'\tilde{\mu} = \beta\overleftarrow{s}. \qquad \#\end{aligned}$$

Definition 3.7.5. Let $N \subset M$ be finite sums of II_1– factors with the same identity and with N of finite index in M. Let $\Lambda = \Lambda_N^M = (\lambda_{i,j})$ be the matrix of indices and $T = T_N^M = (c_{i,j})$ be the row stochastic matrix of traces as above. Form $\tilde{T} = \tilde{T}(\Lambda,T)$, the matrix whose (j,i) entry is 0 if $c_{i,j} = 0$ and $\dfrac{\lambda_{i,j}^2}{c_{i,j}}$ otherwise. Then the index of N in M, $[M:N]$, is the largest eigenvalue of the matrix $T\tilde{T}$.

Remark. It is easy to see that this definition agrees with that of Section 3.4 when N and M are factors. We mention again that P. Jolissaint has recently shown that this definition always coincides with the ring theoretic definition given in Section 2.1 and [Jo4].

Corollary 3.7.6. *If* $N \subset M$ *are as above and* $[M:N] < 4$, *then* $[M:N] \in \{4\cos^2\pi/q : q \geq 3\}$.

Proof. The index is the largest of the numbers $[Mz:Nz]$, where z is a minimal projection in $Z(M) \cap Z(N)$, so we can assume that $M \supset N$ is connected.

By 3.7.4(i), there is a Markov trace tr on M of modulus [M:N].

Then 3.7.4(ii) allows us to iterate the fundamental construction in the usual way to obtain a tower

$$M_0 = N \subset M_1 = M \subset \cdots \subset M_k \subset M_{k+1} \subset \cdots,$$

a sequence of self-adjoint projections $(e_k)_{k\geq 1}$ with $M_{k+1} = \langle M_k, e_k\rangle$ for all k, and a trace tr on $\bigcup_k M_k$ satisfying the Markov property

$$[M:N]\mathrm{tr}(e_k x) = \mathrm{tr}(x) \text{ for } x \in M_k.$$

The projections e_k then satisfy the usual relations and therefore the restriction on [M:N] follows from [Jo1]; see the argument given in Section 3.4. #

Next we provide some examples. Note that by 3.5.6, to construct examples it suffices to give the matrices Λ and T.

Example 3.7.7. The simplest new example is where M is a II_1-factor, p is a projection in M and $N = pMp + (1-p)M(1-p)$. Here the matrix Λ_N^M is (1 1), and T_N^M is (t 1-t), where $t = \mathrm{tr}_M(p)$. Thus $\tilde{T} = \begin{bmatrix} 1/t \\ 1/1-t \end{bmatrix}$ and $T\tilde{T} = 2$. So [M:N] = 2, independent of t! #

Example 3.7.8. Consider an inclusion $N \subset M$ with $\Lambda = \Lambda_N^M = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $T = T_N^M = \begin{bmatrix} t & 1-t \\ 0 & 1 \end{bmatrix}$. Then $\tilde{T} = \begin{bmatrix} 1/t & 0 \\ 1/1-t & 1 \end{bmatrix}$ and $T\tilde{T} = \begin{bmatrix} 2 & 1-t \\ 1/1-t & 1 \end{bmatrix}$. The characteristic equation is $\lambda^2 - 3y + 1 = 0$, so $[M:N] = 4\cos^2\pi/5$, independent of t. #

Example 3.7.9. Take $\Lambda_N^M = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $T_N^M = \begin{bmatrix} a & 1-a \\ b & 1-b \end{bmatrix}$, with $0 < a,b < 1$. Then

$$\tilde{T} = \begin{bmatrix} 1/a & 1/b \\ 1/1-a & 1/1-b \end{bmatrix} \text{ and } T\tilde{T} = \begin{bmatrix} 2 & \frac{a}{b} + \frac{1-a}{1-b} \\ \frac{b}{a} + \frac{1-b}{1-a} & 2 \end{bmatrix}.$$ The characteristic polynomial is

$$(\lambda - 2 - \sqrt{2+\alpha+\alpha^{-1}})(\lambda - 2 + \sqrt{2+\alpha+\alpha^{-1}}),$$

with $\alpha = \frac{a}{b}\frac{(1-b)}{(1-a)}$. So $[M:N] = 2 + \sqrt{2+\alpha+\alpha^{-1}}$, which can be any real number greater than or equal to 4. #

Example 3.7.10. Let M be a II_1-factor, p a projection of trace t in M and $N = pMp + Q$, where Q is a subfactor of index λ in $(1-p)M(1-p)$. Then $\Lambda_N^M = (1\ \lambda^{1/2})$ and $T_N^M = (t\ 1-t)$. So $\tilde{T} = \begin{bmatrix} 1/t \\ \lambda/1-t \end{bmatrix}$ and $T\tilde{T} = 1 + \lambda$. This is ≤ 4 when $\lambda = 1, 2, 4\cos^2\pi/5$, or 3.

Remark. The index matrices in example 3.7.10 correspond to A_3, B_3, H_3, and $G_2^{(1)}$, respectively, under the corespondence of Theorem 1.1.3, when $\lambda = 1, 2, 4\cos^2\pi/5$, and 3. This is no accident, as we will see.

Proposition 3.7.11. *Let $\Lambda = (\lambda_{i,j})$ be an irredundant matrix over $\{2\cos(\pi/q) : q \geq 2\}$ and $T = (c_{i,j})$ is a row stochastic matrix with the same pattern of zero entries as Λ. Let $\tilde{T} = \tilde{T}(\Lambda,T)$ be the matrix whose (j,i)-entry is zero if $c_{i,j} = 0$ and equal to $\frac{\lambda_{i,j}^2}{c_{i,j}}$ otherwise.*

If the spectral radius of $T\tilde{T}$ is less than 4, then it equals $4\cos^2\pi/q$ for some $q \in \{3,4,5\cdots\}$.

Proof. We can suppose that Λ and T are indecomposable. By 3.5.6, there is a connected inclusion $M \supset N$ of finite direct sums of II_1-factors with $\Lambda = \Lambda_N^M$ and $T = T_N^M$. Thus the result is a corollary of 3.7.6. #

Remark. It would be interesting to find a proof of 3.7.11 within usual matrix theory; hopefully this might give information on the spectral radius of $T\tilde{T}$ even when it is larger than 4.

Proposition 3.7.12.

(a) *Let Λ be an irredundant m-by-n matrix with non-negative real values. Then there is a row stochastic m-by-n matrix T with the same pattern of zeros such that*

$$\rho(T\tilde{T}) = \|\Lambda\|^2,$$

where ρ denotes spectral radius and $\tilde{T} = \tilde{T}(\Lambda,T)$ is as above.

(b) *If Λ is irredundant with values in $\{2\cos(\pi/q) : q \geq 2\}$, then there is a pair $N \subset M$ of finite direct sums of II_1-factors with $\Lambda = \Lambda_N^M$ and $[M:N] = \|\Lambda\|^2$.*

(c) *If Λ is any non-zero matrix over $\{2\cos(\pi/q) : q \geq 2\}$ and $\|\Lambda\| < 2$, then*

$$\|\Lambda\| \in \{2\cos(\pi/q) : q \geq 3\}.$$

Proof. (a) As Λ is irredundant, we can define a row stochastic matrix $T = (c_{i,j})$ by

$c_{i,j} = \left[\sum_j \lambda_{i,j}\right]^{-1}\lambda_{i,j}$, or $T = X\Lambda$, where X is the m-by-m diagonal matrix whose (i,i)-entry is $\left[\sum_j \lambda_{i,j}\right]^{-1}$. Then the (j,i) entry of $\tilde{T}$ is $\lambda_{i,j}\left[\sum_j \lambda_{i,j}\right]$, i.e., $\tilde{T} = \Lambda^t X^{-1}$. Thus $T\tilde{T} = X\Lambda\Lambda^t X^{-1}$, which has the same spectrum as $\Lambda\Lambda^t$.

(b) By 3.5.6 there is a pair $N \subset M$ with $\Lambda = \Lambda_N^M$ and $T = T_N^M$. Then $[M:N] = \rho(T\tilde{T}) = \|\Lambda\|^2$.

(c) It suffices to consider Λ irredundant, so the result follows from (b) and 3.7.6. #

Of course, 3.7.12(c) was already known as a consequence of Theorem 1.1.3. Theorem 1.1.3 suggests (but does not immediately imply) the following, which is the main result of this section.

Theorem 3.7.13. *Let* $N \subset M$ *be a connected inclusion of finite direct sums of* II_1*-factors.*

(a) *If* $[M:N] < 4$, *then* Λ_N^M *is the matrix associated (in Theorem 1.1.3) to a bicoloration of one of the following Coxeter graphs:*

$$A_\ell\ (\ell \geq 2),\ B_\ell\ (\ell \geq 3),\ D_\ell\ (\ell \geq 4),\ E_\ell\ (\ell = 6,7,8),$$
$$F_4,\ G_2,\ H_\ell\ (\ell = 3,4),\ I_2(p)\ (p = 5 \text{ or } p \geq 7).$$

Moreover $[M:N] = \|\Lambda_N^M\|^2 = 4\cos^2\pi/h$, *where* h *is the Coxeter number. (See tables 1.4.5, 1.4.6, and 1.4.7.)*

(b) *If* $[M:N] = 4$, *then* Λ_N^M *corresponds to one of*

$$A_\ell^{(1)}\ (\ell \text{ odd}, \ell \geq 1),\ B_\ell^{(1)}\ (\ell \geq 2),\ C_\ell^{(1)}\ (\ell \geq 3),$$
$$D_\ell^{(1)}\ (\ell \geq 4),\ E_\ell^{(1)}\ (\ell = 6,7,8),\ F_4^{(1)},\ G_2^{(1)}.$$

Lemma 3.7.14. (Schwenck, [Sch2]) *Let* $A = (a_{i,j})$ *be a non-negative* n-*by*-n *matrix and let* $G = (g_{i,j})$ *be the matrix with entries* $g_{i,j} = (a_{i,j}a_{j,i})^{1/2}$. *Then* $\|G\| = \rho(G) \leq \rho(A)$, *where* ρ *denotes spectral radius.*

Proof. For any k+1-tuple $\alpha = (i_1, i_2, \cdots, i_{k+1})$ with $1 \leq i_j \leq n$, let α^{-1} denote the reversed tuple $\alpha^{-1} = (i_{k+1}, \cdots, i_2, i_1)$, and let

$$a_\alpha = a_{i_1,i_2}a_{i_2,i_3}\cdots a_{i_k,i_{k+1}}.$$

Let W_k be the set of k+1-tuples with $i_1 = i_{k+1}$; thus $\alpha \mapsto \alpha^{-1}$ is bijection of W_k and

$$\mathrm{tr}(A^k) = \sum_{\alpha\in W_k} a_\alpha = \tfrac{1}{2}\sum_{\alpha\in W_k}(a_\alpha + a_{\alpha^{-1}}).$$

Thus

$$\mathrm{tr}(A^k) \geq \sum_{\alpha\in W_k}(a_\alpha a_{\alpha^{-1}})^{1/2} = \sum_{\alpha\in W_k} g_\alpha = \mathrm{tr}(G^k),$$

for all $k \in \mathbb{N}$. When k is even, we have

$$n\rho(A)^k = n\rho(A^k) \geq \mathrm{tr}(A^k) \geq \mathrm{tr}(G^k) \geq \rho(G)^k,$$

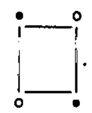
where the first equality and last inequality result from considering canonical forms for A and G, noting that the eigenvalues of G^k are positive. Taking k^{th} roots and then the limit as $k \to \infty$ gives the result. #

Lemma 3.7.15. *Let* $\Lambda = (\lambda_{i,j})$ *be an* m-*by*-n *irredundant matrix over* $\{r \in \mathbb{R} : r = 0 \text{ or } r \geq 1\}$. *Let* $T = (c_{i,j})$ *be a row stochastic matrix with the same pattern of zero entries as* Λ. *Let* $\tilde{T}$ *be the* n-*by*-m *matrix whose* (j,i)-*entry is* 0 *if* $c_{i,j} = 0$ *and* $\frac{\lambda_{i,j}^2}{c_{i,j}}$ *otherwise.*

If $\rho(T\tilde{T}) \leq 4$ *then* $\|\Lambda\|^2 \leq \rho(T\tilde{T})$.

Proof. We may assume without loss of generality that Λ is indecomposable.

Suppose that there exist indices i_1, i_2, j_1, j_2 such that the four entries λ_{i_μ, j_ν} for $\mu,\nu \in \{1,2\}$ are all non-zero; that is, the graph $\tilde{\Gamma}(\Lambda)$ contains a subgraph of the form



Rearranging the rows and columns, we can suppose $i_1 = j_1 = 1$ and $i_2 = j_2 = 2$. Denote the (i,j)-entry of $T\tilde{T}$ by $\gamma_{i,j}$. Then $\gamma_{i,j} = \sum_k{}' \frac{c_{i,k}\lambda_{j,k}^2}{c_{j,k}}$, the sum being over those k for which $c_{j,k} \neq 0$. In particular $\gamma_{i,j} \geq \sum_k{}' \frac{c_{i,k}}{c_{j,k}}$ with equality everywhere if and only if all the non-zero $\lambda_{j,k}$ are equal to one.

By monotonicity of the Perron-Frobenius eigenvalue, we have

$$\rho(T\tilde{T}) \geq \rho\left(\begin{bmatrix} \gamma_{1,1} & \gamma_{1,2} \\ \gamma_{2,1} & \gamma_{2,2} \end{bmatrix}\right),$$

with equality if and only if Λ is 2-by-n. This in turn is no smaller than

$$\rho\left(\left[\sum_k{}' \frac{c_{i,k}}{c_{j,k}}\right]_{1\leq i,j\leq 2}\right),$$

by the observation above, with equality if and only if all the non-zero $\lambda_{j,k}$ with $j \leq 2$ are equal to one. Truncating the sums defining the entries of the last matrix we see that the spectral radius is at least

$$\rho\left(\begin{bmatrix} 2 & \frac{c_{1,1}}{c_{2,1}} + \frac{c_{1,2}}{c_{2,2}} \\ \frac{c_{2,1}}{c_{1,1}} + \frac{c_{2,2}}{c_{1,2}} & 2 \end{bmatrix}\right),$$

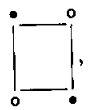
with equality if and only if $\lambda_{j,k} = 0$ for $j = 1,2$ and $k > 2$. If we replace the off-diagonal entries by their geometric mean, we do not alter the spectrum, so the last quantity is equal to

$$\rho\left(\begin{bmatrix} 2 & (2+\alpha+\alpha^{-1})^{1/2} \\ (2+\alpha+\alpha^{-1})^{1/2} & 2 \end{bmatrix}\right),$$

where $\alpha = \frac{c_{1,1}c_{2,2}}{c_{2,1}c_{1,2}}$. Finally, this is at least

$$\rho\left(\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\right) = 4,$$

with equality if and only if $\alpha = 1$. But since $\rho(T\tilde{T}) \leq 4$ by hypothesis, we must have equality at every step: Λ and T are 2-by-2 with $\Lambda = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $\alpha = 1$. Since T is row-stochastic this implies $T = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$, and $\|\Lambda\|^2 = \rho(T\tilde{T}) = 4$.

If on the other hand $\tilde{\Gamma}(\Lambda)$ contains no subgraph of the form , then for $i \neq j$ there is at most one non-zero term in the sum defining $\gamma_{i,j}$,

$$\gamma_{i,j} = \begin{cases} 0 & \text{if the } i^{th} \text{ and } j^{th} \text{ rows of } \Lambda \text{ are orthogonal} \\ \frac{c_{i,k_0}}{c_{j,k_0}} \lambda^2_{j,k_0} & \text{for some } k_0, \text{ otherwise.} \end{cases}$$

Observe that k_0 depends on (i,j), and $k_0(i,j) = k_0(j,i)$. On the other hand, $\gamma_{i,i} = \sum_k \lambda^2_{i,k}$. Note that for all pairs (i,j), the (i,j) entry of $\Lambda\Lambda^t$ is the geometric mean of the (i,j) and (j,i)-entries of $T\tilde{T}$; i.e.,

$$(\Lambda\Lambda^t)_{i,j} = \begin{cases} \sum_k \lambda^2_{i,k} & \text{if } i = j \\ 0 & \text{if the } i^{th} \text{ and } j^{th} \text{ rows of } \Lambda \text{ are orthogonal} \\ \lambda_{i,k_0}\lambda_{j,k_0} & \text{otherwise.} \end{cases}$$

Hence by Lemma 3.7.14

$$\|\Lambda\|^2 = \rho(\Lambda\Lambda^t) \leq \rho(T\tilde{T}). \quad \#$$

Proof of Theorem 3.7.13. (a) Let $T = T_N^M$ and $\tilde{T} = \tilde{T}(\Lambda,T)$. By hypothesis $[M:N] = \rho(T\tilde{T}) < 4$, so by 3.7.15 we have $\|\Lambda\|^2 \leq \rho(T\tilde{T})$. Let $\mathcal{S}$ be the (convex) set of all row-stochastic matrices of the same dimension and zero-pattern as Λ and T. For each $S \in \mathcal{S}$ define $\tilde{S} = \tilde{T}(\Lambda,S)$, the matrix whose (j,i)-entry is 0 if $\lambda_{i,j} = 0$ and $\lambda^2_{i,j}/(S)_{i,j}$ otherwise, and $\varphi(S) = \rho(S\tilde{S})$. Then φ is a continuous function of S by elementary perturbation theory, and φ assumes the value $[M:N] = \rho(T\tilde{T})$ as well as the value $\|\Lambda\|^2$, by 3.7.12(a). But by 3.7.11, the set of values of φ less than 4 is discrete, so by convexity of $\mathcal{S}$, φ is constant and

$$\|\Lambda\|^2 = \rho(T\tilde{T}) = [M:N].$$

The classification of Λ_N^M then follows from Theorem 1.1.3.

(b) We have $\|\Lambda\|^2 \leq \rho(T\tilde{T}) = 4$, by 3.7.15. If $\|\Lambda\|^2 < 4$, then the connectedness argument of part (a) would imply that $\|\Lambda\|^2 = \rho(T\tilde{T}) < 4$, a contradiction. Thus $\|\Lambda\|^2 = 4$, and the classification follows from 1.1.3 again. #

CHAPTER 4

Commuting sqares, subfactors, and the derived tower

4.1. Introduction.

There are two main themes in this chapter. The first is the approximation of a pair $N \subset M$ of hyperfinite II_1 factors by pairs $C_n \subset B_n$ of finite dimensional von Neumann algebras, with

$$\begin{array}{ccc} N & \subset & M \\ \cup & & \cup \\ C_{n+1} & \subset & B_{n+1} \\ \cup & & \cup \\ C_n & \subset & B_n \end{array}$$

and $M = (\cup B_n)''$, $N = (\cup C_n)''$. In order for the approximating "ladder" of finite dimensional algebras to behave well with respect to the fundamental construction and the index, it should behave well with respect to the conditional expectations: $E_N|_{B_n}$ should be the conditional expectation of B_n onto C_n. We are thus led to the following definition which was first introduced by Popa (Lemma 1.2.2 in [Pop1]; see also [Pop2]):

Definition 4.1.1. A diagram

$$\begin{array}{ccc} C_1 & \subset & B_1 \\ \cup & & \cup \\ C_0 & \subset & B_0 \end{array}$$

of finite von Neumann algebras with a finite faithful normal trace tr on B_1 is a commuting square if the diagram

$$\begin{array}{ccc} C_1 & \xleftarrow{E_{C_1}} & B_1 \\ \cup & & \cup \\ C_0 & \xleftarrow[E_{C_0}]{} & B_0 \end{array}$$

commutes.

In Section 4.2, we study the notion of commuting squares and give a number of examples of constructions which produce commuting squares. In particular we consider the behavior of commuting squares under the fundamental construction.

Proposition 4.1.2. *Consider a commuting square* $\begin{array}{ccc} C_1 & \subset & B_1 \\ \cup & & \cup \\ C_0 & \subset & B_0 \end{array}$ *with respect to a trace* tr *which is a Markov trace for the pair* $B_0 \subset B_1$ *of finite von Neumann algebras with finite dimensional centers. Let* $B_2 = \langle B_1, e_1 \rangle$ *be the von Neumann algebra obtained via the fundamental construction for* $B_0 \subset B_1$, *and let* $C_2 = \{C_1, e_1\}''$. *Then*

$$\begin{array}{ccc} C_2 & \subset & B_2 \\ \cup & & \cup \\ C_1 & \subset & B_1 \end{array}$$

is also a commuting square.

Therefore iterating the fundamental construction will produce an infinite ladder of commuting squares. Now suppose that we have a pair $N \subset M$ of finite von Neumann algebras with a Markov trace tr of modulus β and a ladder of commuting squares

$$\begin{array}{ccc} N & \subset & M \\ \cup & & \cup \\ C_{n+1} & \subset & B_{n+1} \\ \cup & & \cup \\ C_n & \subset & B_n \end{array}$$

with $N = (\cup C_n)''$ and $M = (\cup B_n)''$. Let $\langle M, e\rangle$ be the result of the fundamental construction for $N \subset M$ and set $A_n = \langle B_n, e\rangle$ for each $n \geq 1$. We show that the algebras $(A_n)_{n \geq 0}$ generate $\langle M, e\rangle$, and $\begin{array}{ccc} M & \subset & \langle M, e\rangle \\ \cup & & \cup \\ B_n & \subset & A_n \end{array}$ is a commuting square with respect to the Markov extension of the trace to $\langle M, e\rangle$.

In Section 4.3 we prove a theorem of H. Wenzl on pairs $N \subset M$ generated by a ladder of commuting squares satisfying a periodicity assumption. (See Section 4.3, Hypothesis (B), for the exact assumption.)

<u>Theorem 4.1.3.</u> *Suppose* $N \subset M$ *is a hyperfinite pair (with a finite faithful trace* tr *on* M*) generated by a ladder of commuting squares*

$$\begin{array}{ccc} C_{n+1} & \subset & B_{n+1} \\ \cup & & \cup \\ C_n & \subset & B_n \end{array}$$

of finite dimensional von Neumann algebras. Suppose that the inclusion data for the ladder is periodic, in the sense of Hypothesis B of Section 4.3. Then

(i) N and M *are factors and* $[M:N] < \infty$.

Let e *and* $(A_n)_{n \geq 0}$ *be as above and let* z_n *be the central support of* e *in* A_n.

(ii) *For large* n, $z_n = \mathbf{1}$. *Equivalently,* A_n *is isomorphic to the result of the fundamental construction for* $C_n \subset B_n$.

(iii) *For large* n, $[M:N] = [B_n:C_n] = \|\vec{t}^{(n)}\|^2 / \|\vec{s}^{(n)}\|^2$, *where* $\vec{t}^{(n)}$ *and* $\vec{s}^{(n)}$ *are the vectors of the trace on* C_n *and* B_n *respectively.*

Section 4.4 contains a contruction of (necessarily irreducible) pairs of hyperfinite factors with index less than 4, as follows: Start with a connected pair $C_0 \subset B_0$ of finite dimensional von Neumann algebras with index $\beta < 4$, and let $B_1 = \langle B_0, e_{C_0} \rangle$ be the result of the fundamental construction for $C_0 \subset B_0$, with respect to the Markov trace of modulus β on B_0. Define $q \in \mathbf{T}$ by $\beta = 2 + q + q^{-1}$, set $g = qe_{C_0} - (\mathbf{1} - e_{C_0})$, a unitary element in B_1, and set $C_1 = gB_0g^{-1}$. Then

$$\begin{array}{ccc} C_1 & \subset & B_1 \\ \cup & & \cup \\ C_0 & \subset & B_0 \end{array}$$

is a commuting square, with respect to the Markov extension of the trace to B_1. Let $(B_n)_{n \geq 0}$ be the tower obtained by iterating the fundamental construction, with $B_{n+1} = \langle B_n, e_n \rangle$ and set $C_{n+1} = \mathrm{alg}\{C_n, e_n\}$ for $n \geq 1$. Then

$$\begin{array}{ccc} C_{n+1} & \subset & B_{n+1} \\ \cup & & \cup \\ C_n & \subset & B_n \end{array}$$

is a ladder of commuting squares, with respect to the Markov trace on $\cup B_n$. It turns out that the inclusion data is periodic and that $B = (\cup B_n)'' \supset C = (\cup C_n)''$ is a pair of factors with index β.

Let $(M_k)_{k \geq 0}$ be the tower obtained from a pair $N \subset M$ of finite von Neumann algebras with finite dimensional centers, with index $\beta < 4$ and let $(e_k)_{k \geq 1}$ be the usual sequence of projections in the tower construction. We already know another construction of an irreducible pair with index β, namely $\{e_1, e_2, \cdots\}'' \supset \{e_2, e_3, \cdots\}''$ (Theorem 3.4.3). An argument due to C. Skau shows that $\{e_1, e_2, \cdots\}' \cap (\cup M_k)'' = N$ (Theorem 4.4.3).

In Section 4.5 we present a construction which yields *irreducible* pairs of hyperfinite II_1 factors, starting with a Coxeter graph Γ of type A, D, or E and a choice of a distinguished vertex w_1 on Γ. In particular for $\Gamma = E_6$ and w_1 an end vertex on one of the long arms of Γ, we obtain the index value $3 + \sqrt{3}$, which is at present the smallest known value larger than 4 of the index of an irreducible pair. The construction goes as follows.

Give Γ the bicoloration with w_1 white and with r white vertices altogether. Let M_0 denote the abelian von Neumann algebra $\mathbb{C}^r$ and M_1 the finite dimensional von Neumann algebra containing M_0 such that Γ is the Bratteli diagram of the inclusion $M_0 \subset M_1$. Form the tower $(M_j)_{j \geq 0}$, starting with the pair $M_0 \subset M_1$ and the Markov trace tr on M_1, and let $(e_j)_{j \geq 1}$ be the usual sequence of projections. Let M be the factor $(\bigcup_{j \geq 0} M_j)''$ and let N be the subfactor of M generated by $\mathbf{1}$ and the e_j's. Set $\beta = [M_1:M_0] = \|\Gamma\|^2$; since $\beta < 4$, Skau's Lemma 4.4.3 applies and $M \cap N' = M_0$. Let p be the minimal projection of M_0 corresponding to the vertex w_1 and set $C = pN$ and $B = pMp$. Then $C \subset B$ is a pair of factors with $C' \cap B = p(N' \cap M)p = pM_0p = \mathbb{C}p$; that is $C \subset B$ is irreducible. The index of this pair can be computed as follows:

Let Γ also denote the matrix of the bicolored graph Γ, and let $\vec{\xi}$ denote the unique Perron–Frobenius row vector for $\Gamma^t\Gamma$, normalized so that its first co-ordinate (corresponding to the distinguished white vertex w_1) is 1. Let Λ be the Coxeter graph of type A with the same Coxeter number as Γ, and with a bicoloration having at least one white end vertex, which is labelled as the first white vertex. Denote also by Λ the matrix of the graph Λ, and let $\vec{\eta}$ be the Perron–Frobenius row vector of $\Lambda^t\Lambda$, normalized so that its first co-ordinate (corresponding to the chosen white end vertex) is 1. then $[B:C] = \|\vec{\eta}\|^2 / \|\vec{\xi}\|^2$. The proof uses Wenzl's index formula from Section 4.3.

The second main topic of Chapter 4, presented in Sections 4.6 and 4.7, is the <u>derived tower</u> and <u>principal graph</u> of a pair of finite factors $N \subset M$ of finite index. The <u>derived tower</u> is the chain of relative commutants $(N' \cap M_k)_{k \geq 0}$, where $(M_k)_{k \geq 0}$ is the tower for the pair $N \subset M$. It follows from 3.6.2 that $\dim (N' \cap M_k) \leq [M:N]^k$ for $k \geq 0$.

Let $(e_k)_{k \geq 0}$ be the projections in the tower construction, let Y_k denote $N' \cap M_k$ and Λ_k the inclusion matrix for $Y_k \subset Y_{k+1}$. The following summarizes the structure of the derived tower.

Theorem 4.1.4.

(i) *The inclusion* $Y_k \subset Y_{k+1}$ *is connected.*

(ii) $Y_k e_k Y_k$ *is an ideal in* Y_{k+1}, *and if* $z_{k+1} = z(e_k)$ *is the corresponding central projection in* Y_{k+1} *then the homomorphism* $\begin{cases} Y_k \longrightarrow Y_k e_k Y_k \\ x \longmapsto x z_{k+1} \end{cases}$ *has inclusion matrix* Λ_{k-1}^t.

(iii) *For all* k, $\|\Lambda_k\|^2 \le [M:N]$.

(iv) *For* $k \ge 2$, *if* $x \in Y_{k+1}$ *and* $x(Y_k e_k Y_k) = 0$, *then* $x(Y_{k-1}e_{k-1}Y_{k-1}) = 0$.

(v) *For all* $k \ge 1$, *the following are equivalent:*

(a) $Y_k e_k Y_k = Y_{k+1}$.

(b) $\vec{s}^{(k-1)}\Lambda_{k-1}^t\Lambda_{k-1} = [M:N]\,\vec{s}^{(k-1)}$, *where* $\vec{s}^{(k-1)}$ *is the vector of the trace on* Y_{k-1}.

(c) $\Lambda_k = \Lambda_{k-1}^t$.

(d) $\|\Lambda_{k-1}\|^2 = [M:N]$.

(vi) *If the equivalent conditions of (v) hold for* k, *then they also hold for* k+1.

We call the ideal $Y_k e_k Y_k$ "the old stuff", since it is determined by $Y_{k-1} \subset Y_k$; the complementary ideal is called "the new stuff". Then (iv) says that "the new stuff comes only from the old new stuff", or $(xz_k)(1-z_{k+1}) = 0$. The principal graph of the pair $N \subset M$ is obtained as follows: on the Bratteli diagram of the derived tower, delete on each level the vertices corresponding to the old stuff, and the edges emanating from them; the result is a connected bipartite graph with a distinguished vertex $*$, the unique vertex on level 0. The Bratteli diagram of the derived tower can be reconstructed from the principal graph. The pair $N \subset M$ is said to be of finite depth if the principal graph is finite; the depth is the maximum distance from any vertex to $*$.

This analysis, together with the work of Chapter 1, yields a new proof of the restriction on index values:

Corollary. (i) *Suppose* $N \subset M$ *is a pair of* II_1 *factors with* $[M:N] < 4$. *Then*

(a) $[M:N] = 4\cos^2 \pi/h$ *for some integer* $h \ge 3$.

(b) *The depth of* $N \subset M$ *is no greater than* h–2.

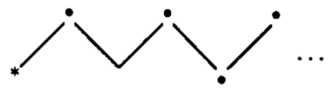
(c) *The principal graph of* $N \subset M$ *is a Coxeter graph of type* A, D, *or* E, *whose norm is* $[M:N]^{1/2}$.

(ii) *Suppose* $N \subset M$ *is a pair of* II_1 *factors with* $[M:N] = 4$.

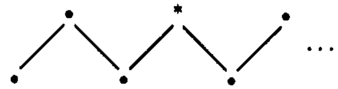
(a) *If* $N \subset M$ *is of finite depth, then the principal graph* Γ *is a completed Coxeter graph of type* A, D, *or* E, *i.e., one of the graphs in Table 1.4.6.*

(b) *If* $N \subset M$ *is of infinite depth, then* Γ *is one of the following:*

A_∞

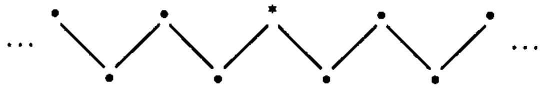


$A_{n,\infty}$



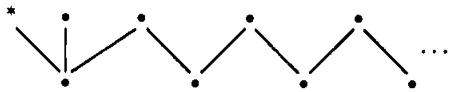
(end vertex at distance n *from* $*$*)*

$A_{\infty,\infty} = \mathbf{T}$

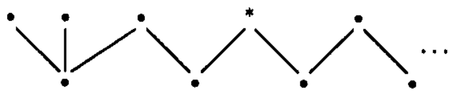


(doubly infinite linear graph)

D_∞



$D_{n,\infty}$



(end vertex at distance n *from* $*$*)*

Section 4.7 is devoted to computing the derived tower for a number of examples: Crossed-products and fixed point algebras for outer actions of finite groups give examples with depth 2. The pairs $R_\beta \subset R$ (of Proposition 3.4.4) when $\beta < 4$ have principal graphs of type A_n; for $\beta = 4$ the principal graph is A_∞. In 4.7.c we give a general method which allows the computation of the derived tower in many examples coming from group actions. In 4.7.d we use this method to obtain the derived towers for the index 4 subfactors $R^G \subset (R \otimes \mathrm{Mat}_2(\mathbb{C}))^G$, where the hyperfinite II_1 factor R is realized as the weak closure of the CAR algebra $\overset{\infty}{\otimes}\mathrm{Mat}_2(\mathbb{C})$ in the trace representation, and G is a closed subgroup of SU(2) acting by the infinite tensor product of its action by conjugation on $\mathrm{Mat}_2(\mathbb{C})$. In this way one obtains as principal graphs all the affine Coxeter graphs of type A, D, and E, as well as the infinite graphs A_∞, $A_{\infty,\infty}$, and D_∞ listed above. Finally we compute the derived tower for the pair $R_\beta \subset R$ when $\beta > 4$. This is the most difficult result of the chapter, involving a representation of the sequence $(e_i)_{i \ge 1}$ in the CAR algebra due to Pimsner and Popa and a theorem of Popa on the tunnel construction (a mirror image of the tower construction). Ultimately one identifies the pair $R_\beta \subset R$ with the pair $N^{\mathbf{T}} \subset (N \otimes \mathrm{Mat}_2(\mathbb{C}))^{\mathbf{T}}$ where N is the completion of the CAR algebra with respect to a certain Powers state. The principal graph is therefore $A_{\infty,\infty}$.

4.2. Commuting squares.

We begin with a proposition, inspired by Lemma 2.1 of [Pop2], which gives a number of equivalent conditions for a commuting square.

Proposition 4.2.1. *Consider a diagram*

$$\begin{matrix} C_1 & \subset & B_1 \\ \cup & & \cup \\ C_0 & \subset & B_0 \end{matrix}$$

of finite von Neumann algebras and a finite faithful normal trace tr *on* B_1. *All conditional expectations being with respect to* tr, *the following are equivalent.*

(i) $E_{C_1}(B_0) \subset C_0$.

(ii) $E_{C_1}E_{B_0} = E_{C_0}$.

(iii) $E_{C_1}E_{B_0} = E_{C_0}E_{B_0}$.

(iv) $E_{C_1}E_{B_0} = E_{B_0}E_{C_1}$ *and* $B_0 \cap C_1 = C_0$.

(v) *The diagram*
$$\begin{tikzcd} C_1 & B_1 \arrow[l, "E_{C_1}"'] \\ C_0 \arrow[u, phantom, "\cup"] & B_0 \arrow[l, "E_{C_0}"] \arrow[u, phantom, "\cup"] \end{tikzcd}$$
commutes.

(vi) $E_{C_0}(b_0c_1) = E_{C_0}(b_0)E_{C_0}(c_1)$ *for* $b_0 \in B_0$ *and* $c_1 \in C_1$.

(vii) $E_{C_0}(b_0c_1) = 0$ *for* $b_0 \in B_0$ *with* $E_{C_0}(b_0) = 0$ *and* $c_1 \in C_1$ *with* $E_{C_0}(c_1) = 0$.

Moreover (i) to (vii) are equivalent with the analogous conditions obtained by interchanging B_0 *with* C_1.

Proof. Let p,q,r be three projections acting on some Hilbert space. The following are clearly equivalent:

(a) $pq = r$

(b) $pq = rq$ and $r \le q$

(c) $pq = qp = r$.

As we may view the conditional expectations as projections on $L^2(B_1,\mathrm{tr})$, this shows the equivalence of (ii), (iii) and (iv). Obviously (ii) implies (i).

Assume (i) holds and let $b_0 \in B_0$. For all $c_0 \in C_0$ one has

$$\mathrm{tr}(E_{C_0}(b_0)c_0) = \mathrm{tr}(b_0c_0) = \mathrm{tr}(E_{C_1}(b_0)c_0).$$

As $E_{C_1}(b_0) \in C_0$, this implies $E_{C_0}(b_0) = E_{C_1}(b_0)$, and (v) follows. As (v) implies (iii), conditions (i) to (v) are equivalent.

The equivalence of (vi) and (vii) follows from the formula

$$E_{C_0}\{(b_0-E_{C_0}(b_0))(c_1-E_{C_0}(c_1))\} = E_{C_0}(b_0c_1) - E_{C_0}(b_0)E_{C_0}(c_1)$$

for $b_0 \in B_0$ and $c_1 \in C_1$.

The next step is to show that (ii) and (vii) are equivalent. Observe first that one has

$$E_{C_1}(E_{B_0}(x) - E_{C_0}(x)) = E_{C_1}E_{B_0}(x) - E_{C_0}(x)$$

for any $x \in B_1$. Thus (ii) can be reformulated as

$$E_{B_0}(x) - E_{C_0}(x) \perp C_1 \quad \text{for all } x \in B_1.$$

Suppose (ii) holds. Then, in particular, $b_0 \perp C_1$ for $b_0 \in B_0$ with $E_{C_0}(b_0) = 0$. Consequently, for all $c_1 \in C_1$ and $c_0 \in C_0$, one has

$$\mathrm{tr}(E_{C_0}(b_0c_1)c_0) = \mathrm{tr}(b_0c_1c_0) = 0.$$

As tr is faithful on C_0, this implies $E_{C_0}(b_0c_1) = 0$ and (vii) holds.

Suppose (vii) holds. For all $x \in B_1$ and for all $c_1 \in C_1$, one has

$$\begin{aligned} &\mathrm{tr}\{(E_{B_0}(x) - E_{C_0}(x))c_1\} \\ &= \mathrm{tr}\{(E_{B_0}(x) - E_{C_0}(x))(c_1 - E_{C_0}(c_1))\} + \mathrm{tr}\{E_{B_0}(x)E_{C_0}(c_1)\} - \mathrm{tr}\{E_{C_0}(x)E_{C_0}(c_1)\} \\ &= 0 + \mathrm{tr}\{E_{B_0}(xE_{C_0}(c_1))\} - \mathrm{tr}\{E_{C_0}(xE_{C_0}(c_1))\}, \end{aligned}$$

which is zero, since the conditional expectations are trace preserving. Consequently (ii) holds.

Finally, as (iv) is symmetric with respect to B_0 and C_1, we may exchange B_0 and C_1 in any of the conditions (i) to (vii). #

It follows for example from (v) that in diagrams like

$$\begin{matrix} C_2 & \subset & B_2 \\ \cup & & \cup \\ C_1 & \subset & B_1 \\ \cup & & \cup \\ C_0 & \subset & B_0 \end{matrix} \qquad \text{or} \qquad \begin{matrix} C_1 & \subset & B_1 & \subset & A_1 \\ \cup & & \cup & & \cup \\ C_0 & \subset & B_0 & \subset & A_0 \end{matrix}$$

the "rectangles" are commuting squares as soon as the "small squares" are commuting squares.

A crucial point about commuting squares is their behavior with respect to fundamental construction defined in Section 3.6.

Proposition 4.2.2. *Consider a pair* $N \subset M$ *of finite von Neumann algebras, a finite faithful normal trace* tr *on* M, *and the algebra* $\langle M, e_N \rangle$ *obtained by the fundamental construction. Assume that* M *[respectively* N*] is generated as a von Neumann algebra by a nested sequence* $(B_j)_{j \geq 0}$ *[resp.* $(C_j)_{j \geq 0}$*] of von Neumann subalgebras in such a way that one has for each* $j \geq 0$ *a commuting square*

$$\begin{matrix} C_{j+1} & \subset & B_{j+1} \\ \cup & & \cup \\ C_j & \subset & B_j \end{matrix}$$

and set $A_j = \{B_j, e_N\}''$. *Then*

(i) $e_N b e_N = E_{C_j}(b) e_N = e_N E_{C_j}(b)$ *for* $b \in B_j$, $j \geq 0$

(ii) *The algebras* $(A_j)_{j \geq 0}$ *generate* $\langle M, e_N \rangle$ *as a von Neumann algebra.*

Suppose moreover that tr *is a Markov trace of modulus* β *for the pair* $N \subset M$, *and denote the Markov extension of* tr *to* $\langle M, e_N \rangle$ *by* tr *again. (See Definition 3.7.1.) Then*

(iii) $\begin{matrix} B_{j+1} & \subset & A_{j+1} \\ \cup & & \cup \\ B_j & \subset & A_j \end{matrix}$ *is a commuting square with respect to* $\mathrm{tr}|_{A_{j+1}}$.

Proof. (i) For each $j \geq 0$ and $k \geq 1$, the diagram $\begin{matrix} C_{j+k} & \subset & B_{j+k} \\ \cup & & \cup \\ C_j & \subset & B_j \end{matrix}$ is a commuting square, by induction on k. It follows that the limit diagram $\begin{matrix} N & \subset & M \\ \cup & & \cup \\ C_j & \subset & B_j \end{matrix}$ is also a commuting square, and thus for any $b \in B_j$ one has $e_N b e_N = E_N(b) e_N = E_{C_j}(b) e_N$. Since elements of N, and in particular $E_{C_j}(b)$, commute with e_N, this shows (i).

Claim (ii) is obvious.

(iii) One has $E_{B_{j+1}}(e_N) = \beta^{-1}$, because

$$\mathrm{tr}(E_{B_{j+1}}(e_N) x) = \mathrm{tr}(e_N x) = \beta^{-1} \mathrm{tr}(x)$$

for all $x \in B_{j+1}$. Consider now $y_0, y'_\alpha, y''_\alpha \in B_j$. Then

$$E_{B_{j+1}}(y_0 + \sum_\alpha y'_\alpha e_N y''_\alpha) = y_0 + \sum_\alpha y'_\alpha \beta^{-1} y''_\alpha \in B_j.$$

Thus $E_{B_{j+1}}(\tilde{A}_j) \subset B_j$ for a dense $*$-subalgebra $\tilde{A}_j$ of A_j.

Let $x \in A_j$. By the density theorem of Kaplansky, there exists a sequence $(x_k)_{k \geq 1}$, with $x_k \in \tilde{A}_j$ and $\|x_k\| \leq \|x\|$ for all $k \geq 1$, such that $x = \lim_{k \to \infty} x_k$ in the topology defined by the norm $\|\cdot\|_2$. It follows that $E_{B_{j+1}}(x) = \lim_{k \to \infty} E_{B_{j+1}}(x_k) \in B_j$. Thus $E_{B_{j+1}}(A_j) \subset B_j$ and this proves (iii). #

Corollary 4.2.3. *Consider a commuting square* $\begin{matrix} C_1 & \subset & B_1 \\ \cup & & \cup \\ C_0 & \subset & B_0 \end{matrix}$ *with respect to a trace* tr *which is a Markov trace for the pair* $B_0 \subset B_1$. *Let* $B_2 = \langle B_1, e_1 \rangle$ *be the von Neumann algebra obtained via the fundamental construction for* $B_0 \subset B_1$, *and let* $C_2 = \{C_1, e_1\}''$.

Then $\begin{matrix} C_2 & \subset & B_2 \\ \cup & & \cup \\ C_1 & \subset & B_1 \end{matrix}$ *is also a commuting square.*

Proof. This is the special case of 4.2.2 applied to $\begin{matrix} N=B_0 & \subset & M=B_1 & \subset & B_2 \\ \cup & & \cup & & \\ C_0 & \subset & C_1 & \subset & C_2 \end{matrix}$. #

Remark. Suppose moreover that B_0 and B_1 have finite dimensional centers. Then the fundamental construction iterates to give the tower $(B_j)_{j\geq 0}$ with $B_{j+1} = \langle B_j, e_j\rangle$ for all j. Define inductively $C_{j+1} = \{C_j, e_j\}''$ for $j \geq 1$. Then we obtain a ladder of commuting squares $\begin{matrix} C_{j+1} & \subset & B_{j+1} \\ \cup & & \cup \\ C_j & \subset & B_j \end{matrix}$. We are going to use this idea to construct examples of subfactors below, starting with a commuting square of finite dimensional algebras. The next two lemmas concern conditions which cause the inclusion matrices for the resulting ladder of finite dimensional algebras to be repeated with period 2.

Lemma 4.2.4. *Consider a commuting square* $\begin{matrix} C_1 & \subset & B_1 \\ \cup & & \cup \\ C_0 & \subset & B_0 \end{matrix}$ *of finite dimensional von Neumann algebras, with respect to some trace* tr *on* B_1. *Let* $B_2 = \langle B_1, e_1\rangle$ *be the finite dimensional von Neumann algebra obtained via the fundamental construction for* $B_0 \subset B_1$ *and let* $C_2 = \mathrm{alg}\{C_1, e_1\}$.

Suppose that $C_2 = C_1e_1C_1$ *(or equivalently, by 2.6.9, that* $\sum x_i e_{C_0} y_i \stackrel{\sigma}{\longmapsto} \sum x_i e_1 y_i$ *is an isomorphism from the algebra* $\langle C_1, e_{C_0}\rangle$ *obtained by the fundamental construction for* $C_0 \subset C_1$ *onto* C_2*). Then*

(i) $\Lambda_{C_2}^{B_2} = \Lambda_{C_0}^{B_0}$. *More exactly, let* q, p *be minimal central projections in* C_0, B_0 *respectively. Let* $\tilde{q} = \sigma(J_{C_1} q J_{C_1})$ *and* $\tilde{p} = J_{B_1} p J_{B_1}$ *be the corresponding minimal central projections in* C_2, B_2 *respectively. Then* $[(B_0)_{pq} : (C_0)_{pq}] = [(B_2)_{\tilde{p}\tilde{q}} : (C_2)_{\tilde{p}\tilde{q}}]$.

Suppose in addition that tr *is a Markov trace with respect to* $B_0 \subset B_1$. *Let* $(B_j)_{j\geq 0}$ *be the tower obtained by iterating the fundamental construction for* $B_0 \subset B_1$, *with* $B_{j+1} = \langle B_j, e_j\rangle$, *and let* $C_{j+1} = \mathrm{alg}\{C_j, e_j\}$ *for all* $j \geq 1$. *Then*

(ii) *For all* $j \geq 1$, $C_j e_j C_j = C_{j+1}$ *and* $\Lambda_{C_{j+1}}^{B_{j+1}} = \Lambda_{C_{j-1}}^{B_{j-1}}$. *The inclusion matrices for* $C_{j-1} \subset C_j$ *are alternately* $\Lambda_{C_0}^{C_1}$ *and* $\Lambda_{C_1}^{C_2} = (\Lambda_{C_0}^{C_1})^t$.

Proof. (i) Let f be a minimal projection in $(C_0)q$ and let $pf = \sum_{i=1}^{n} g_i$ be a decomposition of pf into orthogonal minimal projections in $(B_0)p$ (so $n = [(B_0)_{pq} : (C_0)_{pq}]^{1/2}$). Then (by 2.6.4) $fe_1 = \sigma(fe_{C_0})$ is a minimal projection in $C_2\tilde{q}$ and

$$\begin{aligned} fe_1\tilde{p} &= fe_1p && \text{(by 3.6.9)} \\ &= fpe_1 && \text{(because } p \in B_0) \\ &= pfe_1 && \text{(because } p \in Z(B_0)) \\ &= \sum_{i=1}^{n} g_i e_1. \end{aligned}$$

Thus $(fe_1)\tilde{p}$ is a sum of n orthogonal minimal projections in $(B_2)_{\tilde{p}\tilde{q}}$.

(ii) We are now supposing that tr is a Markov trace. The statement $C_{j+1} = C_je_jC_j$ is valid for $j = 1$ by hypothesis. Suppose it is valid for some j. Then $C_{j+1}e_{j+1}C_{j+1}$ is an ideal in C_{j+2} containing $\beta e_j e_{j+1} e_j = e_j$, where β is the modulus of the Markov trace. Then $C_{j+1}e_{j+1}C_{j+1} \supset C_je_jC_j \ni \mathbf{1}$, so $C_{j+1}e_{j+1}C_{j+1} = C_{j+2}$.

It follows that for all j, the tower $C_{j-1} \subset C_j \subset C_{j+1}$ is isomorphic to $C_{j-1} \subset C_j \subset \mathrm{End}_{C_{j-1}}(C_j)$, so the inclusion matrices $\Lambda_{C_{j-1}}^{C_j}$ are alternately $\Lambda_{C_0}^{C_1}$ and $(\Lambda_{C_0}^{C_1})^t$. Finally the statement regarding $\Lambda_{C_j}^{B_j}$ follows from (i) and induction. #

Lemma 4.2.5. *Consider a commuting square* $\begin{matrix} C_1 & \subset & B_1 \\ \cup & & \cup \\ C_0 & \subset & B_0 \end{matrix}$ *of finite dimensional von Neumann algebras, with respect to a trace* tr *on* B_1. *Suppose* $\Lambda_{C_0}^{C_1} = \Lambda_{C_0}^{B_0} = \Lambda^t$ *and* $\Lambda_{B_0}^{B_1} = \Lambda_{C_1}^{B_1} = \Lambda$ *for some* Λ. *Let* $B_2 = \langle B_1, e_1\rangle$ *be the algebra obtained via the fundamental construction for* $B_0 \subset B_1$ *and let* $C_2 = \mathrm{alg}\{C_1, e_1\}$. *Then*

(i) $C_2 = C_1e_1C_1$, $\Lambda_{C_1}^{C_2} = \Lambda$, *and* $\Lambda_{C_2}^{B_2} = \Lambda^t$.

Suppose in addition that tr *is a Markov trace with respect to* $B_0 \subset B_1$. *Let* $(B_j)_{j\geq 0}$ *be the tower obtained by iterating the fundamental construction, with* $B_{j+1} = \langle B_j, e_j \rangle$, *and* set $C_{j+1} = \mathrm{alg}\{C_j, e_j\}$ *for all* $j \geq 1$. *Then*

(ii) *The chain* $C_{j-1} \subset C_j \subset C_{j+1}$ *is isomorphic to* $C_{j-1} \subset C_j \subset \mathrm{End}_{C_{j-1}}(C_j)$ *for all* j. *The inclusion matrices* $\Lambda_{C_j}^{C_{j+1}}$ *are alternately* Λ^t *and* Λ $(j \geq 0)$, *and the inclusion matrices* $\Lambda_{C_j}^{B_j}$ *are alternately* Λ^t *and* Λ $(j \geq 0)$.

Proof. (i) We have $C_2 = C_1 e_1 C_1 \oplus K$ and $\Lambda_{C_1}^{C_2} = \begin{bmatrix} \Lambda \\ \Omega_1 \end{bmatrix}$ for some matrix Ω_1, by 2.6.9. Also $\Lambda_{C_1 e_1 C_1}^{B_2} = \Lambda^t$ by the argument of 4.2.4(i), so $\Lambda_{C_2}^{B_2} = \begin{bmatrix} \Lambda^t & \Omega_2 \end{bmatrix}$ for some Ω_2. Therefore $\Lambda_{C_1}^{B_2} = \Lambda_{C_2}^{B_2} \Lambda_{C_1}^{C_2} = \Lambda^t \Lambda + \Omega_2 \Omega_1$. On the other hand $\Lambda_{C_1}^{B_2} = \Lambda_{C_1}^{B_2} \Lambda_{C_1}^{B_1} = \Lambda^t \Lambda$. This is only possible if $K = (0)$, because otherwise $\Omega_2 \Omega_1 \neq 0$. The remainder of (i) and (ii) now follows from the previous lemma. #

The next result is that commuting squares are preserved under reduction by certain projections.

Proposition 4.2.6. *Consider a commuting square* $\begin{matrix} C_1 & \subset & B_1 \\ \cup & & \cup \\ C_0 & \subset & B_0 \end{matrix}$ *with respect to a trace* tr *on* B_1 *and a projection* $p \in B_0 \cap C_1'$, *not zero. Then*

$$\begin{matrix} pC_1 & \subset & pB_1p \\ \cup & & \cup \\ pC_0 & \subset & pB_0p \end{matrix}$$

is a commuting square with respect to $\mathrm{tr}|_{pB_1p}$.

Proof. Let $y \in pB_1p$. Then $E_{pB_0p}(y) = pE_{B_0}(y)p$ because one has $\mathrm{tr}(pE_{B_0}(y)pu) = \mathrm{tr}(pypu) = \mathrm{tr}(yu)$ for all $u \in pB_0p$. Consider $z \in pC_1$, say $z = pc$, with $c \in C_1$. Then

$$E_{pB_0p}(z) = pE_{B_0}(pcp)p = pE_{B_0}(c)p \in pC_0$$

and the claim follows. #

Remark. A similar result holds for reduction by projections in C_0.

Next we give some examples of commuting squares involving relative commutants, fixed-point algebras of groups, and crossed-products.

Proposition 4.2.7. *Let* $N \subset M$ *be a pair of von Neumann algebras, let* tr *be a finite faithful normal trace on* M, *and let* S *be a self-adjoint subset of* N. *Then*

$$\begin{matrix} S' \cap M & \subset & M \\ \cup & & \cup \\ S' \cap N & \subset & N \end{matrix}$$

is a commuting square.

Proof. We may suppose that S is a von Neumann subalgebra of N. Choose $x \in M$. Denote by C the $\|\cdot\|_2$-closure of the convex hull of $\{uxu^* : u$ is unitary and $u \in S\}$ in $L^2(M,\mathrm{tr})$, and denote by y the projection of the origin onto C. Then $y \in M$ because the ball of radius $\|x\|$ in M is a $\|\cdot\|_2$-closed subset of $L^2(M,\mathrm{tr})$. Moreover, by the uniqueness of the projection onto a closed convex set, $uyu^* = y$ for any unitary $u \in S$. It follows that y is also in S'.

For any $z \in S' \cap M$ and for any unitary $u \in S$, one has $\mathrm{tr}(uxu^*z) = \mathrm{tr}(xu^*zu) = \mathrm{tr}(xz)$, so that $E_{S'\cap M}(uxu^*) = E_{S'\cap M}(x)$. Consequently $E_{S'\cap M}(C) = E_{S'\cap M}(x)$, and $y = E_{S'\cap M}(y) = E_{S'\cap M}(x)$. In particular, if $x \in N$, then $C \subset N$ and $E_{S'\cap M}(x) = y \in S' \cap N$. #

Proposition 4.2.8. *Let* M *be a von Neumann algebra given with a finite faithful normal trace* tr. *Let* $G = H \rtimes K$ *be a semi-direct product group which acts on* M *and preserves* tr. *Assume that* K *is a compact group and that the restricted action of* K *on* M *is continuous. Denote by* M^G *the algebra of vectors in* M *fixed by* G, *and similarly for* M^H *and* M^K. *Then*

$$\begin{matrix} M^K & \subset & M \\ \cup & & \cup \\ M^G & \subset & M^H \end{matrix}$$

is a commuting square.

Proof. For each $x \in M$, one has

$$E_{M^K}(x) = \int_K k(x)dk.$$

Suppose moreover that $x \in M^H$. Then $k(x) \in M^H$ for any $k \in K$, so that $E_{M^K}(x) \in M^H \cap M^K = M^G$. #

We leave it to the reader to formulate the details of a proposition involving the diagram

$$\begin{array}{ccc} M \rtimes H & \subset & M \rtimes G \\ \cup & & \cup \\ M & \subset & M \rtimes K \end{array}$$

where $\rtimes$ indicates now a crossed product.

We next describe three examples which are interesting in light of the connections between the theory of subfactors and that of the braid groups.

Example 4.2.9. Let $e_1, \cdots, e_n$ be a sequence of projections acting on some Hilbert space such that

$$\begin{array}{ll} \beta\, e_i e_j e_i = e_i & \text{if } |i-j| = 1 \\ e_i e_j = e_j e_i & \text{if } |i-j| \geq 2 \end{array}$$

for some real number $\beta \geq 1$ (see the last remark of Appendix IIc). Let tr be a normalized faithful trace on the algebra generated by the identity and the e_j's, and assume that the Markov relation

$$\beta\, \mathrm{tr}(we_j) = \mathrm{tr}(w), \quad w \in \mathrm{alg}\{\mathbf{1}, e_1, \cdots, e_{j-1}\}, \quad 1 \leq j \leq n$$

holds (see Section 3.4). Then the diagram

$$\begin{array}{ccc} C_1 = \mathrm{alg}\{\mathbf{1}, e_2, \cdots, e_{n-1}, e_n\} & \subset & B_1 = \mathrm{alg}\{\mathbf{1}, e_1, e_2, \cdots, e_{n-1}, e_n\} \\ \cup & & \cup \\ C_0 = \mathrm{alg}\{\mathbf{1}, e_2, \cdots, e_{n-1}\} & \subset & B_0 = \mathrm{alg}\{\mathbf{1}, e_1, e_2, \cdots, e_{n-1}\} \end{array}$$

is a commuting square.

Proof. Let us show that $E_{B_0}(x) \in C_0$ for any $x \in C_1$. This is obvious when $x \in C_0$. By Proposition 2.8.1, one may then assume without loss of generality that $x = ye_n z$ with $y, z \in C_0$. As $E_{B_0}(e_n) = \beta^{-1}$ (see the proof of 4.2.2.iii), one has $E_{B_0}(ye_n z) = y\beta^{-1}z \in C_0$.

#

Example 4.2.10. Let $N \subset M$ be a connected pair of finite von Neumann algebras with finite dimensional centers, of finite index (Definition 3.5.3). Let tr be the normalized Markov trace on $N \subset M$ (Corollary 3.7.4.i), and let $\beta = [M:N]$ be its modulus (Definition 3.7.5). Then tr has an extension to $\langle M, e_N \rangle$ which is again a Markov trace of modulus β on $M \subset \langle M, e_N \rangle$ (Corollary 3.7.4.ii), and that we denote by tr again.

Suppose moreover that $\beta \leq 4$, write $\beta = 2 + q + q^{-1}$, define

$$g = qe_N - (\mathbf{1} - e_N)$$

and observe that g is a unitary which commutes with N. Then

$$\begin{array}{ccc} gMg^{-1} & \subset & \langle M, e_N \rangle \\ \cup & & \cup \\ N & \subset & M \end{array}$$

is a commuting square.

Proof. Let $x \in gMg^{-1}$. If $y = g^{-1}xg \in M$, one has

$$x = \{(q+1)e_N - \mathbf{1}\}y\{(q^{-1}+\mathbf{1})e_N - \mathbf{1}\}.$$

Since $E_M(e_N) = \beta^{-1}$, we have

$$\begin{aligned} E_M(x) &= \beta\, E_M(e_N y e_N) + \{1 - (q+1)\beta^{-1} - (q^{-1}+1)\beta^{-1}\}y \\ &= \beta\, E_M(E_N(y)e_N) = E_N(y). \qquad \# \end{aligned}$$

Remarks.

(1) Up to scalars, g and g^{-1} are the only unitaries in $\mathrm{alg}\{\mathbf{1}, e_N\}$ for which the above construction works. Observe that g is precisely the element involved in the braid group representation of [Jo2].

(2) This example is the basis for the examples of Section 4.4 below.

Example 4.2.11. Let $N \subset M$ be a pair of factors, of finite index β, and let tr denote the normalized trace on M. Assume that there exists a projection $e_0 \in M$ such that

$$e_0 \text{ and } N \text{ generate } M$$
$$\mathrm{tr}(e_0 y) = \beta\, \mathrm{tr}(y) \text{ for all } y \in N.$$

Let $(M_j)_{j \geq 0}$ be the tower and let $(e_j)_{j \geq 1}$ be as usual. (See Section 3.4; of course $M_1 = M$.) Let M_∞ denote the von Neumann algebra generated by $\bigcup_{j \geq 0} M_j$. Then

$$\begin{array}{ccc} \{\mathbf{1}, e_0, e_1, \cdots\}'' & \subset & M_\infty \\ \cup & & \cup \\ \mathbb{C} & \subset & N \end{array}$$

is a commuting square.

Proof. We want to check that $\mathrm{tr}(xy) = \mathrm{tr}(x)\mathrm{tr}(y)$ for all $x \in \{\mathbf{1}, e_0, e_1, \cdots\}''$ and for all $y \in N$. Because of the density theorem of Kaplansky (see the proof of Proposition 4.2.2.iii), we may check this for all $x \in \mathrm{alg}\{\mathbf{1}, e_0, \cdots, e_n\}$ and for all $n \geq 0$. If $n = 0$, this follows from the hypothesis on e_0. To end the proof, we may assume that $n \geq 1$ and that the claim holds up to $n-1$.

For $a_\alpha, b_\alpha \in \mathrm{alg}\{\mathbf{1}, e_0, \cdots, e_{n-1}\}$ and $x = \sum_\alpha a_\alpha e_n b_\alpha$,

$$\mathrm{tr}(xy) = \mathrm{tr}(\sum_\alpha a_\alpha e_n b_\alpha y) = \beta^{-1}\mathrm{tr}(\sum_\alpha a_\alpha b_\alpha y)$$

which is by induction

$$\beta^{-1}\mathrm{tr}(\sum_\alpha a_\alpha b_\alpha)\mathrm{tr}(y) = \mathrm{tr}(x)\mathrm{tr}(y).$$

This shows that the claim holds up to n. #

Remark. It would be interesting to have a systematic classification of commuting squares

$$\begin{array}{ccc} C_1 & \subset & (B_1, \mathrm{tr}) \\ \cup & & \cup \\ C_0 & \subset & B_0 \end{array}$$

of finite dimensional von Neumann algebras.

4.3 Wenzl's index formula.

In this section we prove a formula due to H. Wenzl [Wen 2] for the index of a pair of factors generated by a ladder of commuting squares. The set up is as follows: We are given a chain $(B_j)_{j \geq 0}$ of finite dimensional von Neumann algebras and a faithful tracial state tr on $B_\infty = \bigcup_j B_j$. Since the GNS representation π of tr (on $\mathcal{H}_\pi$) is faithful, we regard B_∞ as a subalgebra of $B = \pi(\bigcup_j B_j)''$, a finite hyperfinite von Neumann algebra. We suppose we have a chain $(C_j)_{j \geq 0}$ of finite dimensional von Neumann algebras such that $\mathbf{1} \in C_j \subset B_j$ and:

Hypothesis (A). For each j, $\begin{array}{ccc} C_{j+1} & \subset & B_{j+1} \\ \cup & & \cup \\ C_j & \subset & B_j \end{array}$ is a commuting square.

Then $C = (\bigcup_j C_j)''$ is a von Neumann subalgebra of B. In the *periodic* case which we consider below, tr is the unique tracial state on $\bigcup_j C_j$ and $\bigcup_j B_j$, so that C and B are factors.

If $E : B \to C$ and $E_j : B_j \to C_j$ denote the conditional expectations with respect to tr, then $E|_{B_j} = E_j$; that is $\begin{array}{ccc} C & \subset & B \\ \cup & & \cup \\ C_j & \subset & B_j \end{array}$ is a commuting square for each j. Let $A = \langle B, e \rangle$ be the result of the fundamental construction for $C \subset B$ with respect to tr, and let $A_j = \{B_j, e\}''$ for each j. Then A_j is an E_j-extension of B_j, in the terminology of Section 2.6. Hence if $\langle B_j, f_j \rangle$ is the result of the fundamental construction for $C_j \subset B_j$, then the formula $\sigma_j(\sum_i a_i f_j b_i) = \sum_i a_i e b_i$ $(a_i, b_i \in B_j)$ defines an isomorphism from $\langle B_j, f_j \rangle$ onto the two sided ideal $B_j e B_j$ generated by e in A_j, by 2.6.9.

Lemma 4.3.1.

(i) *The central support* z_j *of* e *in* A_j *is* $\sigma_j(\mathbf{1})$; *this is also the central support of the ideal* $B_j e B_j$.

(ii) $\lim_{j \to \infty} z_j = \mathbf{1}$ *in the strong operator topology.*

Proof. (i) This is straightforward, since the central support of f_j in $\langle B_j,f_j\rangle$ is $\mathbf{1}$, by 3.6.1(vi).

(ii) Since $(\bigcup_j A_j)'' = A$ and the central support of e in A is $\mathbf{1}$, one has $\bigvee_j [A_j e \mathcal{X}_\pi] = [Ae\mathcal{X}_\pi] = \mathcal{X}_\pi$. That is, z_j increases to $\mathbf{1}$. #

Next we introduce a very strong periodicity assumption on the inclusion data for the ladder of inclusions

$$\begin{array}{ccc} C_{j+1} & \subset & B_{j+1} \\ \cup & & \cup \\ C_j & \subset & B_j \end{array}.$$

Hypothesis (B). We assume there is a $j_0 \geq 0$ and a $p \geq 1$ and a suitable ordering of the factors in the B_j's and C_j's such that for all $j \geq j_0$:

(i) The inclusion matrix for $B_j \subset B_{j+1}$ is the same as that for $B_{j+p} \subset B_{j+p+1}$. Similarly for $C_j \subset C_{j+1}$ and $C_{j+p} \subset C_{j+p+1}$.

(ii) The inclusion matrices Φ_j for $B_j \subset B_{j+p}$ and Ψ_j for $C_j \subset C_{j+p}$ are primitive.

(iii) The inclusion matrix Λ_j for $C_j \subset B_j$ is the same as that for $C_{j+p} \subset B_{j+p}$.

We remark that under this hypothesis, tr is the unique tracial state on $\bigcup_j C_j$ and $\bigcup_j B_j$. In fact the trace vector $\vec{s}^{(j)}$ for B_j [resp. $\vec{t}^{(j)}$ for C_j] is a Perron–Frobenius (row) vector for Φ_j [resp. Ψ_j] for $j \geq j_0$, because $\vec{s}^{(j)} = \vec{s}^{(j+\ell p)}\Phi_j^\ell$ for all $\ell \geq 0$. Furthermore the dimension vectors $\vec{\mu}^{(j)}$ of B_j [resp. $\vec{\nu}^{(j)}$ of C_j] approach Perron–Frobenius eigenvectors of Φ_j [resp. Ψ_j]. More precisely, if φ_j is the spectral radius of Φ_j, then $\lim_{\ell\to\infty} \vec{\mu}^{(j+\ell p)}/\varphi_j^\ell = \lim_{\ell\to\infty}(\Phi_j^\ell \vec{\mu}^{(j)})/\varphi_j^\ell$ exists, and is a Perron–Frobenius eigenvector for Φ_j, and similarly for the vectors $\vec{\nu}^{(j)}$.

Lemma 4.3.2. *Assuming hypotheses* (A) *and* (B),

(i) B *and* C *are factors and* $[B:C] < \infty$.

(ii) $\vec{t}^{(j)}\Lambda_j^t\Lambda_j \leq [B:C]\,\vec{t}^{(j)}$ *for all* j, *the inequality holding component-by-component.*

(iii) *If* z_k *is the central support of* e *in* A_k *and* ψ_j *denotes the spectral radius of* Ψ_j, *then for* $j \geq j_0$,

$$\mathrm{tr}(\mathbf{1}-z_{j+\ell p}) = \langle \vec{t}^{(j)} - [B:C]^{-1}\vec{t}^{(j)}\Lambda_j^t\Lambda_j,\ \psi_j^{-\ell}\vec{\nu}^{(j+\ell p)}\rangle.$$

Proof. (i) That B and C are factors follows from the uniqueness of the trace on $\bigcup_j B_j$ and $\bigcup_j C_j$. We have to show that $A = \langle B,e\rangle$ is a finite factor. In any case A is semi-finite, so has a faithful normal semi-finite trace Tr; we have to show that $\mathrm{Tr}(\mathbf{1}) < \infty$. Now $eAe = Ce$ is isomorphic to C, which is a finite factor, so e is a finite projection and $\mathrm{Tr}(e) < \infty$. Fix some $j \geq j_0$ and let q_i be a minimal central projection in C_j, and $\tilde{q}_i = J_{B_j} q_i J_{B_j}$ and $\bar{q}_i = \sigma_j(\tilde{q}_i)$ the corresponding minimal central projections in $\langle B_j,f_j\rangle$ and $B_j e B_j = z_j A_j$. Then

$$\frac{\mathrm{Tr}(e\bar{q}_i)}{\mathrm{Tr}(\bar{q}_i)} = \frac{\mathrm{Tr}(eq_i)}{\mathrm{Tr}(\bar{q}_i)} \qquad \text{(using 3.6.9)}$$
$$= \nu_i^{(j)}/(\Lambda_j^t\Lambda_j\vec{\nu}^{(j)})_i,$$

because eq_i is the sum of $\nu_i^{(j)}$ minimal projections in $\bar{q}_i A_j$ (by 2.6.4) while $\bar{q}_i$ is the sum of $(\Lambda_j^t\Lambda_j\vec{\nu}^{(j)})_i$ minimal projections in $\bar{q}_i A_j$. Let $d_j = \min_i\{\nu_i^{(j)}/((\Lambda_j^t\Lambda_j\vec{\nu}^{(j)})_i\}$. Then we have $d_j > 0$ and

$$\mathrm{Tr}(e) = \mathrm{Tr}(ez_j) = \sum_i \mathrm{Tr}(e\bar{q}_i) \geq d_j \sum_i \mathrm{Tr}(\bar{q}_i) = d_j\mathrm{Tr}(z_j).$$

Since $\vec{\nu}^{(j+\ell p)}/\psi_j^\ell$ converges to a Perron–Frobenius vector for Ψ_j, it follows that $d := \inf_{\ell\geq 0} d_{j+\ell p}$ is positive. We have

$$\mathrm{Tr}(e) \geq d_{j+\ell p}\mathrm{Tr}(z_{j+\ell p}) \geq d\,\mathrm{Tr}(z_{j+\ell p})$$

for all ℓ, and since $\lim_\ell z_{j+\ell p} = \mathbf{1}$, it follows that $\mathrm{Tr}(e) \geq d\,\mathrm{Tr}(\mathbf{1})$, and Tr is finite.

Since $\beta = [B:C] < \infty$, the normalized trace tr on A has the Markov property: $\mathrm{tr}(ex) = \beta^{-1}\mathrm{tr}(x)$ for $x \in B$. It follows from this and 2.6.4(c) that the weight vector of tr on $B_j e B_j = z_j A_j$ is $\beta^{-1}\vec{t}^{(j)}$.

(ii) It follows from 2.4.1(b) and 2.6.9 that the inclusion matrix of $B_j \subset A_j$ is of the form $\begin{bmatrix}\Lambda_j^t \\ \Omega_j\end{bmatrix}$ for some Ω_j, Λ_j^t being the inclusion matrix of $B_j \subset z_jA_j$. By the remark above, the weight vector of tr on A_j has the form $(\beta^{-1}\vec{t}^{(j)}, \vec{r}^{(j)})$, so that