NOTES ON SUBFACTORS AND STATISTICAL MECHANICS

V.F.R. JONES
Dept. of Mathematics,
University of California at Berkeley,
Berkeley, CA 94720, USA

0. Introduction

A lot has been made in the last few years of connections between knot theory, statistical mechanics, field theory and von Neumann algebras. Because of their more technical nature, the von Neumann algebras have tended to be neglected in surveys. This is not an accurate reflection of their fundamental role in the subject, both as a continuing inspiration and as the vehicle of the discovery of the original ties between statistical mechanics and knot theory. In this largely expository article, we attempt to redress this balance by talking almost entirely about von Neumann algebras and how they occur as algebras of transfer matrices in statistical mechanical models. We shall focus mostly on the Potts model and the model of Fateev and Zamolodchikov, with a brief exposition of how vertex models are related to type III factors.

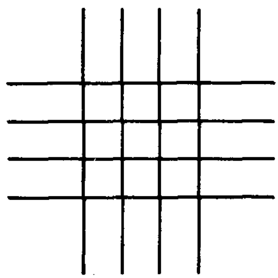
1. Spin Models in Statistical Mechanics

A spin model may be defined on any graph G. One is given a set S, with Q elements, of spin states per site, and a state of the system is a function from the set of vertices V of G to S. For each edge e in G we suppose given an energy function $E_e : S \times S \longrightarrow \mathbb{R}$ such that $E_e(\sigma, \tau)$ is the energy of the "bond" represented by e if the ends of e are in states σ and τ. (If the edges of G are directed, E_e does not have to be symmetric — see [A–YBP] — but we will not consider this more general case). We then define $w_e : S \times S \longrightarrow \mathbb{R}^+$ by $w_e(\sigma, \tau) = \exp(-\beta E_e(\sigma, \tau))$, β being a constant. Given a state $\sigma : V \longrightarrow S$, its energy is then $\Sigma_e E_e(\sigma(\alpha), \sigma(\beta))$, α and β being the ends of e, and the corresponding Boltzmann weight is then $w(\sigma) = \Pi_e w_e(\sigma(\alpha),\ \sigma(\beta))$.

The *partition function* of the system is then defined to be

$$Z = \sum_{\sigma:V \to S} \prod_{e} w_e(\sigma(\alpha), \sigma(\beta)) \ .$$

If the graph G has a special form, one may make restrictions on w_e. The most commonly encountered situation in 2 dimensions is a graph which is part of a square lattice

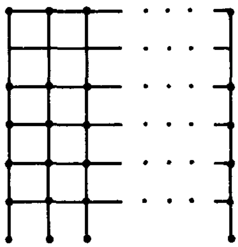


where the boundary conditions may require that identifications are made on the boundary (e.g., periodic boundary conditions when the graph is wrapped on a torus). The edges of the graph can now be divided horizontally and vertically, and the Boltzmann weights are only allowed to be of two kinds, one for each direction: w_{H} and w_{V}.

To calculate the partition function in the above situation, it is very convenient to use a *transfer matrix*, which builds the lattice up row by row. Thus we consider a vector space V with basis S and define the linear map $T : \otimes^n V \longrightarrow \otimes^n V$ by

$$T(\sigma_1 \otimes \sigma_2 \ldots \otimes \sigma_n) = \sum_{\tau_1, \tau_2, \ldots, \tau_n} \left(\prod_{i=1}^{n-1} w_{\mathrm{H}}(\sigma_i, \sigma_{i+1}) \right) \times \left(\prod_{i=1}^{n} w_{\mathrm{V}}(\sigma_i, \tau_i) \right) \tau_1 \otimes \tau_2 \ldots \otimes \tau_n \ .$$

It is easy to see then that the $(\sigma_1, \ldots, \sigma_n; \tau_1, \ldots \tau_n)$ entry of T^m is then the partition function for the system:



with m rows and n columns, where the spins at the top and bottom are fixed to be $(\tau_1, \tau_2, \ldots, \tau_n)$ and $(\sigma_1, \sigma_2, \ldots, \sigma_n)$, respectively. We see that the row-to-row transfer matrix T further splits up as a product

$$T = R_1 R_3 \ldots R_{2n-1} R_2 R_4 R_6 \ldots R_{2n-2} \ ,$$

where

$$R_{2i-1}(\sigma_1 \otimes \ldots \otimes \sigma_n) = \sum_{\tau} w_{\mathrm{V}}(\sigma_i, \tau)\sigma_1 \otimes \ldots \otimes \tau \otimes \ldots \otimes \sigma_n$$

$$R_{2i}(\sigma_1 \otimes \ldots \otimes \sigma_n) = w_{\mathrm{H}}(\sigma_i, \sigma_{i+1})\sigma_1 \otimes \ldots \otimes \sigma_n \ .$$

The individual R_i's can generate some very interesting algebras and this was the reason for the connection between von Neumann algebras, statistical mechanics and knot theory referred to in the introduction. The reason for the interesting algebra from the statistical mechanics point of view is the following: Suppose all Boltzmann weights depend on a parameter λ. Suppose further that for each pair λ, λ' in parameter space, there is a third λ'' such that $R_i(\lambda)R_{i+1}(\lambda')R_i(\lambda'') = R_{i+1}(\lambda'')R_i(\lambda')R_{i+1}(\lambda)$ (the so-called star triangle relation), then an elegant argument (on p. 93 of [B]) shows that the diagonal-to-diagonal transfer matrices for the system wrapped on a torus commute for all parameters in the family. Note that the star triangle relation only needs to be checked for $i = 1$ and 2.

We give two examples of the algebraic relations we have in mind.

Example 1.1. The Potts model

This is the model defined by $w_{\mathrm{H}}(\sigma, \tau) = \exp(K_1 \delta_{\sigma,\tau})$ and $w_{\mathrm{V}}(\sigma, \tau) = \exp(K_2 \delta_{\sigma,\tau})$. Then it is clear that if we define e_{2i-1} by $e_{2i-1}(\sigma_1 \otimes \ldots \otimes \sigma_n) = \frac{1}{Q}\Sigma_\tau \sigma_1 \otimes \ldots \otimes \tau \otimes \ldots \sigma_n$, then $R_{2i-1} = Qe_{2i-1} + (e^{K_2} - 1)\mathbb{1}$, $\mathbb{1}$ being the identity matrix, and if $e_{2i}(\sigma_1 \otimes \ldots \otimes \sigma_n) = \delta_{\sigma_i, \sigma_{i+1}}(\sigma_1 \otimes \ldots \otimes \sigma_n)$ then

$R_{2i} = (e^{K_1} - 1)e_{2i} + 1$. It is easy to verify the following relations

$$e_i^2 = e_i$$

(TL)
$$e_i e_{i\pm 1} e_i = \frac{1}{Q} e_i$$
$$e_i e_j = e_j e_i \text{ if } |i-j| \geq 2 \ .$$

The usefulness of these relations was first noticed by Temperley and Lieb ([TL]), who used the easily proven fact that if we define $p = e_1 e_3 \ldots e_{2n-1}$, then $p^2 = p$ and $pxp = \varphi(x)p$ for any x in the algebra generated by $e_1, e_2, \ldots e_{2n-1}$. The proof shows that the linear functional φ is completely determined by the algebraic relations (TL). The linear functional φ is also essentially the sum of all the matrix entries so that the partition function for the Potts model is also determined by (TL). We will see that these relations occur again in our discussion of von Neumann algebras. This remarkable coincidence was first pointed out by D. Evans in 1984. The above star-triangle equation is true for the Potts model but an analysis of the argument shows that the transfer matrices will commute only provided K_1 and K_2 satisfy a relation. This relation is believed to be the same as criticality for the model.

Example 1.2. The model of Fateev and Zamolodchikov

The Potts model has an obvious symmetry group given by the full symmetric group S_Q, since the Boltzmann weight $w(\sigma, \tau)$ depends only on whether $\sigma = \tau$. It is equally clear that the Potts model is the only model with this property. If one reduces the size of the symmetry group one can expect to find other models. The most natural starting case would be $\mathbb{Z}/Q\mathbb{Z}$ so that the Boltzmann weights would have to satisfy the property that $w(a,b)$ depends only on $a - b \pmod{Q}$. Write $w_{\rm H}(a,b) = x_{\rm H}(a-b)$ and $w_{\rm V}(a,b) = x_{\rm V}(a-b)$. Such a model, with a parameter λ, has been found by Fateev and Zamolodchikov in [FZ]. The functions $x_{\rm H}$ and $x_{\rm V}$ are:

$$x_{\rm H}(n,\lambda) = \prod_{k=0}^{n-1} \frac{\sin\left(\pi\frac{k}{Q} + \frac{\lambda}{2Q}\right)}{\sin\left(\pi(k+1)/Q - \frac{\lambda}{2Q}\right)} \ .$$

The function $x_{\rm V}(n,\lambda)$ is the finite Fourier transform of the function $x_{\rm H}(n,\lambda)$.

To take advantage of the $\mathbb{Z}/Q\mathbb{Z}$ symmetry, it is convenient to analyze the transfer matrices in terms of operators,

$$U_1, U_2, \ldots U_n, \text{ where } U_i U_{i+1} = e^{2\pi i/Q} U_{i+1} U_i,$$
$$U_i U_j = U_j U_i \text{ if } |i-j| \geq 2 \text{ and } U_i^Q = 1 \ .$$

These relations are represented on $\otimes^n V$ by

$$U_{2i-1}(\sigma_1 \otimes \ldots \otimes \sigma_n) = \sigma_1 \otimes \ldots \otimes (\sigma_i + 1) \otimes \ldots \otimes \sigma_n \ ,$$
$$U_{2i}(\sigma_1 \otimes \ldots \otimes \sigma_n) = \omega^{\sigma_i - \sigma_{i+1}}(\sigma_1 \otimes \ldots \otimes \sigma_n) \ .$$

The horizontal transfer matrices for the above model are then $\Sigma_{k=0}^{Q-1} x_{\mathrm{H}}(k, \lambda)\ U_{2i-1}^k$ and the vertical ones are (at least up to a constant) $\Sigma_{k=0}^{Q-1} x_{\mathrm{H}}(k, \lambda) U_{2i}^k$.

2. von Neumann Algebras

If $\mathcal{H}$ is a complex Hilbert space with inner product $\langle\ ,\rangle$, a *von Neumann algebra* M is an algebra of bounded linear operators on $\mathcal{H}$ such that

a) $id \in M$

b) If $a \in M$, then $a^* \in M$.

c) If a_n is a set of elements of M with $\langle a_n \xi, \eta \rangle \longrightarrow \langle a\xi, \eta \rangle$ for some a, then $a \in M$.

The simplest example of such an M is the algebra of all bounded operators $\mathcal{B}(\mathcal{H})$. This example has "atoms", i.e., projections p on to one-dimensional subspaces, so that $pMp = \mathbb{C}p$. In finite dimensions any von Neumann algebra is abstractly (i.e. as a $*$-algebra) isomorphic to a direct sum of finitely many copies of $\mathcal{B}(\mathcal{H})$. Thus it is determined as an algebra by integers $n_1, n_2, \ldots n_k$ which can be conveniently represented as the row of integers $n_1, n_2, \ldots n_k$. For instance

$$2 \qquad 3 \qquad 1$$

stands for the algebra $M_2(\mathbb{C}) \oplus M_3(\mathbb{C}) \oplus \mathbb{C}$. What we have just said completely describes finite-dimensional von Neumann algebras up to abstract isomorphism. To understand how they can act on Hilbert spaces, one need only consider the case of a simple direct summand $M_n(\mathbb{C})$. Elementary arguments show that, since the identity operator on $\mathcal{H}$ must be in $M = M_n(\mathbb{C})$, the dimension of $\mathcal{H}$ is a multiple mn of n and that there is a tensor product splitting $\mathcal{H} \simeq \mathbb{C}^n \otimes \mathcal{H}_2$ with the action of $a \in M$ being as $a \otimes id_{\mathcal{H}_2}$. The integer m is then called the *multiplicity* of the algebra M on $\mathcal{H}$.

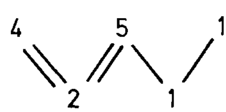
Already, in the finite-dimensional case, we see the importance of the direct summands in the decomposition. They are characterized as algebras with trivial centre and this is taken as a definition in the general case: a *factor* is a von Neumann algebra whose centre is just the scalar multiples of the identity. Our only infinite-dimensional example $\mathcal{B}(\mathcal{H})$ is certainly a factor. Any factor

M with "atoms" (i.e. elements p with $p^* = p^2 = p$ and $pMp = \mathbb{C}p$) is called *type I* and is abstractly isomorphic to $\mathcal{B}(\mathcal{H})$ for some $\mathcal{H}$, not necessarily the $\mathcal{H}$ on which it acts. As in finite dimensions, type I factors acting on Hilbert spaces induce tensor product factorizations.

Not all factors are of type I. This was the great discovery of Murray and von Neumann in [MvN]. There are infinite-dimensional factors that in some ways are more like $M_n(\mathbb{C})$ than $\mathcal{B}(\mathcal{H})$ (dim $\mathcal{H} = \infty$). These are the *type* II_1 factors which admit a trace tr: $M \longrightarrow \mathbb{C}$ which is a (unique) linear function such that $\text{tr}(ab) = \text{tr}(ba)$, $\text{tr}(1) = 1$. We shall construct some below. One can show that if one restricts the trace to the projections of a II_1 factor, the allowed values are precisely the unit interval $[0, 1]$, where one would have obtained $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}$ for $M_n(\mathbb{C})$. This is *continuous dimensionality.*

Beyond II_1 factors, one may construct type II_∞ ones as tensor products $\text{II}_1 \otimes \mathcal{B}(\mathcal{H})$. There are still other factors, called type III, which admit no projections p such that the algebra pMp has a trace as above. The type III factors do admit a complicated splitting as "crossed products" of type II objects by groups. For an excellent survey of the global structure of factors, see [Co].

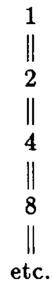
We now want to take up the construction of a type II factor. The method we shall give is quite limited and somewhat complicated but it will introduce ideas that will be crucial later. We will start with finite-dimensional von Neumann algebras and build them up to II_1 factors. For this, it behooves us to study inclusions of finite-dimensional von Neumann algebras. These are described by "Bratteli diagrams" as amply illustrated in the following example:



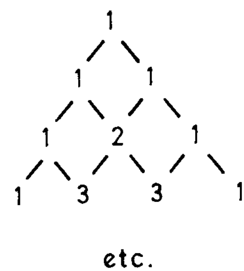
This diagram means two algebras $N \subseteq M$, where $N = M_2(\mathbb{C}) \oplus \mathbb{C}$, $M = M_4(\mathbb{C}) \oplus M_5(\mathbb{C}) \oplus \mathbb{C}$ and an element $A \oplus x$ of N is the element

$$\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \oplus \begin{pmatrix} A & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & x \end{pmatrix} \oplus x \text{ of } M\,.$$

One can stack such diagrams on top of each other to obtain infinite-dimensional abstract algebras as limits of finite-dimensional ones. Thus



is the Bratteli diagram for an algebra A which may alternatively be described as the infinite tensor product $\otimes_{i=1}^{\infty} M_2(\mathbb{C})$. A more interesting case is the following

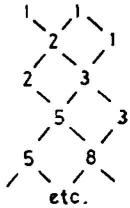


This is Pascal's triangle and this algebra may be obtained from the previous one by looking at the fixed points for the action of the circle group $\mathbb{T}$, an element $e^{i\theta}$ of which acts by

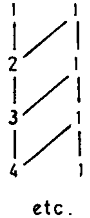
$$x \longrightarrow \left[\bigotimes_{i=1}^{\infty} \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}\right] x \left[\bigotimes_{i=1}^{\infty} \begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix}\right] .$$

The algebras A obtained by this process do not act naturally on a Hilbert space. To make them into von Neumann algebras, one uses the so-called GNS construction ([Ar]). (Note that for this construction, A need not be of the form described above). One begins with a linear functional $\varphi : A \longrightarrow \mathbb{C}$ with the properties $\varphi(a^*a) \geq 0$ and $\varphi(1) = 1$ (a "state"). The set of all a such that $\varphi(a^*a) = 0$ is then a subspace I and A/I inherits a pre-Hilbert structure: $\langle a, b\rangle = \varphi(b^*a)$. The algebra A acts on A/I, by left multiplication, as bounded operators which therefore extend to the Hilbert space completion $\mathcal{H}_\varphi$ of A/I. The relevant von Neumann algebra M_φ is the closure of A in the topology implicit in the definition of a von Neumann algebra. M_φ is to be thought of as the *von Neumann completion of A*.

The simplest situation is when φ is a trace, i.e., $\varphi(ab) = \varphi(ba)$. Then φ extends to a trace on M_φ which thus, if a factor, will be a II_1 factor. Now traces are easy to keep track of on finite-dimensional von Neumann algebras. Moreover, it is easy to show that if the algebra A admits precisely one trace tr, then the von Neumann algebra M_{tr} is a II_1 factor. It is very easy to give criteria on a Bratteli diagram of A for this to happen. Basically, if the width of the diagram is finite and it is not too lopsided, there is a unique trace. For instance, the algebra



admits a unique trace, whereas the algebra



does not. (Note that a given trace on a Bratteli diagram may define a II_1 factor without the diagram satisfying these conditions — there is a one-parameter family of such traces on our Pascal's triangle example.)

Thus we have a whole host of II_1 factors. Just choose any Bratteli diagram satisfying the criteria and it has a unique trace, which gives a II_1 factor as the von Neumann algebra completion of A. It is a remarkable result of Murray and von Neumann that *all such II_1 factors are isomorphic*, and the resulting factor is called the hyperfinite II_1 factor, R. By a deep result of Connes [Co2], R is a "minimal" II_1 factor. Any II_1 subfactor is again isomorphic to R, and R is contained in any II_1 factor. It can also be constructed by many other means. It is naturally the Clifford algebra of an infinite-dimensional Hilbert space.

Before passing to subfactors, we must say a few words about the "multiplicity" of a II_1 factor on a Hilbert space. This is the analogue of the multiplicity defined by the tensor product splitting that comes with the action of a type I factor on a Hilbert space. It is most conveniently thought of by considering the II_1 factor M as an abstract object capable of acting on many Hilbert spaces $\mathcal{H}$ which become "Hilbert spaces over M". In this analogy, M behaves like a field and the Hilbert space as a vector space over that field. The multiplicity then plays the role of the dimension of the vector space and we shall write it $\dim_M(\mathcal{H})$. The striking fact is that for II_1 factors, all positive real values of $\dim_M(\mathcal{H})$ (and ∞) are allowed, where we would have found only integers for a type I factor (and no obvious coherent way of measuring dimension for a type III factor). If M were constructed from A by the GNS construction with respect to a trace tr (with $\mathrm{tr}(a^*a) > 0$ for $a \neq 0$) the space $\mathcal{H}_{\mathrm{tr}}$ would be a "free module of rank 1", thus $\dim_M(\mathcal{H}_{\mathrm{tr}}) = 1$. (This Hilbert space is often denoted $L^2(M, \mathrm{tr})$ as it is also obtained by applying the GNS construction to the algebra M itself, and if M were $L^\infty(X, \mu)$ with trace $\int f d\mu$, one would obtain $L^2(X, \mu)$.)

To see Hilbert spaces of all real dimensions over M, first take $L^2(M, \mathrm{tr})$ and look at the space $\mathcal{H} = L^2(M, \mathrm{tr})p$ where p is a projection in M. (Since tr is a trace, M also acts on the right on the GNS Hilbert space). Not surprisingly, $\dim_M(\mathcal{H}) = \mathrm{tr}(p)$ and by continuous dimensionality, this can be any real between 0 and 1. Values larger than 1 can be obtained simply by taking direct sums of the Hilbert spaces constructed above.

3. Subfactors

Pursuing the analogy which would have it that a II_1 factor is a field, the study of subfactors is an extension of Galois theory and we can expect to see finite groups come into play. Thus the situation we want to consider is two II_1 factors $N \subseteq M$, with the same identity. If N and M were fields, the most immediate entity to attach would be the integer giving the degree of the extension, the dimension of M over N. But following our previous discussion, we can define this number using our "$\dim_N$" ideas by the simple expedient of applying the GNS construction to complete M to the Hilbert space $L^2(M)$. Thus we define

$$[M : N] = \dim_N(L^2(M)), \text{ called the } \textit{index of } N \textit{ in } M.$$

We shall restrict our attention in this article to the case $[M : N] < \infty$.

The simplest example of all for a finite index subfactor is obtained by tensoring by a matrix algebra, thus $M = N \otimes M_n(\mathbb{C})$, N considered as a subfactor of M via the obvious diagonal embedding. A moment's thought shows that the index in this example must be n^2. Despite its triviality, many constructions to follow are easily worked out for this example and it is instructive to do so.

The next most interesting example is to take a finite group G and suppose that it acts by automorphisms on M. If all the automorphisms are outer (i.e., not of the form $x \longrightarrow uxu^{-1}$) except the identity, it is well known that M^G is again a II_1 factor. It is no surprise that $[M : M^G] = |G|$. This example is very important in the theory, especially as it is known ([JI]) that any finite group can act on R in one and only one way by outer automorphisms. It is also important that this example is irreducible in the sense that no element of M commutes with all of M^G except the identity. This makes it quite different from the previous $N \otimes M_n(\mathbb{C})$ example.

Note that we have only seen integers so far as the indices of subfactors, whereas according to the continuous dimension idea, the index is just a real number ≥ 1. The now obvious question of deciding what real numbers occur as index values is completely answered by the following:

Theorem (J2)

a) If $N \subseteq M$ are II_1 factors and $[M : N] < 4$, there is an integer $n \geq 3$ with $[M : N] = 4\cos^2 \pi/n$.

b) For each number $r = 4\cos^2 \pi/n$ as in a), and for each real number $r \geq 4$, there is a subfactor $R_r \subseteq R$ with $[R : R_r] = r$.

The original proof of this theorem involved a basic construction which has since proved to be quite powerful and useful in many situations (see [We1]). The key ingredient is the conditional expectation $E_N : M \longrightarrow N$ (see [U]), which is the restriction to M of the orthogonal projection e_N from $L^2(M, \mathrm{tr})$ to $L^2(N, \mathrm{tr})$. Identifying M with its image under the GNS construction on $L^2(M, \mathrm{tr})$, it is easy to show that e_N satisfies the following properties:

(3.1) $e_N^2 = e_N^* = e_N$

(3.2) If $x \in M$, then $xe_N = e_N x$ iff $x \in N$.

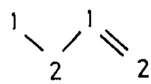
(3.3) For $x \in M$, $e_N x e_N = E_N(x) e_N$ so $e_N M e_N = N e_N$.

The situation presented by these equations is quite interesting and one may enquire as to the algebra generated by M and e_N in the cases a) where M and N are as constructed above b) where M, N and e_N are all in some larger algebra, and E_N in some map.

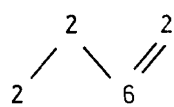
In case a), we write the algebra as $\langle M, e_N \rangle$ and call this the *basic construc-*

tion. In case b), the algebra is not as easy to pin down but can, in general, be determined by an analysis of traces. It invariably contains $\langle M, e_N \rangle$ as a two-sided ideal. For a precise statement, see theorem 1.1 of [We2].

The basic construction is even useful, and quite enjoyable, in finite dimensions. The result is that if $N \subseteq M$ is given by a Bratteli diagram, then $M \subseteq \langle M, e_N \rangle$ is given by the mirror image of that diagram (independent of the trace used, provided $\mathrm{tr}(a^*a) > 0$ for $a \neq 0$). For instance, if $N \subseteq M$ is



then $M \subseteq \langle M, e_N \rangle$ is



We have digressed a little, so let us return to the II_1 factor case. The crucial result is:

Lemma 3.1. If $[M : N] < \infty$, then $\langle M, e_N \rangle$ is a II_1 factor and
(i) $[\langle M, e_N \rangle : M] = [M : N]$
(ii) $\mathrm{tr}(xe_N) = [M : N]^{-1}\ \mathrm{tr}(x)$ if $x \in M$.
(Note that we do not need to specify which trace we mean since II_1 factors admit only one normalized trace.)

Property (i) of the lemma meakes it tempting to iterate the basic construction. One obtains a tower $M_i \subseteq M_{i+1}$ of II_1 factors defined by $M_1 = N, M_2 = M, M_{i+1} = \langle M_i, e_{M_{i-1}} \rangle$.

Writing $e_{M_i} = e_i$, properties (3.1), (3.2) and (3.3) make it easy to show:

$$\text{(TL)} \qquad \begin{aligned} &e_i^2 = e_i \\ &e_i e_{i\pm 1} e_i = \tau e_i \qquad\qquad (\tau = [M : N]^{-1}) \\ &e_i e_j = e_j e_i \quad |i - j| \geq 2\ . \end{aligned}$$

Note that these are *precisely* the relations used by Temperley and Lieb for the Potts model! Part a) of the theorem is now proved by using positivity of the restriction of $\mathrm{tr}(aa^*)$ to the algebra generated by $e_1, e_2, \ldots$, and (ii) of lemma 3.1.

Part b) of the theorem is proved for $[M : N] < 4$ by iterating the basic construction for a pair of finite-dimensional von Neumann algebras and then using the resulting von Neumann algebra generated by $\{e_1, e_2, \dots\}$ as R and R_r the von Neumann algebra generated by $\{e_2, e_3, \dots\}$.

4. The Tower of Relative Commutants

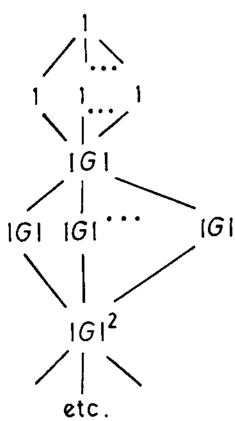
Although only the e_i's were used in the proof of the theorem of Sec. 3, there is a much richer structure available and attention was turned to it immediately after the proof of the theorem. Its existence follows from the next almost trivial result.

Lemma 4.1. If $N \subseteq M$ has finite index then $N' \cap M = \{x \in M | xy = yx$ for all y in $N\}$ is a finite-dimensional von Neumann algebra.

Thus if we consider the tower

$$N \subseteq M \subseteq M_3 \subseteq M_4 \subseteq M_4 \subseteq \dots$$

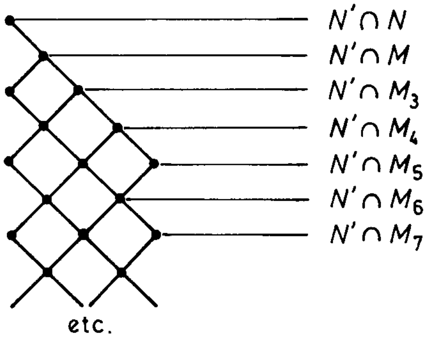
previously constructed, then all the algebras $N' \cap M_i$ and $M' \cap M_i$ are finite-dimensional. They also contain the e_i's which is precious information. We thus have two Bratteli diagrams, for $N' \cap M_i$ and $M' \cap M_i$ which are also connected by "horizontal" inclusions. The simplest nontrivial example to analyze is that in which M carries an outer action of a finite abelian group G and $N = M^G$, the fixed point algebra. Standard duality arguments then give the following Bratteli diagram for both towers $N' \cap M_i$, $M' \cap M_i$



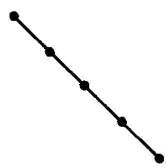
The inclusion diagram between the towers is the obvious one. If we had chosen a non-abelian group, the two towers would be different — the $M' \cap M_i$ tower would be as above but the second level of the $N' \cap M_i$ tower would be given by the group algebra.

The most penetrating analysis of the two towers has been made by Ocneanu ([O1]) who has announced a characterization of the possible combinatorial structures. The presence of the e_i's in the towers means they always have the property that the mirror image of the previous inclusion is always part of the next inclusion. Thus we may throw away all the reflected stuff and look only at what Ocneanu calls the *principal graph.*

For instance, the following is an example of a tower



Throwing away the reflections, we just see



One may recognize the graph A_5 of the Coxeter-Dynkin theory. It was known to the author in 1982 that if $[M : N] < 4$, the principal graph must be an A, D, E Coxeter graph (for a full account, see [GHJ]). For instance the diagram D_4 can be realized by choosing $G = \mathbb{Z}/3\mathbb{Z}$ in the previous example. According to Ocneanu's announcements, the graphs D_{2n+1} and E_7 cannot be obtained.

The principal graph may be either finite or infinite and the subfactor is said to be of finite or infinite depth, accordingly. (The subfactors R_r for $r \geq 4$ are of infinite depth.)

The most impressive result about the towers from a von Neumann algebra point of view is the following, proved by Popa in [Po] and announced by Ocneanu in [O2].

If $R_0 \subseteq R$ is a finite depth subfactor, then the inclusion $R_0 \subseteq R$ is isomorphic to the inclusion coming from the appropriately completed towers of relative commutants.

This result reduces the classification of finite depth subfactors to the combinatorial elucidation of all possible towers. Ocneanu has developed another approach to this problem using bimodules rather than algebras, the equivalence of the two approaches hinging on the fact that $\langle M, e_N \rangle$ is $M \otimes_N M$ as an $M - M$ bimodule. (See [O2].)

From the statistical mechanics point of view, the interest of the towers lies more in the possibility of defining solvable models using their combinatorial structure. This led Pasquier to define his *A-D-E* models ([Pa]) generalizing the Andrews-Baxter-Forrester models [(ABF)]. These are IRF models, not spin models, so we just refer the interested reader to [Pa] and references therein.

This thoroughly motivates the search for examples of subfactors. Although Ocneanu's announced machinery is in principle computable, most of the beautiful examples of subfactors have been found by other means. In [We2], Wenzl constructed many examples using Hecke algebras. The combinatorial ingredients are the Young diagrams, and corresponding statistical mechanical models occur in the work of Date, Miwa and Jimbo (see Jimbo's article in this volume). Wenzl has also found many examples using the Birman-Wenzl algebra ([BW]) which was discovered through the knot theory connection. There are also many sporadic examples (see [GHJ]) and some new ones of small index (< 5) have been announced recently by Ocneanu, and Haagerup. Presumably, they all have solvable models associated with them.

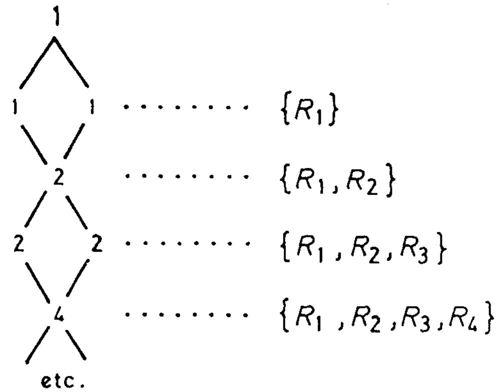
5. Some Examples of Bratteli Diagrams

We have seen many examples of algebras obtained as limits of finite-dimensional von Neumann algebras, both in statistical mechanics and subfactors. In this section, we given the Bratteli diagrams for some of them. In all cases except the Hecke algebra, the structure can be calculated by an elaboration of the basic construction technique in finite dimensions.

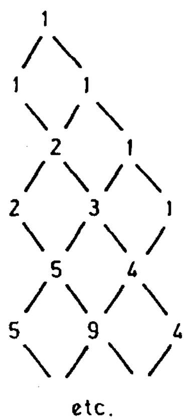
1) The Q-state Potts model

The Temperley-Lieb algebra generated by the elementary transfer matrices R_i has the following Bratteli diagram.

$Q = 1$:

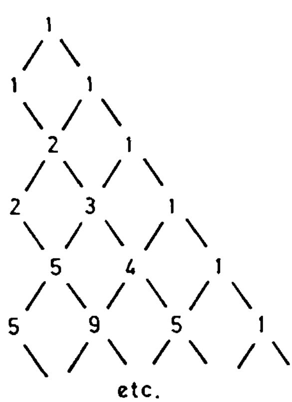


$Q = 3$:



(R_i's as above)

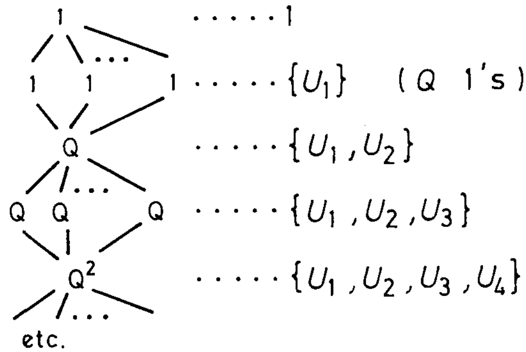
$Q \geqslant 4$:



(R_i's as above)

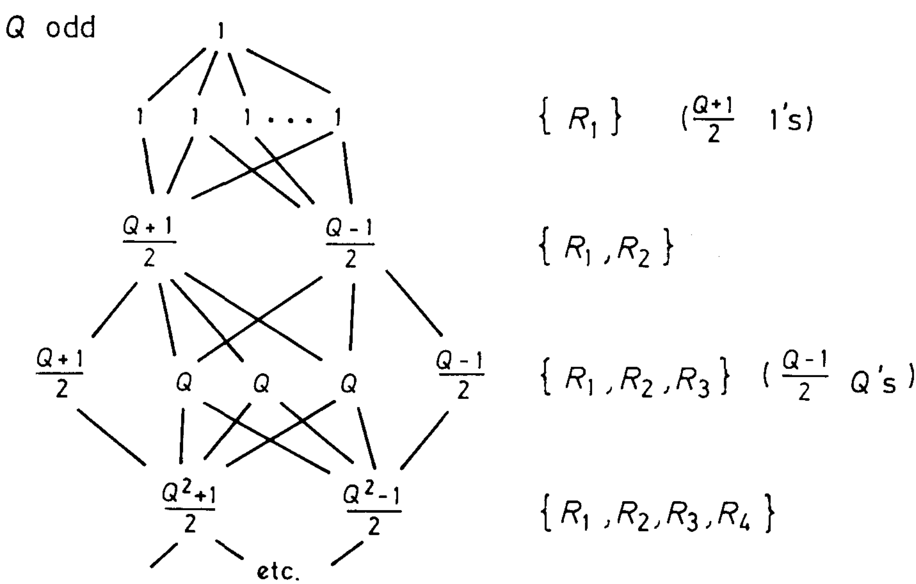
2) The Fateev-Zamolodchikov model

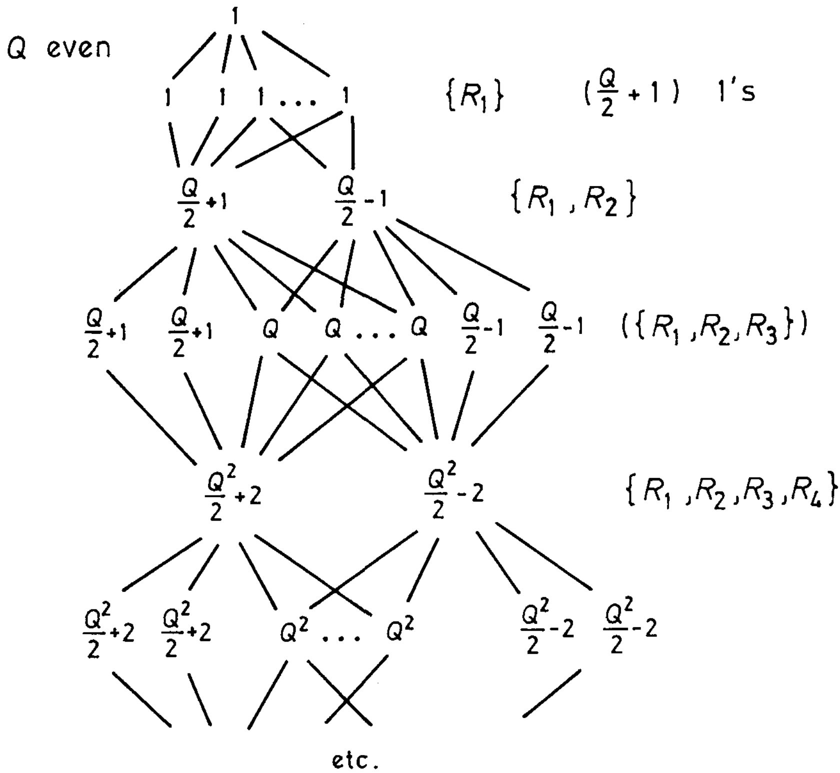
a) The algebra generated by the U_i's of Sec. 1 (common to any $\mathbb{Z}/Q\mathbb{Z}$ symmetric model).



(Note – this is the same as the tower diagram for $G = \mathbb{Z}/Q\mathbb{Z}$ in Sec. 4.

b) The algebra generated by the R_i's $(R_i = \Sigma x_{\mathrm{H}}(k, \lambda))(U_{2i-1})$





A few comments are perhaps in order since this diagram does not seem to be in the literature (a full proof is available on request).

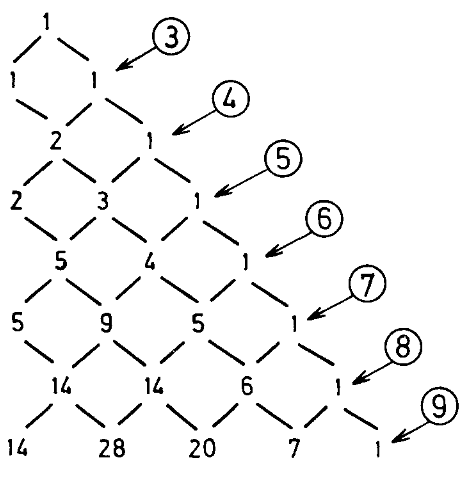
The key to the analysis is to note that $x_{\mathrm{H}}^{n}(\lambda) = x_{\mathrm{H}}^{m}(\lambda)$ iff $m = -n$. A spectral analysis of R_i then reveals that the algebra generated by R_1 and R_2 is the fixed point algebra for the period 2 automorphism of the $Q \times Q$ matrix algebra defined by $U_1 \longrightarrow U_1^{-1}, U_2 \longrightarrow U_2^{-1}$. Use of the basic construction technique then gives the whole diagram.

Note that if Q is a prime, one may take the limit $\lambda \longrightarrow i\infty$, calling σ_i the limit of R_i. Since $m^2 = n^2 \iff m = \pm n$ in this case, the algebra generated by the σ_i's is the same as that generated by the R_i's. The σ_i's define the metaplectic representation which may be analyzed by the usual techniques of character theory to give the Bratteli diagram (see [(GJ]). The method fails when Q is not prime for two reasons. The first is that the metaplectic representation is more difficult to analyze if Q is not prime, and the second, more significant, reason is that if Q is not prime, the $R_i(\lambda)$'s do not belong to the algebra generated by the braid group.

One could define a subfactor of the hyperfinite II_1 factor in the following way: the large factor is the von Neumann algebra generated by $R_1(\lambda)$, $R_2(\lambda)$, $\ldots$, and the subfactor is the von Neumann algebra generated by $R_2(\lambda)$, $R_3(\lambda)$, $\ldots$. It is clear from our analysis that the resulting subfactor is of the form $M^{D_n} \subseteq M^{\mathbb{Z}/2\mathbb{Z}}$ where D_n is the dihedral group of order $2n$ acting in some outer fashion on M. This generalizes [J3].

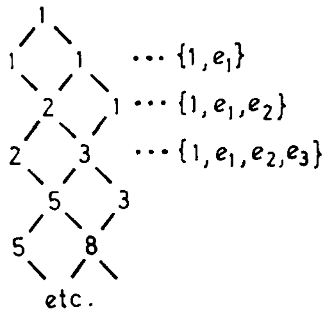
The Bratteli diagram for the Temperley-Lieb algebra in the subfactor context

For each value $4\cos^2 \pi/n$ of the index $[M : N]$ of a subfactor of a II_1 factor, the algebra generated by the e_i's in the tower of II_1 factors, is different. To obtain its structure, it is convenient to begin with the "generic" Bratteli diagram:



etc.

We have attached integers to the 1's down the diagonal. To obtain the structure of alg$\{1, e_1, e_2, \ldots\}$ when $[M : N] = 4\cos^2 \pi/n$, one eliminates the diagonal "1" with label n and everything below and to the right of it, and adjusts the remaining numbers so that the addition rule holds. Thus for $n = 5$, one obtains the diagram:



For $n = 4$ and 6, one obtains the same diagram as for the 2 and 3 state Potts models, respectively.

One feature of these Bratteli diagrams is that there is a simple way to represent the corresponding algebras on the vector space whose basis is the descending paths from the initial "1" to a given level. Thus one defines an algebra whose basis elements are pairs (γ_1, γ_2) of such paths, which end at the same point on the diagram. The algebra structure is given by $(\gamma_1, \gamma_2)(\gamma_3, \gamma_4) = \delta_{\gamma_2, \gamma_3}(\gamma_1, \gamma_4)$. It is a trivial exercise to see that this algebra has the right structure. Although this may seem to be a banal observation, it is suggestive in that one might look for simple representations of the appropriate commutation relations (e.g. Temperly-Lieb, Fateev-Zamolodchikov) on the space of paths. The existence of some representation is of course guaranteed by the knowledge that one has the correct Bratteli diagram. It is this line of reasoning that led Pasquier to his *A-D-E* models and Ocneanu to his first combinatorial descriptions of subfactors. Wenzl also used this idea to construct appropriate Hecke algebra representations in [W2].

6. Type III Factors

We return now to the GNS construction of Sec. 2, this time with a view to constructing type III factors, i.e. factors which possess no trace of any sort. So let us begin with an appropriate algebra A which we are trying to "complete" to form a type III factor. It is obviously no use beginning with a trace for the state(positive linear functional) φ, but let us simplify our discussion by assuming that $\varphi(x^*x) \neq 0$ for $x \neq 0$. A fundamental difference arises in the GNS construction if φ is not a trace: it is that, while *left* multiplication by elements of A is bounded, the same need not be, and often is not, true of right multiplication. Let us immediately give an example.

Example 6.1. The Powers factors

For A we take the infinite tensor product $\bigotimes_{i=1}^{\infty} M_2(\mathbb{C})$ just as in Sec. 2, but we define the (Powers) state φ_λ by

$$\varphi_\lambda(x_1 \otimes x_2 \otimes \ldots \otimes x_n \otimes 1) = \prod_{i=1}^{n} \text{trace}\left(x_i \begin{pmatrix} \frac{1}{1+\lambda} & 0 \\ 0 & \frac{\lambda}{1+\lambda} \end{pmatrix}\right) ,$$

where "trace" means the sum of the diagonal elements, and $0 < \lambda < 1$. Powers, in [P], showed that, if R_λ is defined to be the von Neumann algebra resulting from the GNS construction, then the R_λ are all of type III and mutually non-isomorphic. Thus type III factors exist.

In this example, only left multiplication operators are automatically bounded and extend to the whole Hilbert space. Treatment of this left-right asymmetry is the Tomita-Takesaki theory which we briefly outline. If $\mathcal{H}_\varphi$ is the GNS Hilbert space, it contains the dense subspace A. We define the (conjugate linear) operator S to be the (unbounded) operator with domain A, $S(x) = x^*$. Under suitable conditions, satisfied in the case of the Powers factor, the closure of the graph of S is again the graph of an operator also denoted S. By the von Neumann theory, S then has a polar decomposition $S = J\Delta$, where J is a conjugate linear isometry and Δ is positive and self-adjoint. If M is the von Neumann algebra generated by left multiplication operators from A on $\mathcal{H}_\varphi$, the main result of Tomita-Takesaki theory (see [Ta]) is that

a) $JMJ = M'$ (the commutant of M)

b) $\Delta^{it} M \Delta^{-it} = M$.

This is a remarkable result since it shows that to any appropriately continuous state φ on a von Neumann algebra, there is a canonical 1-parameter automorphism group σ_t^φ of M defined by applying the above procedure with $A = M$ and letting

$$\sigma_t^\varphi(x) = \Delta^{it} x \Delta^{-it} .$$

Connes ([Co3]) has shown that if the state φ is changed, σ_t^φ changes only by an inner perturbation so that one has the following invariant of a von Neumann algebra:

$$T(M) = \{t \in \mathbb{R} | \sigma_t^\varphi \text{ is an inner automorphism}\} .$$

This invariant is particularly easy to calculate for Powers factors, since the group σ_t^φ is easily seen to be conjugation by

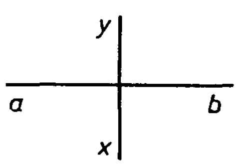
$$\left[\bigotimes_{i=1}^{\infty}\begin{pmatrix}\frac{1}{1+\lambda} & 0\\ 0 & \frac{\lambda}{1+\lambda}\end{pmatrix}\right]^{it}.$$

This is inner iff $\lambda^{it}=1$, i.e. $t\in(2\pi/\log\lambda)\mathbb{Z}$.

In general, the fixed point algebra M^{σ_φ} is called the *centralizer* of φ and it is part of the theory that φ is a trace when restricted to the centralizer (indeed, $\varphi(xy)=\varphi(yx)$ for $x\in M^{\sigma_\varphi}$, $y\in M$). In the case of the Powers factor, this centralizer is the closure of the algebra of Sec. 2, whose Bratteli diagram was Pascal's triangle (page 7).

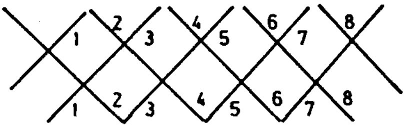
7. Vertex Models in Statistical Mechanics

The difference between vertex models and the spin models of Sec. 1 is that the states of a vertex model are functions from the edges of the graph to the Q-element set S and the energy contribution is associated with a vertex. Thus if the graph is the two-dimensional lattice, a vertex model will be defined by Boltzmann weights $w(a,b|x,y)$ associated with a vertex in the following state configuration:



where, as before, a distinction can be drawn between the horizontal and vertical directions.

A simple transfer matrix approach is to use diagonal-to-diagonal transfer matrices. Consider the diagram below:



Clearly, if we impose periodic horizontal boundary conditions and define a vector space $\otimes^n V$ ($n=8$ above, n always even) with basis vectors $\otimes_{i=1}^n q_i$, $q_i\in S$, then the transfer matrix associated with the above diagram is $(R_{23}R_{45}\dots R_{n,1})$ $(R_{12}R_{34}\dots R_{n-1,n})$, where

$$R_{i,i+1}\left(\bigotimes_{j=1}^{n} q_i\right) = \sum_{r_i,r_{i+1}=1}^{Q} w(q_i,r_{i+1}|q_{i+1},r_i)\bigotimes_{j=1}^{n} q'_j$$

with $q'_j = q_j, j \neq i, i+1, q'_j = r_j, q'_{j+1} = r_{j+1}$. The simplest such model that has been extensively studied is the "ice-type" model with $Q = 2$, solved by Lieb (see [L]). In this case, the R matrices can be expressed in the form $R_{i,i+1} = \text{const}(1 + xe_i)$, where e_i (index modulo m) is given by

$$e_i\left(\bigotimes_{i=1}^{n} q_i\right) = \sum_{r_i,r_{i+1}} x(q_i,r_{i+1}|q_{i+1}r_i)\bigotimes_{j=1}^{n} q'_i \ ,$$

with notation as before, and x is the 4×4 matrix below:

$$\frac{1}{1+t}\begin{pmatrix} 0 & 0 & 0 & 0\\ 0 & 1 & \sqrt{t} & 0\\ 0 & \sqrt{t} & t & 0\\ 0 & 0 & 0 & 0\end{pmatrix} .$$

It follows immediately that for $1 \leq i,j \leq n-1$,

$$e_i^2 = e_i, \quad e_ie_{i\pm1}e_i = \frac{t}{(1+t)^2}e_i, \quad e_ie_j = e_je_i \quad \text{if } |i-j| \geq 2 .$$

These are the same relations as we met in the Potts model and this was the key to the equivalence, established first by Temperley and Lieb in [TL], of the Potts and ice-type models. (see also [Ba] Chap. 12).

In fact, we saw in the Potts model that the partition function for the system with periodic boundary conditions in the vertical direction is given by the trace of some power of the row-to-row transfer matrix and that this trace, when normalized so that $\text{tr}(1) = 1$, has the following useful property, called the Markov property,

$$\text{tr}(we_{n+1}) = \frac{1}{Q}\text{tr}(w) \quad \text{if } w \in \text{alg}\{1, e_1, \dots, e_n\} .$$

This property completely defines tr on the algebra generated by the e_i's and it was this fact that was crucial in [J2] for the proof of the restrictions on the index values for subfactors.

It is remarkable that this trace is *not* given in the ice type model by the usual trace but rather by the restriction of the Powers state φ_λ when $2+\lambda+\lambda^{-1}=Q$ (or $\lambda=t$)! The e_i's are easily seen to be in the centralizer of the Powers state. One of the most striking simple facts in the theory of subfactors is that the von Neumann algebra generated by the e_i's is precisely the centralizer of φ_λ. This was proved by Pimsner and Popa in [PP] using entropy considerations. A direct proof is not yet available. It suffices to prove that the element $\begin{pmatrix}\lambda & 0\\ 0 & 1\end{pmatrix}\otimes 1\otimes 1\otimes$ is in the von Neumann algebra generated by the e_i's.

As far as subfactors are concerned, this example gives an interesting and useful procedure. The obvious subfactor around is the subfactor given by sums of elements of the form $\{1\otimes x_1\otimes x_2\otimes\dots\}$ in the Powers factor R_λ. This is a trivial subfactor of the form $N\subseteq N\otimes M_2(\mathbb{C})$. But restricting to the centralizer of the Powers state, one gets the interesting subfactor generated by the $\{e_i\text{'s}\}$. This procedure can obviously be generalized to arbitrary finite-dimensional unitary representations of groups. In the case of compact groups, and using the trace for the GNS construction, much analysis has been done by A. Wassermann (see [Wa]). The subfactor has trivial relative commutant precisely when the representation is irreducible.

8. Quantum Groups

Quantum groups in the spirit of Drinfeld ([Dr]) and Jimbo provided vast generalizations of the ice-type model, and in a sense this is what they were invented to do. To be precise, to every simple finite-dimensional Lie algebra, G, and every finite-dimensional representation of G, on V say, one can construct a family $R(\lambda)\in \mathrm{End}(V\otimes V)$ satisfying the Yang-Baxter equation $R_{12}(\lambda)R_{13}(\lambda+\mu)R_{23}(\mu)=R_{23}(\mu)R_{13}(\lambda+\mu)R_{12}(\lambda)$ in $\mathrm{End}(V\otimes V\otimes V)$, with obvious notation.

I do not want to go into the construction of R which is beautifully detailed in [Dr]. The six-vertex model is precisely the case $G=sl_2, V=\mathbb{R}^2$. It is clear that one is also able to use this machine to construct subfactors which generalize the $\{e_i\}$ subfactors. For $G=sl_n, V=\mathbb{C}^n$, these subfactors were considered by Wenzl in [We2]. It is not clear *a priori* what the Powers state should be in this general situation. At this point, we would have to appeal to the knot theory to justify that the Powers state on $\otimes_{r=1}^\infty \mathrm{End}(V)$ should be $\varphi_q(\otimes_{i=1}^\infty x_i)=\Pi_{i=1}^\infty$ trace $(q^{\frac{H}{2}}x_i)$, where H is the element of the Cartan subalgebra of G corresponding via the Killing form to the half sum of the positive roots. This state, which is positive definite for $q\in\mathbb{R}^+$, does have the Markov property with respect to the R_i's as shown by Rosso and Reshetikhin

([Ro], [Re]).

The subfactors created by this process do not seem to have been completely, systematically investigated. They are to be thought of as quantized versions of the subfactors considered by Wassermann.

References

[Ar] W. Arveson, *An Invitation to C^*-algebras, Graduate Texts in Mathematics* **No. 39**, (Springer-Verlag, 1976)

[A-Y BP] H. Au-Yang, R. Baxter and J. Perk, "New solutions of the star-triangle relations for the chiral Potts model", *Phys. Lett.* **A128** (1988) 138-142.

[Ba] R. Baxter, *Exactly solved models in statistical mechanics*, (Academic Press, 1982).

[BW] J. Birman and H. Wenzl, "Braids, link polynomials and a new algebra", to appear in *Trans. A.M.S.*

[Co1] A. Connes, "Classification des facteurs", *Proc. Symp. Pure Math.* **38** (1982), Part 2, 43–110.

[Co2] A. Connes, "Classification of injective factors", *Ann. Math.* **104** (1976) 73-115.

[Co3] A. Connes, "Une classification des facteurs de type III", *Ann. Sci. Ec. Norm-Sup.* **6** (1973) 133-252.

[Dr] V. Drinfeld, "Quantum Groups" in *Proc. ICM 86* Vol. 1, 798-820.

[FZ] V. Fateev and A. Zamolodchikov, "Self-dual solutions of the star-triangle relations in $\mathbb{Z}_N$-models", *Phys. Lett.* **92A** (1982) 37-39.

[GHJ] F. Goodman, P. de la Harpe and V. Jones, "Coxeter graphs and towers of algebras", to appear (Springer-Verlag).

[GJ] D. Goldschmidt and V. Jones, "Metaplectic link invariants", to appear (Geometrica Dedicata).

[J1] V. Jones, "Actions of finite groups on the hyperfinite II_1 factor", *Memoirs A.M.S.*, No. 237 (1980).

[J2] V. Jones, "Index for subfactors", *Invent. Math.* **72** (1983) 1-25.

[J3] V. Jones, "On a certain value of the Kauffman polynomial", to appear in *Comm. Math. Phys.*

[O1] A. Ocneanu, "Subalgebras are canonically fixed point algebras", *A.M.S. Abstracts* **6** (1986) 822-99-165.

[O2] A. Ocneanu, "Quantized groups, string algebras and Galois theory for algebras" — an outline of results, 1987.

[Pa] V. Pasquier, "Etiology of IRF models", Preprint Saclay SPhT/88-20.

[P] R. Powers, "Representations of uniformly hyperfinite algebras and the associated von Neumann algebras", *Ann. of Math.* **86** (1967) 138-171.

[Po] S. Popa, "Classification of subfactors — the finite depth case", UCLA preprint 1989.

[PP] M. Pimsner and S. Popa, "Entropy and index for subfactors" *Ann. Sci. Ec. Norm. Sup.* **19** (1986) 57-106.

[Re] N. Reshetikhin, "Quantized universal enveloping algebra, the Yang-Baxter equation and invariants of links I", Steklov preprint, 1988.

[Ro] M. Rosso, "Quantum groups and V. Jones' vertex models for knots", preprint 1988.

[TL] H. Temperley and E. Lieb, "Relations between the percolation...", *Proc. Roy. Soc. London* (1971) 251-280.

[Ta] M. Takesaki, "Tomita's theory of modular Hilbert algebras and its applications", *Lecture notes in Math,* No. 128 (Springer-Verlag, 1970).

[Wa] A. Wassermann, "Coactions and Yang-Baxter equations for ergodic actions and subfactors", in *Operator algebras and applications* Vol. 2, eds. D. Evans and M. Takesaki, *L.M.S. Lecture Notes* **136** (1988) 202-236.

[We1] H. Wenzl, "On the structure of Brauer's centralizer algebras", *Ann. Math.* **128** (1988) 173-194.

[We2] H. Wenzl, "Hecke algebras of type A_n and subfactors", *Invent. Math.* **92** (1988) 349-383.

[U] H. Umegaki, "Conditional expectation in an operator algebra I", *Tohoku Math. J.* **6** (1954) 358-362.