Operators
and
Matrices
Volume 2, Number 4 (2008), 555–582

GENERATORS OF II_1 FACTORS

KEN DYKEMA *, ALLAN SINCLAIR, ROGER SMITH ** AND STUART WHITE

(*communicated by D. Hadwin*)

Abstract. In 2005, Junhao Shen introduced a new invariant, $\mathscr{G}(N)$, of a diffuse von Neumann algebra N with a fixed faithful trace, and he used this invariant to give a unified approach to showing that large classes of II_1 factors M are singly generated. This paper focuses on properties of this invariant. We relate $\mathscr{G}(M)$ to the number of self-adjoint generators of a II_1 factor M: if $\mathscr{G}(M) < n/2$, then M is generated by $n+1$ self-adjoint operators, whereas if M is generated by $n+1$ self-adjoint operators, then $\mathscr{G}(M) \leqslant n/2$. The invariant $\mathscr{G}(\cdot)$ is well-behaved under amplification, satisfying $\mathscr{G}(M_t) = t^{-2}\mathscr{G}(M)$ for all $t > 0$. In particular, if $\mathscr{G}(\mathscr{L}\mathbb{F}_r) > 0$ for any particular $r > 1$, then the free group factors are pairwise non-isomorphic and are not singly generated for sufficiently large values of r. Estimates are given for forming free products and passing to finite index subfactors and the basic construction. We also examine a version of the invariant $\mathscr{G}_{\mathrm{sa}}(M)$ defined only using self-adjoint operators; this is proved to satisfy $\mathscr{G}_{\mathrm{sa}}(M) = 2\mathscr{G}(M)$. Finally we give inequalities relating a quantity involved in the calculation of $\mathscr{G}(M)$ to the free-entropy dimension δ_0 of a collection of generators for M.

1. Introduction

An old problem in von Neumann algebra theory is the question of whether each separable von Neumann algebra N is singly generated. A single generator x leads to two self-adjoint generators $\{x+x^*, i(x-x^*)\}$ and any pair $\{h,k\}$ of self-adjoint generators yields a single generator $h+ik$. Thus the single generation problem has an equivalent formulation as the existence of two self-adjoint generators. Earlier work in this area solved all cases except for the finite von Neumann algebras, [1, 16, 22, 26]. Here there has been progress in special situations, [8, 9, 19], but a general solution is still unavailable. Recently Junhao Shen, [19], introduced a numerical invariant $\mathscr{G}(N)$, and was able to show that single generation for II_1 factors was a consequence of $\mathscr{G}(N) < 1/4$. He proved that $\mathscr{G}(N) = 0$ for various classes of II_1 factors, giving a new approach to the single generation of II_1 factors with Cartan masa, with property Γ and those factorising as tensor products of II_1 factors. His work settled some previously unknown cases as well as giving a unified approach to various situations that had been determined by diverse methods. It should be noted that 0 is the only value of Shen's

Mathematics subject classification (2000): 46L10, 46L54.
Keywords and phrases: II_1 factors, generators, generator invariant, free orbit dimension.
* Research supported in part by NSF grant DMS-0600814.
** Research supported in part by NSF grant DMS-0401043.

Paper OAM-02-35

invariant that is currently known. If strictly positive values are possible, then Corollary 5.1 guarantees examples of separable II_1 factors which are not singly generated.

In this paper, our purpose is to undertake a further investigation of this invariant, and to relate it to a quantity $\mathscr{G}^{\min}(M)$ which counts the minimal number of generators for M. A related quantity $\mathscr{G}_{\mathrm{sa}}^{\min}(M)$ counts the minimal number of self-adjoint generators, and there is a parallel invariant $\mathscr{G}_{\mathrm{sa}}(M)$ to $\mathscr{G}(M)$ which has a similar definition (given below) but which restricts attention to self-adjoint generating sets.

The contents of the paper are as follows. The second section gives the definitions of $\mathscr{G}(N)$ and $\mathscr{G}_{\mathrm{sa}}(N)$ in terms of generating sets and finite decompositions of 1 as sums of orthogonal projections. This is a slightly different but equivalent formulation of the original one in [19]. These are related by the inequalities $\mathscr{G}(N) \leqslant \mathscr{G}_{\mathrm{sa}}(N) \leqslant 4\mathscr{G}(N)$, although it is shown subsequently that $\mathscr{G}(M) = 2\mathscr{G}_{\mathrm{sa}}(M)$ for all II_1 factors M. The main result of the third section is that the relation $\mathscr{G}(M) < n/2$ for II_1 factors M implies generation by $n+1$ self-adjoint elements. The case $n = 1$ is of particular interest since single generation is then a consequence of $\mathscr{G}(M) < 1/2$.

The fourth section relates the generator invariant of a II_1 factor M to that of a compression pMp. If $\tau(p) = t$, then $\mathscr{G}(pMp) = t^{-2}\mathscr{G}(M)$. Up to isomorphism, M_t can be uniquely defined as pMp for any projection $p \in M$ with $\tau(p) = t$, $0 < t < 1$. In a standard way, M_t can be defined for any $t > 0$ as $p(\mathbb{M}_n \otimes M)p$ where n is any integer greater than t, and $p \in \mathbb{M}_n \otimes M$ is a projection of trace $t/n < 1$. In this more general situation, the scaling formula $\mathscr{G}(M_t) = t^{-2}\mathscr{G}(M)$ for $t > 0$ also holds. The subsequent section contains some consequences of the scaling formula, and also establishes it for the related invariant $\mathscr{G}_{\mathrm{sa}}(M)$. This requires a more indirect argument since the method of passing between generating sets for M and for M_t does not preserve self-adjointness and so cannot apply to $\mathscr{G}_{\mathrm{sa}}(M_t)$ although it is suitable for $\mathscr{G}(M_t)$. The equality $\mathscr{G}_{\mathrm{sa}}(M) = 2\mathscr{G}(M)$ is also established in this section.

The sixth section is concerned with finite index inclusions $N \subseteq M$ of II_1 factors. The main results are that $\mathscr{G}(\langle M, e_N\rangle) \leqslant \mathscr{G}(M)$ and that $\mathscr{G}(N) = \lambda^2\mathscr{G}(\langle M, e_N\rangle)$, where $\langle M, e_N\rangle$ is the basic construction and λ denotes the index $[M : N]$. A standard result of subfactor theory is that M is the basic construction $\langle N, e_P\rangle$ for an index λ inclusion $P \subseteq N$, so two of these basic constructions scale $\mathscr{G}(\cdot)$ by λ^2. This suggests the formula $\mathscr{G}(\langle M, e_N\rangle) = \lambda\mathscr{G}(M)$, but this is still an open problem.

Section 7 concentrates on free group factors and their generalisations, the interpolated free group factors. For $r \in (0, \infty]$, the formula $\mathscr{G}(\mathscr{L}\mathbb{F}_{1+r}) = r\alpha$ is established, where α is a fixed constant in the interval $[0, 1/2]$. This leads to two possibilities, depending on the value of α. If $\alpha = 0$, then $\mathscr{L}(\mathbb{F}_{1+r})$ is singly generated for all $r > 0$, while if $\alpha > 0$, then the free group factors are pairwise non-isomorphic, being distinguished by the generator invariant. The paper concludes with a discussion of Voiculescu's modified free entropy dimension $\delta_0(X)$, where X is a finite generating set for M. A quantity $\mathscr{I}(X)$ is introduced in the second section on the way to defining $\mathscr{G}(M)$. The main results of the last section are the inequalities $\delta_0(X) \leqslant 1 + 2\mathscr{I}(X)$ for general finite generating sets, and the stronger form $\delta_0(X) \leqslant 1 + \mathscr{I}(X)$ for generating sets of self-adjoint elements. These have the potential for providing lower bounds for $\mathscr{G}(M)$.

We thank the referee for bringing the following result to our attention, which works in a general unital C^*-algebra. We include a proof for completeness. The scaling formula for the generator invariant in section 4 can be thought of as a continuous version of this fact.

PROPOSITION 1.1. *Let $n,t \in \mathbb{N}$ with $t \geqslant 2$. Suppose that M is a unital C^*-algebra generated by $(t-1)n^2+1$ hermitian operators. Then $M \otimes \mathbb{M}_n$ is generated by t hermitian operators.*

Proof. When $t = 2$, consider n^2+1 hermitian generators of M labeled as $x_1,\dots,x_n,y_1,\dots,y_n,z_1,\dots,z_{n-1}$ and $f_{j,k},g_{j,k}$ for $1 \leqslant j \leqslant n-1$ and $j+1<k \leqslant n$. By scaling and adding copies of the identity, we can assume that the spectrum of each x_j lies in $[2j,2j+1]$ and that each z_j is positive and invertible. Then the C^*-algebra A generated by the hermitian operators

$$\begin{pmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_n \end{pmatrix}$$

and

$$\begin{pmatrix} y_1 & z_1 & f_{1,3}+ig_{1,3} & \dots & f_{1,n}+ig_{1,n} & \\ z_1 & y_2 & z_2 & \dots & f_{2,n}+ig_{2,n} & \dots \\ f_{1,3}-ig_{1,3} & z_2 & y_3 & \dots & f_{3,n}+ig_{3,n} & \\ \vdots & \ddots & \ddots & \ddots & \vdots & \\ f_{1,n}-ig_{1,n} & \dots & f_{n-2,n}-ig_{n-2,n} & z_{n-1} & y_n & \end{pmatrix}$$

is all of $\mathbb{M}_n(M)$. This will follow immediately once we have shown that all scalar matrices lie in A. By applying the functional calculus to the first generator, we see that scalar diagonal matrices lie in A. Regard $\mathbb{M}_n(M)$ as $M \otimes \mathbb{M}_n$, writing $(e_{i,j})$ for the canonical matrix units of $\mathbb{M}_n$. Pre and post-multiplication by diagonal matrices shows that $z_1 \otimes e_{1,2} \in A$. Then multiply this element by its adjoint to obtain $z_1^2 \otimes e_{1,1} \in A$ and apply the functional calculus to get $z_1^{-1} \otimes e_{1,1} \in A$. Multiplying this by $z_1 \otimes e_{1,2}$ shows that the matrix unit $1 \otimes e_{1,2}$ lies in A. The same technique can be used to see that $e_{j,j+1} \in A$ for each j. Thus all scalar matrices lie in A, as claimed.

When $t > 2$, suppose the additional $(t-2)n^2$ hermitians $a_1^{(s)},\dots,a_n^{(s)}$ and $b_{j,k}^{(s)}, c_{j,k}^{(s)}$ for $1 \leqslant i < j \leqslant n$ and $1 \leqslant s \leqslant t-2$ are required to generate M. Adding the generators

$$\begin{pmatrix} a_1^{(s)} & b_{1,2}^{(s)}+ic_{1,2}^{(s)} & \dots & b_{1,n}^{(s)}+ic_{1,n}^{(s)} \\ b_{1,2}^{(s)}-ic_{1,2}^{(s)} & a_2^{(s)} & \dots & b_{2,n}^{(s)}+ic_{2,n}^{(s)} \\ \vdots & \ddots & \ddots & \vdots \\ b_{1,n}^{(s)}-ic_{1,n}^{(s)} & \dots & b_{n-1,n}^{(s)}-ic_{n-1,n}^{(s)} & a_n^{(s)} \end{pmatrix}$$

to those listed above, gives a total of $(t-1)n^2+1$ hermitian operators which generate $\mathbb{M}_n(M)$. $\square$

Finally, a word on notation. For a subset X of a von Neumann algebra M, $W^*(X)$ will denote the von Neumann algebra generated by X. It is not assumed that $W^*(X)$ automatically contains the identity of M. For example, $W^*(p) = \mathbb{C}p$ for a projection $p \in M$.

2. The generator invariant

The main focus of the paper is on II_1 factors. However we define the generator invariant and establish basic results in the context of diffuse finite von Neumann algebras with a fixed faithful trace τ, which is normalised with $\tau(1) = 1$. Diffuse finite von Neumann algebras will be denoted by N, while M is reserved for II_1 factors.

DEFINITION 2.1. Let (N,τ) be a finite von Neumann algebra with fixed trace. Let $\mathscr{P}$ (or $\mathscr{P}(N)$ when the underlying algebra is unclear) denote the collection of all finite sets of mutually orthogonal projections in N which sum to 1. An important subclass of $\mathscr{P}$ is the collection $\mathscr{P}_{\mathrm{eq}}$ of those P which consist of projections of equal trace. Note that in the context of a finite factor M, $P = \{p_1,\dots,p_k\} \in \mathscr{P}_{\mathrm{eq}}(M)$ can be diagonalised in the sense that there exist matrix units $(e_{i,j})_{i,j=1}^k$ in M satisfying $e_{i,i} = p_i$ for all i. In this factor context, the elements of $\mathscr{P}_{\mathrm{eq}}$ are referred to as the *diagonalisable* elements of $\mathscr{P}$.

DEFINITION 2.2. Consider $P, Q \in \mathscr{P}$. Then Q *refines* P, written $Q \preceq P$, when every $p \in P$ is a sum of elements of Q. The sets P and Q are *equivalent*, written $P \sim Q$, when there is a unitary $u \in N$ with $uPu^* = Q$. Note that in a II_1 factor, $P \sim Q$ if, and only if, the multiset of the traces of the elements of P is the same as the multiset of the traces of the elements in Q.

The ordering $Q \preceq P$ defined above is chosen so that the map $P \mapsto \mathscr{I}(X;P)$ below is order preserving. This will be established in Lemma 2.5.

DEFINITION 2.3. Let N be a finite von Neumann algebra. Given $x \in N$ and $P \in \mathscr{P}$, define

$$\mathscr{I}(x;P) = \sum_{\substack{p,q\in P\\ pxq\neq 0}} \tau(p)\tau(q).$$

For a finite subset $X \subset N$ and $P \in \mathscr{P}$, define

$$\mathscr{I}(X;P) = \sum_{x\in X} \mathscr{I}(x;P),$$

and

$$\mathscr{I}(X) = \inf_{P\in\mathscr{P}} \mathscr{I}(X;P). \tag{2.1}$$

On occasion, the notation $\mathscr{I}_N(X;P)$ and $\mathscr{I}_N(X)$ will emphasise the algebra N and hence the choice of trace.

The definition of $\mathscr{I}$ given above is formally different from that of Shen [19, Definition 2.1], in that Shen only considers families from $\mathscr{P}_{\mathrm{eq}}$ and performs the limiting procedure in a slightly different order. Nevertheless, the resulting invariant $\mathscr{G}(N)$

defined in Definition 2.7 below agrees with [19, Definition 2.1] in the case of diffuse von Neumann algebras. Before proceeding, a few elementary observations are recorded.

REMARKS 2.4.

1. The inequality
$$0 \leqslant \mathscr{I}(X) \leqslant \mathscr{I}(X;P) \leqslant |X|$$
holds for all finite subsets X in N and all $P \in \mathscr{P}$.
2. If X is a finite subset of N, then define $X^* = \{x^* | x \in X\}$. The equality
$$\mathscr{I}(X^*;P) = \mathscr{I}(X;P)$$
holds for all $P \in \mathscr{P}$.
3. Let $P \in \mathscr{P}$ be a set of k projections. The estimate
$$\mathscr{I}(x;P) = \sum_{\substack{p\in P\\ px\neq 0}} \tau(p)^2 \leqslant k \max\{\tau(p)^2 | p \in \mathscr{P}\}$$
is valid for each $x \in N \cap P'$. In particular, if $P \in \mathscr{P}_{\mathrm{eq}}$, then
$$\mathscr{I}(x;P) \leqslant k^{-1}.$$
4. If $X = \{x_1, \dots, x_m\} \subset N$ and $P = \{p_1, \dots, p_k\} \in \mathscr{P}_{\mathrm{eq}}$ then
$$\mathscr{I}(X;P) = k^{-2} \left|\{(i,j,l) | p_i x_l p_j \neq 0\}\right|.$$
Writing $\mathscr{I}(X;P)$ in this way is often useful in calculations, an example being [19, Theorem 4.1].
5. For any $z_1, z_2 \in N$ and $P \in \mathscr{P}$,
$$\mathscr{I}(z_1 + z_2;P) \leqslant \mathscr{I}(z_1;P) + \mathscr{I}(z_2;P). \tag{2.2}$$
6. If X is a finite generating set of N, then so also is
$$Y = \{x + ix^*, i(x - ix^*) | x \in X\}$$
and, furthermore,
$$\mathscr{I}(X;P) \leqslant \mathscr{I}(Y;P) \leqslant 4\mathscr{I}(X;P).$$
This follows from the preceding remark. Take $z_1 = x + ix^*$ and $z_2 = x - ix^*$ for each $x \in X$ to obtain $\mathscr{I}(X;P) \leqslant \mathscr{I}(Y;P)$. Now let $Y_1 = \{x + ix^* | x \in X\}$ and $Y_2 = \{i(x - ix^*) | x \in X\}$ so that (2.2) and item 2 above combine to give
$$\mathscr{I}(Y_1;P) \leqslant 2\mathscr{I}(X;P) \quad \text{and} \quad \mathscr{I}(Y_2;P) \leqslant 2\mathscr{I}(X;P). \tag{2.3}$$

LEMMA 2.5. *Let N be a finite von Neumann algebra. Consider a finite subset $X \subset N$ and $P, Q \in \mathscr{P}$ with $Q \preceq P$. Then*
$$\mathscr{I}(X;Q) \leqslant \mathscr{I}(X;P).$$

Proof. Take $x \in N$ and two pairs of orthogonal projections (e_1, e_2) and (f_1, f_2) in N. If $(e_1+e_2)x(f_1+f_2) \neq 0$, then

$$\sum_{\substack{i,j \\ e_i x f_j \neq 0}} \tau(e_i)\tau(f_j) \leqslant \tau(e_1+e_2)\tau(f_1+f_2).$$

The result now follows by induction. □

In many applications it will be useful to know that the infimum defining $\mathscr{I}(X)$ in (2.1) can be taken through $\mathscr{P}_{\mathrm{eq}}$. Recall that a separable diffuse abelian von Neumann algebra A is isomorphic to $L^\infty[0,1]$ and, if A is equipped with a trace, then this isomorphism can be chosen so that the trace is given by $\int_0^1 \cdot \,\mathrm{d}t$ on $L^\infty[0,1]$.

LEMMA 2.6. *Let N be a diffuse von Neumann algebra and let X be a finite subset of N. For each $n \in \mathbb{N}$,*

$$\begin{aligned}\mathscr{I}(X) &= \inf\{\mathscr{I}(X;P) | P \in \mathscr{P}_{\mathrm{eq}}, \quad |P| = nl \text{ some } l \in \mathbb{N}\} \\ &= \lim_{l\to\infty}(\inf\{\mathscr{I}(X;P) | P \in \mathscr{P}_{\mathrm{eq}}, \quad |P| = nl\}).\end{aligned}$$

Proof. Fix $\varepsilon > 0$ and find $Q = \{q_1, \ldots, q_k\} \in \mathscr{P}$ with

$$\mathscr{I}(X;Q) \leqslant \mathscr{I}(X) + \varepsilon. \tag{2.4}$$

Let A be a diffuse abelian subalgebra of N with $Q \subset A$, and let

$$\delta = \frac{\varepsilon}{2|X|k}. \tag{2.5}$$

By approximating the family $\tau(q_1), \ldots, \tau(q_k)$ by rational numbers with common denominator nl, there exists $l_0 \in \mathbb{N}$ such that if $l \geqslant l_0$, then there are $r_1, \ldots, r_k \in \mathbb{N}$ with

$$\frac{r_i}{nl} \leqslant \tau(q_i) \leqslant \frac{r_i}{nl} + \delta$$

for each i. Choose projections $p_1, \ldots, p_k$ in A with $0 \leqslant p_i \leqslant q_i$ and $\tau(p_i) = r_i/nl$ for each i. Let $p_{k+1} = 1 - \sum_{i=1}^k p_i$, so $\tau(p_{k+1}) \leqslant k\delta$. Let $P_0 = \{p_1, \ldots, p_{k+1}\} \in \mathscr{P}$. Thus $q_i x q_j \neq 0$ whenever $p_i x p_j \neq 0$. Hence

$$\mathscr{I}(X;Q) = \sum_{x\in X} \sum_{\substack{1\leqslant i,j\leqslant k \\ q_i x q_j \neq 0}} \tau(q_i)\tau(q_j) \geqslant \sum_{x\in X} \sum_{\substack{1\leqslant i,j\leqslant k \\ p_i x p_j \neq 0}} \tau(p_i)\tau(p_j). \tag{2.6}$$

Now

$$\begin{aligned}\mathscr{I}(X;P_0) \leqslant{}& \sum_{x\in X} \sum_{\substack{1\leqslant i,j\leqslant k \\ p_i x p_j \neq 0}} \tau(p_i)\tau(p_j) \\ &+ |X| \left(\sum_{i=1}^k \tau(p_i)\tau(p_{k+1}) + \sum_{j=1}^k \tau(p_{k+1})\tau(p_j) + \tau(p_{k+1})^2 \right) \\ \leqslant{}& \mathscr{I}(X;Q) + 2|X|\tau(p_{k+1}) \\ \leqslant{}& \mathscr{I}(X;Q) + 2k|X|\delta \leqslant \mathscr{I}(X;Q) + \varepsilon,\end{aligned}$$

by inequality (2.6), the estimate $\tau(p_{k+1}) \leqslant k\delta$ and the choice of δ in (2.5). Finally, refine P_0 to find a family $P \preceq P_0$ in $\mathscr{P}_{\mathrm{eq}}$ with nl elements. Lemma 2.5 gives

$$\mathscr{I}(X) \leqslant \mathscr{I}(X;P) \leqslant \mathscr{I}(X;P_0) \leqslant \mathscr{I}(X;Q) + \varepsilon < \mathscr{I}(X) + 2\varepsilon,$$

by (2.4). Since this can be done for any $l \geqslant l_0$, and $\varepsilon > 0$ was arbitrary, the result follows. □

These preliminaries allow the *generator invariant* of a diffuse finite von Neumann algebra to be defined.

DEFINITION 2.7. If N is a finitely generated diffuse finite von Neumann algebra, define the *generator invariant* $\mathscr{G}(N)$ by

$$\mathscr{G}(N) = \inf\{\mathscr{I}(X) | W^*(X) = N, \quad X \text{ is a finite subset of } N\},$$

and the *hermitian, (or self–adjoint), generator invariant*, $\mathscr{G}_{\mathrm{sa}}(N)$ by

$$\mathscr{G}_{\mathrm{sa}}(N) = \inf\{\mathscr{I}(X) | W^*(X) = N, \quad X \text{ is a finite subset of } N_{\mathrm{sa}}\}.$$

If N is not finitely generated, then define $\mathscr{G}(N) = \mathscr{G}_{\mathrm{sa}}(N) = \infty$.

REMARK 2.8. By item 2.3 of Remarks 2.4, the inequalities

$$\mathscr{G}(N) \leqslant \mathscr{G}_{\mathrm{sa}}(N) \leqslant 4\mathscr{G}(N)$$

hold for any diffuse finite von Neumann algebra N. In Theorem 5.5 it will be shown that $\mathscr{G}(M) = 2\mathscr{G}_{\mathrm{sa}}(M)$ for all II_1 factors M.

3. The generation theorem

Theorem 4.1 of [19] states that if a II_1 factor M has $\mathscr{G}(M) < 1/4$, then it is generated by a projection and a hermitian element. It is subsequently remarked that the same proof shows that a II_1 factor M with $\mathscr{G}(M) < 1/2$ is singly generated. The theorem below strengthens this result to consider II_1 factors which may not be singly generated, with Shen's remark arising from the case $n = 1$. The basic idea is the same combinatorical counting argument of [19] which dates back through [8] and [9] to work in the 1960's by Douglas, Pearcy and Wogen, [1, 16, 26]. A good account of this material can be found in the book by Topping, [22].

Recall that if $(e_{i,j})_{i,j=1}^{k}$ are matrix units for the $k \times k$ matrices, $\mathbb{M}_k(\mathbb{C})$, then the self–adjoint elements $\sum_{i=1}^{k-1}(e_{i,i+1} + e_{i+1,i})$ and $e_{k,k}$ generate $\mathbb{M}_k(\mathbb{C})$.

THEOREM 3.1. (The generation theorem) *Let M be a separable II_1 factor and $n \in \mathbb{N}$. If $\mathscr{G}(M) < n/2$, then M is generated by $n+1$ hermitian elements.*

Proof. Suppose that $\mathscr{G}(M) < n/2$ for some $n \in \mathbb{N}$. There exists $k_0 \in \mathbb{N}$ such that

$$\mathscr{G}(M) < \frac{n}{2} - \left(\frac{n+2}{2k} - \frac{1}{k^2}\right)$$

for all $k \geqslant k_0$. By Lemma 2.6, there is a finite set $X = \{x_1, \ldots, x_m\}$ of generators for M, some $k \geqslant k_0$ and a diagonalisable family $P = \{e_1, \ldots, e_k\} \in \mathscr{P}_{\mathrm{eq}}$ such that

$$\mathscr{I}(X;P) < \frac{n}{2} - \left(\frac{n+2}{2k} - \frac{1}{k^2}\right). \tag{3.1}$$

Choose a set of matrix units $(e_{i,j})_{i,j=1}^k$ in M with $e_{i,i} = e_i$ for each i. Define a set of triples

$$T = \{(i,j,l) | 1 \leqslant i,j \leqslant k, \quad 1 \leqslant l \leqslant m, \quad e_i x_l e_j \neq 0\} \tag{3.2}$$

so the definition of $\mathscr{I}(X;P)$ gives

$$\mathscr{I}(X;P) = k^{-2}|T|. \tag{3.3}$$

For each $r = 1, \ldots, n-1$, let S_r be the set of triples (s,t,r) with $1 \leqslant s < t \leqslant k$ and let S_n be the set of triples (s,t,n) with $1 \leqslant s < t \leqslant k-1$. Then

$$\left|\bigcup_{q=1}^n S_q\right| = \frac{(n-1)k(k-1)}{2} + \frac{(k-1)(k-2)}{2} = \frac{nk^2}{2} - \frac{(n+2)k}{2} + 1.$$

By the choice of $k \geqslant k_0$,

$$|T| < \frac{nk^2}{2} - \frac{(n+2)k}{2} + 1 = \left|\bigcup_{r=1}^n S_r\right|,$$

from (3.1) and (3.3).

Decompose T into a partition $\bigcup_{r=1}^n T_r$ with each $|T_r| \leqslant |S_r|$ and find, for each r, injections $T_r \to S_r$, written as $(i,j,l) \mapsto (s(i,j,l), t(i,j,l), r)$. This really defines n maps indexed by r, but as the domains T_r and ranges S_r are disjoint, these are regarded as a single map from $T = \bigcup T_r$ to $\bigcup S_r$. Define self–adjoint operators

$$y_r = \sum_{(i,j,l)\in T_r} \left(e_{s(i,j,l),i} x_l e_{j,t(i,j,l)} + e_{t(i,j,l),j} x_l^* e_{i,s(i,j,l)}\right)$$

for $1 \leqslant r \leqslant n$.

If $1 \leqslant r \leqslant n$ and $(i_1,j_1,l_1) \in T_r \subset T$, then $(s(i_1,j_1,l_1), t(i_1,j_1,l_1))$ appears exactly once in the set

$$\{(s,t),(t,s) | (s,t,r) \in S_r\}$$

as S_r is disjoint from its transpose on the first two variables and the map $T_r \to S_r$ is an injection. For such $(i_1,j_1,l_1) \in T_r$,

$$\begin{aligned} e_{i_1,s(i_1,j_1,l_1)} y_r e_{t(i_1,j_1,l_1),j_1} &= e_{i_1,s(i_1,j_1,l_1)} e_{s(i_1,j_1,l_1),i_1} x_{l_1} e_{j_1,t(i_1,j_1,l_1)} e_{t(i_1,j_1,l_1),j_1} \\ &= e_{i_1} x_{l_1} e_{j_1} \neq 0. \end{aligned} \tag{3.4}$$

As $T = \bigcup_{r=1}^n T_r$, equation (3.4) and the definition of T in (3.2) imply that

$$x_l = \sum_{(i,j,l)\in T} e_{i,i} x_l e_{j,j} = \sum_{r=1}^n \sum_{(i,j,l)\in T_r} e_{i,s(i,j,l)} y_r e_{t(i,j,l),j)},$$

for each $l = 1, \dots, m$. Thus the set $\{y_1, \dots, y_n\} \cup \{e_{i,j} | 1 \leqslant i, j \leqslant k\}$ generates the II_1 factor M.

Finally note that $e_{k,k} y_n = y_n e_{k,k} = 0$ so that $y_n, e_{k,k} \in W^*(\{y_n + \lambda e_{k,k}\})$ for any $\lambda > \|y_n\|$ by the spectral theorem. In this way M is generated by the $n+1$ hermitian elements

$$y_1, \dots, y_{n-1}, \quad y_n + \lambda e_{k,k}, \text{ and } \sum_{i=1}^{k-1} (e_{i,i+1} + e_{i+1,i}),$$

as required. □

REMARK 3.2. In Corollary 5.7, it will be shown that if M is generated by $n+1$ hermitian elements, then $\mathscr{G}(M) \leqslant n/2$, which almost gives a converse to this result. The remaining gap is summarised in the following question.

QUESTION 3.3. Let M be a II_1 factor. If $\mathscr{G}(M) = n/2$ for some $n \in \mathbb{N}$, is M generated by $n+1$ hermitians?

4. A scaling formula

This section examines the behaviour of $\mathscr{G}(M)$ under amplifications and compressions. Theorem 4.5 establishes that

$$\mathscr{G}(M_t) = t^{-2} \mathscr{G}(M) \tag{4.1}$$

for all II_1 factors M and $t > 0$. Note that it is a consequence of equation (4.1) that M is finitely generated if and only if M_t is finitely generated for all $t > 0$. This result is deduced in Lemma 4.4 from the lemmas that will be needed to prove equation (4.1).

LEMMA 4.1. *Let M be a separable II_1 factor and let $n \in \mathbb{N}$. Then*

$$\mathscr{G}(M_{n^{-1}}) \leqslant n^2 \mathscr{G}(M).$$

Proof. Assume that $\mathscr{G}(M) < \infty$ as otherwise the lemma is vacuous. Let $\varepsilon > 0$. By Lemma 2.6, there is a finite generating set X for M, some $k \in \mathbb{N}$ and a diagonalisable $P = \{e_1, \dots, e_{nk}\} \in \mathscr{P}_{\mathrm{eq}}$ such that $\mathscr{I}(X; P) < \mathscr{G}(M) + \varepsilon$. Find matrix units $(e_{i,j})_{i,j=1}^{nk}$ in M with $e_{i,i} = e_i$. Define

$$f_{r,s} = \sum_{i=1}^{k} e_{(r-1)k+i,(s-1)k+i}$$

for $r, s = 1, \dots, n$. The family $(f_{r,s})_{r,s=1}^{n}$ is a system of matrix units in M. Consider $f = f_{1,1}$, a projection of trace n^{-1} so fMf is a representative of $M_{n^{-1}}$.

The von Neumann algebra fMf is generated by $\bigcup_{r,s=1}^{n} f_{1,r} X f_{s,1}$ using induction on the equation

$$fxyf = \sum_{r=1}^{n} f_{1,1} x f_{r,1} f_{1,r} y f_{1,1},$$

see for example [25, Lemma 5.2.1]. Consider $Q = \{e_1, \ldots, e_k\}$, a family of orthogonal projections in fMf with sum f, so Q is a diagonalisable element of $\mathscr{P}_{\mathrm{eq}}(fMf)$. For $x \in X$, $i,j = 1, \ldots, k$ and $r, s = 1, \ldots, n$ the relation $e_i f_{1,r} x f_{s,1} e_j \neq 0$ is equivalent to $e_{(r-1)k+i} x e_{(s-1)k+j} \neq 0$ since

$$e_i f_{1,r} x f_{s,1} e_j = e_{i,(r-1)k+i} x e_{(s-1)k+j,j}.$$

Let C be the number of quintuples (i,j,r,s,x) with this property.

Since each projection in Q has trace k^{-1} in fMf, it follows that

$$\mathscr{I}_{fMf}\Big(\bigcup_{r,s=1}^{n} f_{1,r} X f_{s,1}; Q\Big) = C\frac{1}{k^2},$$

whereas

$$\mathscr{I}_M(X;P) = C\frac{1}{n^2k^2}.$$

Therefore

$$\begin{aligned} \mathscr{G}(M_{n^{-1}}) &= \mathscr{G}(fMf) \\ &\leqslant \mathscr{I}_{fMf}\left(\bigcup_{r,s=1}^{n} f_{1,r} X f_{s,1}; Q\right) \\ &= n^2 \mathscr{I}_M(X;P) \leqslant n^2 \mathscr{G}(M) + n^2\varepsilon, \end{aligned}$$

from which the lemma follows. □

REMARK 4.2. It does not appear to be possible to obtain

$$\mathscr{G}_{\mathrm{sa}}(M_{n^{-1}}) \leqslant n^2 \mathscr{G}_{\mathrm{sa}}(M) \tag{4.2}$$

using the methods of Lemma 4.1, since a self–adjoint generating set X leads to a generating set $\bigcup_{r,s=1}^{n} f_{1,r} X f_{s,1}$ for fMf which is not necessarily self–adjoint. Nevertheless inequality (4.2) is true, as will be established in Corollary 5.6 from the scaling formula of Theorem 4.5 and Theorem 5.5.

The next lemma provides one inequality in (4.1) and also the parallel inequality for the hermitian–generator invariant.

LEMMA 4.3. *Let M be a separable II_1 factor and $0 < t < 1$. Then*

$$\mathscr{G}(M) \leqslant t^2 \mathscr{G}(M_t) \quad \text{and} \quad \mathscr{G}_{\mathrm{sa}}(M) \leqslant t^2 \mathscr{G}_{\mathrm{sa}}(M_t).$$

Proof. Assume that M_t is finitely generated, otherwise there is nothing to prove. Fix a projection $p \in M$ of trace t so that pMp is a representative of M_t and let X be an arbitrary finite set of generators for pMp.

For $\varepsilon > 0$, find orthogonal projections $E = \{e_1, \ldots, e_n\} \in \mathscr{P}_{pMp}$ such that

$$\mathscr{I}_{pMp}(X;E) = t^{-2} \sum_{x \in X} \sum_{e_i x e_j \neq 0} \tau_M(e_i)\tau_M(e_j) < \mathscr{I}_{pMp}(X) + \varepsilon. \tag{4.3}$$

The factor t^{-2} arises in (4.3) as $\tau_{M_t}(y) = t^{-1}\tau_M(y)$ for $y \in M_t$. Let m be the maximal integer such that $mt \leqslant 1$ and find a family of orthogonal projections $p_1, \dots, p_{m+1}$ in M such that:

i. $p_1 = p$;
ii. $\tau(p_i) = \tau(p)$, for $i = 2, \dots, m$;
iii. $\sum_{i=1}^{m+1} p_i = 1$. In this way $0 \leqslant \tau(p_{m+1}) < \tau(p)$ with p_{m+1} possibly the zero projection. Let $v_1 = p_1$ and find partial isometries $v_2, \dots, v_{m+1} \in M$ such that:

1. $v_i v_i^* = p_i$, for each $i = 2, \dots, m+1$;
2. $v_i^* v_i = p_1$, for $i = 2, \dots, m$;
3. $v_{m+1}^* v_{m+1}$ is a subprojection of p_1 of the form $\sum_{j=1}^{k-1} e_j + \widetilde{e_{k+1}}$ for some $k \in \{1, \dots, n\}$ and some $\widetilde{e_k} \leqslant e_k$.

It will now be shown that

$$Y = X \cup \{v_2, \dots, v_{m+1}\} \tag{4.4}$$

generates the II_1 factor M. Since X generates pMp, it follows that $pMp \subset W^*(Y)$. The relations

$$p_i M p_j = v_i v_i^* M v_j v_j^* = v_i p v_i^* M v_j p v_j^*,$$

for $1 \leqslant i,j \leqslant m+1$, then imply that

$$p_i M p_j \subset v_i p M p v_j^* \subset v_i W^*(Y) v_j^* \subset W^*(Y),$$

and so Y generates M.

The final step is to estimate $\mathscr{I}(Y)$. Let F be a set of mutually orthogonal projections with sum p that refines $\{e_1, \dots, e_{k-1}, \widetilde{e_k}, e_k - \widetilde{e_k}, e_{k+1}, \dots, e_n\}$ such that

$$\max_{f \in F} \tau(f) < \frac{\varepsilon}{m}. \tag{4.5}$$

Thus $F \in \mathscr{P}_{pMp}$ and $F \preceq E$. Extend F to a family

$$G = \{v_i f v_i^* | 1 \leqslant i \leqslant m+1,\ f \in F\}$$

of projections. Then $G \supset F$ and the sum of the orthogonal projections in G is 1 so $G \in \mathscr{P}_M$. Fix $i \in \{2, \dots, m+1\}$ and note that if $v_j f_1 v_j^* v_i v_k f_2 v_k^* \neq 0$ for some j,k and $f_1, f_2 \in F$, then $j = i$ (as otherwise $v_j^* v_i = 0$) and $k = 1$ (as otherwise $v_i v_k = 0$); when these conditions hold $f_1 = f_2$ (as $f_1 v_j^* v_i v_k f_2 = f_1 p_1 f_2 = f_1 f_2$). Thus

$$\mathscr{I}_M(v_i; G) \leqslant \sum_{f \in F} \tau(f)^2 \leqslant \max_{f \in F} \tau(f) < \frac{\varepsilon}{m}, \tag{4.6}$$

and Lemma 2.5 and (4.3) imply that

$$\begin{aligned} \mathscr{I}_M(X; G) = \mathscr{I}_M(X; F) &\leqslant \mathscr{I}_M(X; E) \\ &= \sum_{\substack{e_1, e_2 \in E, x \in X \\ e_1 x e_2 \neq 0}} \tau(e_1)\tau(e_2) < t^2 \mathscr{I}_{pMp}(X) + \varepsilon t^2. \end{aligned} \tag{4.7}$$

Finally

$$\begin{aligned}\mathscr{G}(M) &\leqslant \mathscr{I}_M(Y;G) = \mathscr{I}_M(X;F) + \sum_{i=2}^{m+1} \mathscr{I}_M(v_i;G) \\ &\leqslant \mathscr{I}_M(X;E) + \varepsilon < t^2 \mathscr{I}_{pMp}(X) + \varepsilon t^2 + \varepsilon. \qquad (4.8)\end{aligned}$$

The inequality $\mathscr{G}(M) \leqslant t^2\mathscr{G}(M_t)$ then follows, as $\varepsilon > 0$ was arbitrary and X was an arbitrary finite generating set for pMp.

Only minor modifications are necessary for the hermitian case. Assume that each element of X is self-adjoint, and replace the partial isometries v_i in (4.4) by the self-adjoint elements $v_i + v_i^*$ and $i(v_i - v_i^*)$. Then the same arguments lead to the inequality

$$\mathscr{G}_{\rm sa}(M) < t^2 \mathscr{I}_{pMp}(X) + \varepsilon t^2 + 4\varepsilon,$$

for any finite generating set X consisting of self-adjoint elements. The factor of 4, which was not present in the counterpart inequality (4.8), arises from the estimate

$$\sum_{i=2}^{m+1} \mathscr{I}_M(v_i + v_i^*;G) + \sum_{i=2}^{m+1} \mathscr{I}_M(i(v_i - v_i^*);G) \leqslant 4\sum_{i=2}^{m+1} \mathscr{I}_M(v_i;G) < 4\varepsilon.$$

This change reflects the replacement of v_i by $v_i + v_i^*$ and $i(v_i - v_i^*)$ in the generating set Y of (4.4). $\square$

LEMMA 4.4. *Let M be a II_1 factor and let $t > 0$. Then M_t is finitely generated if, and only if, M is finitely generated.*

Proof. Without loss of generality, suppose that $0 < t < 1$, otherwise replace t by t^{-1}. Lemma 4.3 shows that if M_t is finitely generated so too is M. Conversely, if M is finitely generated, choose $n \in \mathbb{N}$ with $n^{-1} < t$. Lemma 4.1 shows that $M_{n^{-1}}$ is finitely generated. Let $s = n^{-1}t^{-1}$. Since $0 < s < 1$ and $(M_t)_s = M_{n^{-1}}$ is finitely generated, another application of Lemma 4.3 shows that M_t is finitely generated. $\square$

THEOREM 4.5. (The Scaling Formula) *Let M be a separable II_1 factor. For each $t > 0$,*

$$\mathscr{G}(M_t) = t^{-2}\mathscr{G}(M). \qquad (4.9)$$

Proof. This formula will be established by considering successively the cases $t \in \mathbb{Q}$, $t \in (0,1)$ and $t \in (1,\infty)$.

By Lemma 4.4, M is infinitely generated if, and only if, M_t is infinitely generated. Assume then that both M and M_t are finitely generated. For $n \in \mathbb{N}$ and any separable II_1 factor M, the equation

$$\mathscr{G}(M_{n^{-1}}) = n^2\mathscr{G}(M) \qquad (4.10)$$

is a consequence of the inequalities of Lemma 4.1 and Lemma 4.3. Let $t = p/q$ be a rational and apply (4.10) twice. This gives

$$\mathscr{G}(M_{t^{-1}}) = \mathscr{G}((M_q)_{p^{-1}}) = p^2\mathscr{G}(M_q)$$

and

$$\mathscr{G}(M) = \mathscr{G}((M_q)_{q^{-1}}) = q^2\mathscr{G}(M_q) = \frac{q^2}{p^2}\mathscr{G}(M_{t^{-1}}).$$

This proves the theorem for rational t.

For arbitrary $0 < t < 1$, consider $0 < s < 1$ such that $st \in \mathbb{Q}$. Then

$$\mathscr{G}(M) \leqslant t^2\mathscr{G}(M_t) \leqslant s^2t^2\mathscr{G}(M_{st}) = \mathscr{G}(M) \tag{4.11}$$

by applying Lemma 4.3 twice and the rational case above. Hence $\mathscr{G}(M) = t^2\mathscr{G}(M_t)$.

If $t > 1$, then

$$\mathscr{G}(M) = \mathscr{G}((M_t)_{t^{-1}}) = t^2\mathscr{G}((M_t)_{t^{-1}t}) = t^2\mathscr{G}(M_t).$$

This deals with all possible cases and so completes the proof. □

5. Consequences of scaling

This section contains an initial collection of deductions from the scaling formula. Results regarding the free group factors and estimates involving free products are reserved to Section 7. Note that there is currently no example for which the hypothesis of the first corollary is known to hold.

COROLLARY 5.1. *If there exists a separable* II_1 *factor* M *such that* $\mathscr{G}(M) > 0$*, then there exist separable* II_1 *factors which are not singly generated.*

Proof. If the separable II_1 factor M satisfies $\mathscr{G}(M) = \infty$, then this is already an example with no finite set of generators. Thus the additional assumption that $0 < \mathscr{G}(M) < \infty$ can be made, in which case

$$\lim_{t\to 0+} \mathscr{G}(M_t) = \lim_{t\to 0+} t^{-2}\mathscr{G}(M) = \infty,$$

by Theorem 4.5. It is immediate from the definition of $\mathscr{G}(\cdot)$ that any singly generated factor has $\mathscr{G}(\cdot) \leqslant 1$, and so M_t is not singly generated for sufficiently small values of t. □

COROLLARY 5.2. *Any finitely generated* II_1 *factor* M *with non–trivial fundamental group* $\mathscr{F}(M)$ *has* $\mathscr{G}(M) = 0$ *and is singly generated.*

Proof. If $t \in \mathscr{F}(M) \setminus \{1\}$, then $\mathscr{G}(M) = \mathscr{G}(M_t) = t^{-2}\mathscr{G}(M)$ by Theorem 4.5 as $M \cong M_t$. Since $\mathscr{G}(M) < \infty$, it follows that $\mathscr{G}(M) = 0$ so M is singly generated by Theorem 3.1 or [19]. □

The following lemma leads to the upper bound for $\mathscr{G}(M)$ in Remark 3.2, and which will finally be established in Corollary 5.7. A simple modification yields the odd integer case of Corollary 5.7 without the need for further work.

LEMMA 5.3. *Let* M *be a* II_1 *factor which is generated by* n *hermitian elements for some* $n \in \mathbb{N}$*. Then* $\mathscr{G}_{\mathrm{sa}}(M) \leqslant n - 1$.

Proof. Let $X = \{x_1, \dots, x_n\}$ be n hermitian elements generating N. Let $\varepsilon > 0$ and let A be a masa containing x_n. Since $A \cong L^\infty[0,1]$, there exists some diagonalisable $P \in \mathscr{P}_{\text{eq}}(A)$ with $|P| > \varepsilon^{-1}$. Then $\mathscr{I}(x_n; P) \leqslant \varepsilon$ from item 3 of Remarks 2.4. For each $i \in \{1, \dots, n-1\}$, the estimate $\mathscr{I}(x_i; P) \leqslant 1$ gives

$$\mathscr{G}_{\text{sa}}(M) \leqslant \mathscr{I}(X; P) \leqslant n - 1 + \varepsilon$$

and the result follows. □

The next lemma is the easy direction of Theorem 5.5.

LEMMA 5.4. *Let M be a* II_1 *factor. Then*

$$2\mathscr{G}(M) \leqslant \mathscr{G}_{\text{sa}}(M).$$

Proof. Suppose that M is finitely generated, as otherwise the inequality is vacuous. Let $\varepsilon > 0$. Use Lemma 2.6 to find a self–adjoint set of generators $X = \{x_1, \dots, x_n\}$ for M and a diagonalisable set of projections $P \in \mathscr{P}_{\text{eq}}$ with $\mathscr{I}(X; P) \leqslant \mathscr{G}_{\text{sa}}(M) + \varepsilon$. Take $k = |P|m > n/\varepsilon$ for some $m \in \mathbb{N}$ and choose a diagonalisable refinement $Q = \{e_1, \dots, e_k\} \in \mathscr{P}_{\text{eq}}$ of P. For $l \in \{1, \dots, n\}$ let

$$y_l = \sum_{1 \leqslant i < j \leqslant k} e_i x_l e_j, \quad \text{and} \quad z_l = \sum_{i=1}^{k} e_i x_l e_i.$$

Then let $Y = \{y_1, \dots, y_n\}$ and $Z = \{z_1, \dots, z_n\}$.

Since $x_l = z_l + y_l + y_l^*$ for each l, the finite set $Y \cup Z$ generates M. Moreover,

$$2\mathscr{I}(Y; Q) + \mathscr{I}(Z; Q) = \mathscr{I}(Y; Q) + \mathscr{I}(Y^*; Q) + \mathscr{I}(Z; Q) = \mathscr{I}(X; Q)$$

since the regions in $\{1, \dots, k\} \times \{1, \dots, k\}$ depending on Y, Y^* and Z are disjoint. This gives the estimate

$$\mathscr{I}(Y; Q) \leqslant \mathscr{I}(Y; Q) + \frac{1}{2}\mathscr{I}(Z; Q) \leqslant \frac{1}{2}\mathscr{I}(X; Q).$$

As $z_l \in Q'$ for all l, item 3 of Remarks 2.4 gives $\mathscr{I}(Z; Q) \leqslant n/k < \varepsilon$. Hence

$$\mathscr{G}(M) \leqslant \mathscr{I}(Y; Q) + \mathscr{I}(Z; Q) \leqslant \frac{1}{2}\mathscr{I}(X; Q) + \varepsilon \leqslant \frac{1}{2}\mathscr{G}_{\text{sa}}(M) + \frac{3}{2}\varepsilon$$

as $Q \preceq P$. The lemma follows. □

The generator invariant can now be related to the hermtian generator invariant.

THEOREM 5.5. *Let M be a* II_1 *factor. Then*

$$\mathscr{G}(M) = \frac{1}{2}\mathscr{G}_{\text{sa}}(M).$$

Proof. By Lemma 5.4, it suffices to prove that $\mathscr{G}_{\mathrm{sa}}(M) \leqslant 2\mathscr{G}(M)$ for a finitely generated II_1 factor M. Let $\varepsilon > 0$ and choose $k, n \in \mathbb{N}$ with $k > 1$ such that

$$\mathscr{G}(M) < \frac{n}{2k^2} \leqslant \mathscr{G}(M) + \varepsilon.$$

Write $t = 1/k$ so that $0 < t < 1$. The scaling formula (Theorem 4.5) gives

$$\mathscr{G}(M_t) = t^{-2}\mathscr{G}(M) < \frac{n}{2}.$$

By the generation theorem (Theorem 3.1), M_t is generated by $n+1$ hermitian elements so that $\mathscr{G}_{\mathrm{sa}}(M) \leqslant n$ by Lemma 5.3. Lemma 4.3 on scaling by small t implies that

$$\mathscr{G}_{\mathrm{sa}}(M) \leqslant t^2\mathscr{G}_{\mathrm{sa}}(M_t) \leqslant t^2 n = \frac{n}{k^2} < 2\mathscr{G}(M) + 2\varepsilon.$$

This proves the theorem. □

The scaling for the generator invariant can now be extended to the hermitian case, as stated in Remark 4.2.

COROLLARY 5.6. *If M is a separable II_1 factor and $t > 0$, then*

$$\mathscr{G}_{\mathrm{sa}}(M_t) = t^{-2}\mathscr{G}_{\mathrm{sa}}(M).$$

Proof. This follows directly from Theorems 4.5 and 5.5. □

The next corollary gives the partial converse to the generation theorem which was indicated in Remark 3.2.

COROLLARY 5.7. *If n is the minimal number of hermitian generators of a finitely generated II_1 factor M, then*

$$n - 1 \leqslant 2\mathscr{G}(M) + 1 \leqslant n.$$

Proof. Lemma 5.3 and Theorem 5.5 give $\mathscr{G}(M) \leqslant (n-1)/2$ which yields the second inequality. If $2\mathscr{G}(M) + 1 < n - 1$, then $\mathscr{G}(M) < \frac{n-2}{2}$ so that the generation theorem (Theorem 3.1) gives $n - 1$ hermitian generators for M, contradicting the minimality of n. This gives the first inequality. □

COROLLARY 5.8. *If G is a countable discrete I.C.C. group generated by n elements, then $\mathscr{G}(\mathscr{L}G) \leqslant (n-1)/2$.*

Proof. If G is generated by $g_1, \ldots, g_n$, then there are hermitian elements h_i in the II_1 factor $\mathscr{L}G$ with $W^*(h_i) = W^*(g_i)$ for all i by the spectral theorem, since each g_i is normal. The previous corollary completes the proof. □

Corollary 5.10 gives a formula describing the generator invariant in terms of the minimal number of generators required to generate compressions of the von Neumann algebra, based on the following:

DEFINITION 5.9. Let M be a II_1 factor. Write $\mathscr{G}^{\min}(M)$ for the minimal number of generators of M if M is finitely generated and let $\mathscr{G}^{\min}(M) = \infty$ if M is not finitely generated. The quantity $\mathscr{G}^{\min}_{\mathrm{sa}}(M)$ has a similar definition in terms of the number of self-adjoint generators.

The following simple estimates will be used below.

1. $\mathscr{G}(M) \leqslant \mathscr{G}^{\min}(M)$ and $\mathscr{G}_{\mathrm{sa}}(M) \leqslant \mathscr{G}_{\mathrm{sa}}^{\min}(M)$.
2. $\mathscr{G}_{\mathrm{sa}}^{\min}(M) \leqslant 2\mathscr{G}^{\min}(M) \leqslant \mathscr{G}_{\mathrm{sa}}^{\min}(M) + 1$.

COROLLARY 5.10. *Let M be a separable II_1 factor. Then*

$$\mathscr{G}(M) = \lim_{k\to\infty} \frac{\mathscr{G}^{\min}(M_{1/k})}{k^2} \tag{5.1}$$

and

$$\mathscr{G}_{\mathrm{sa}}(M) = \lim_{k\to\infty} \frac{\mathscr{G}_{\mathrm{sa}}^{\min}(M_{1/k})}{k^2}. \tag{5.2}$$

Proof. By Lemma 4.4, assume that M is finitely generated. The first step is to establish (5.1). The scaling formula (Theorem 4.5) and the estimate 1 above give

$$\mathscr{G}(M) = \frac{\mathscr{G}(M_{1/k})}{k^2} \leqslant \frac{\mathscr{G}^{\min}(M_{1/k})}{k^2}$$

for all k. For $\varepsilon > 0$ find $k_0 \in \mathbb{N}$ with $k_0^2 > \varepsilon^{-1}$. For $k \geqslant k_0$, find $n \in \mathbb{N}$ with

$$2k^2\mathscr{G}(M) < n \leqslant 2k^2\mathscr{G}(M) + 1.$$

The scaling formula gives

$$\mathscr{G}(M_{1/k}) = k^2\mathscr{G}(M) < n/2$$

so the generation theorem (Theorem 3.1) ensures that $M_{1/k}$ is generated by $n+1$ hermitians. Estimate 2 preceding the corollary gives

$$\mathscr{G}^{\min}(M_{1/k}) \leqslant \frac{1}{2}\left(\mathscr{G}_{\mathrm{sa}}^{\min}(M_{1/k}) + 1\right) \leqslant \frac{n}{2} + 1$$

so that

$$\frac{\mathscr{G}^{\min}(M_{1/k})}{k^2} \leqslant \frac{n}{2k^2} + \frac{1}{k^2} \leqslant \mathscr{G}(M) + \frac{2}{k^2} \leqslant \mathscr{G}(M) + 2\varepsilon.$$

This gives equation (5.1). For (5.2), note that the estimate 2 above gives

$$\lim_{k\to\infty} \frac{\mathscr{G}_{\mathrm{sa}}^{\min}(M_{1/k})}{k^2} = 2\lim_{k\to\infty} \frac{\mathscr{G}^{\min}(M_{1/k})}{k^2}.$$

The result follows from (5.1) and Theorem 5.5. □

The last result of this section records that the previous corollary is part of a general phenomenon regarding invariants of II_1 factors which scale and are bounded by the numbers of generators involved. The proof is omitted.

COROLLARY 5.11. *Suppose that $\mathscr{H}$ is an invariant of a II_1 factor which satisfies $\mathscr{H}(M) = t^2\mathscr{H}(M_t)$ for all $t > 0$. If there exist constants $a, b \geqslant 0$ and $\alpha, \beta \in \mathbb{R}$ with*

$$a\mathscr{G}^{\min}(M) + \alpha \leqslant \mathscr{H}(M) \leqslant b\mathscr{G}^{\min}(M) + \beta$$

for all separable II_1 factors M, then

$$a\mathscr{G}(M) \leqslant \mathscr{H}(M) \leqslant b\mathscr{G}(M).$$

6. Finite index subfactors

This brief section examines the generator invariant for finite index inclusions of II_1 factors. Recall from [11] that if $N \subset M$ is a unital inclusion of II_1 factors and e_N is the orthogonal projection from $L^2(M)$ onto $L^2(N)$, then the basic construction $\langle M, e_N \rangle$ is the von Neumann subalgebra of $\mathbb{B}(L^2(M))$ generated by M and e_N and $\langle M, e_N \rangle = JN'J$, where J is the usual modular conjugation operator on $L^2(M)$ given by extending the map $x \mapsto x^*$. This last equation holds for any von Neumann subalgebra of a II_1 factor M, and implies that $\langle M, e_N \rangle$ is a factor precisely when the same is true for N. Recall also that $\{e_N\}' \cap \langle M, e_N \rangle = N$. In this situation, N is said to be a *finite index* subfactor if $\langle M, e_N \rangle$ is a type II_1 factor, and one formulation of the index $[M : N]$ is given by $[M : N] = \mathrm{Tr}(1)$, where Tr is the unique trace on $\langle M, e_N \rangle$ normalised with $\mathrm{Tr}(e_N) = 1$.

LEMMA 6.1. *Suppose that $N \subset M$ is a unital inclusion of* II_1 *factors with $[M : N] < \infty$. Then*

$$\mathscr{G}(\langle M, e_N \rangle) \leqslant \mathscr{G}(M).$$

Note that in this lemma, $\mathscr{G}(\langle M, e_N \rangle)$ is computed with respect to the trace $\lambda^{-1}\mathrm{Tr}$, where $\lambda = [M : N]$, since this is the trace on the basic construction algebra $\langle M, e_N \rangle$, normalised to take the value 1 at the identity.

Proof. Take $\varepsilon > 0$, a finite generating set X for M and (by Lemma 2.6) a collection of projections $P \in \mathscr{P}_{\mathrm{eq}}(M)$ with $\mathscr{I}(X; P) < \mathscr{G}(M) + \varepsilon$. Since N is also a factor, there exists a unitary $u \in \mathscr{U}(N)$ such that $uPu^* = Q_0 \in \mathscr{P}(N)$. Choose a refinement Q of Q_0 in $\mathscr{P}_{\mathrm{eq}}(N)$ with $|Q| = k$ for some $k > \varepsilon^{-1}$. The inequality

$$\mathscr{I}(e_N; Q) \leqslant k^{-1} < \varepsilon$$

is implied by item 3 of Remarks 2.4. Since uXu^* also generates M, it follows that

$$\mathscr{G}(\langle M, e_N \rangle) \leqslant \mathscr{I}(uXu^* \cup \{e_N\}; Q) \leqslant \mathscr{I}(X; P) + \varepsilon < \mathscr{G}(M) + 2\varepsilon,$$

proving the result. □

LEMMA 6.2. *Let $N \subset M$ be a finite index unital inclusion of* II_1 *factors. Then*

$$\mathscr{G}(N) \geqslant \mathscr{G}(M).$$

Proof. Recall from [11, Lemma 3.1.8], that given a finite index inclusion $N \subset M$, there exists a subfactor $P \subset N$ with $[N : P] = [M : N]$ and $\langle N, e_P \rangle \cong M$. The result follows immediately from the previous lemma. □

Note that if there is a II_1 factor M with $\mathscr{G}(M) > 0$, then the previous lemma does not hold for infinite index subfactors as there is always a copy of the hyperfinite II_1 factor R inside any II_1 factor.

THEOREM 6.3. *Let $N \subset M$ be a finite index unital inclusion of* II_1 *factors. Then $\mathscr{G}(M) = 0$ if, and only if, $\mathscr{G}(N) = 0$.*

Proof. Suppose that $\mathscr{G}(M) = 0$. By Lemma 6.1, it follows that $\mathscr{G}(\langle M, e_N \rangle) = 0$. Now $N \cong e_N \langle M, e_N \rangle e_N$, so $N \cong \langle M, e_N \rangle_{\lambda^{-1}}$, where $\lambda = [M : N]$. The scaling formula (Theorem 4.5) immediately gives $\mathscr{G}(N) = 0$. The reverse direction is Lemma 6.2 above. □

More generally, suppose that $N \subset M$ is a finite index unital inclusion of II_1 factors and write $\lambda = [M : N]$. The isomorphism $N \cong e_N \langle M, e_N \rangle e_N$ and the scaling formula lead to the equality

$$\mathscr{G}(N) = \lambda^2 \mathscr{G}(\langle M, e_N \rangle).$$

Furthermore, Lemmas 6.1 and 6.2 give

$$\mathscr{G}(N) \geqslant \mathscr{G}(M) \geqslant \mathscr{G}(\langle M, e_N \rangle) = \lambda^{-2} \mathscr{G}(N).$$

It is then natural to pose the following question.

QUESTION 6.4. Suppose that $N \subset M$ is a finite index unital inclusion of II_1 factors and write $\lambda = [M : N]$. Is it the case that

$$\mathscr{G}(N) = \lambda \mathscr{G}(M)?$$

7. Free group factors and free products

The main result in this section is Theorem 7.1 below, which is obtained by using the quadratic scaling of the generator invariant and a similar property of the interpolated free group factors $\mathscr{L}\mathbb{F}_s$ of the first author and Rădulescu [4, 18]. For $r > 1$ and $\lambda > 0$, the quadratic scaling formula for the interpolated free group factors states that

$$(\mathscr{L}\mathbb{F}_r)_\lambda = \mathscr{L}\mathbb{F}_{1+\frac{r-1}{\lambda^2}}, \tag{7.1}$$

from [4, Theorem 2.4].

THEOREM 7.1. *There exists a constant* $0 \leqslant \alpha \leqslant 1/2$ *such that*

$$\mathscr{G}(\mathscr{L}\mathbb{F}_{1+r}) = r\alpha. \tag{7.2}$$

for all $r \in (0, \infty]$.

Proof. For $r \in (0, \infty)$ write $\beta(r) = \mathscr{G}(\mathscr{L}\mathbb{F}_{r+1})$. Equation (7.1) above and Theorem 4.5 combine to give

$$\beta\left(\frac{r-1}{\lambda^2}\right) = \mathscr{G}((\mathscr{L}\mathbb{F}_r)_\lambda) = \lambda^{-2}\mathscr{G}(\mathscr{L}\mathbb{F}_r) = \lambda^{-2}\beta(r-1). \tag{7.3}$$

Take $r - 1 = t$ and $\lambda^{-2} = s$ in (7.3) to obtain

$$\beta(st) = s\beta(t)$$

for all $s, t \in (0, \infty)$. Hence, there is a constant $\alpha = \beta(1) \geqslant 0$ with $\beta(t) = \alpha t$. The estimate $\alpha \leqslant \frac{1}{2}$ follows from Corollary 5.8 as $\alpha = \mathscr{G}(\mathscr{L}\mathbb{F}_2)$ and $\mathbb{F}_2$ is certainly generated by two elements.

It remains to extend (7.2) to the $r = \infty$ case. Since the fundamental group of $\mathscr{L}\mathbb{F}_\infty$ is $\mathbb{R}^+$, [17], or more easily since $\mathbb{N} \subset \mathscr{F}(\mathscr{L}\mathbb{F}_\infty)$, [25, Corollary 5.2.3], it follows that $\mathscr{G}(\mathscr{L}\mathbb{F}_\infty)$ is either 0 or ∞, by Corollary 5.2. Suppose that $\alpha = 0$ so that $\mathscr{G}(\mathscr{L}\mathbb{F}_k) = 0$ for each $k \geqslant 2$. Consideration of the chain

$$\mathscr{L}\mathbb{F}_2 \subset \mathscr{L}\mathbb{F}_3 \subset \mathscr{L}\mathbb{F}_4 \subset \cdots \subset \mathscr{L}\mathbb{F}_\infty$$

gives $\mathscr{G}(\mathscr{L}\mathbb{F}_\infty) = 0$ by Shen's main technical theorem, [19, Theorem 5.1]. Conversely, if $\mathscr{G}(\mathscr{L}\mathbb{F}_\infty) = 0$, then the relation $\alpha = \mathscr{G}(\mathscr{L}\mathbb{F}_2) = 0$ follows from the isomorphism

$$\mathbb{F}_2 \cong \mathbb{F}_\infty \rtimes \mathbb{Z}.$$

This is essentially in [19] and a proof can also be found in [20, Chapter 15]. Hence $\mathscr{G}(\mathscr{L}\mathbb{F}_\infty) = 0$ if and only if $\alpha = 0$ and the result follows. □

As an immediate consequence, there is a direct link between the free group isomorphism problem and the generator invariant.

COROLLARY 7.2. *If $\mathscr{G}(\mathscr{L}\mathbb{F}_2) > 0$, then all the interpolated free group factors $\mathscr{L}\mathbb{F}_t$ are pairwise non-isomorphic for $t > 1$.*

REMARK 7.3. Since $\mathscr{L}\mathbb{F}_1 \cong L^\infty[0,1]$ has generator invariant 0, the previous theorem extends to include the case $r = 0$.

The next results give some estimates regarding the generator invariant and free products. Reverse inequalities to either Theorem 7.5 or Theorem 7.6 would immediately combine with Corollary 7.2 to show the non–isomorphism of the free group factors. Lemma 7.4 originates in a conjugacy idea used repeatedly by Shen to obtain [19, Theorem 5.1].

LEMMA 7.4. *Let M and N be II_1 factors containing a common diffuse von Neumann subalgebra B. Then*

$$\mathscr{G}(M *_B N) \leqslant \mathscr{G}(M) + \mathscr{G}(N).$$

Proof. Given $\varepsilon > 0$, choose finite subsets $X \subset M$ and $Y \subset N$ and families $P \in \mathscr{P}(M)$ and $Q \in \mathscr{P}(N)$ with $W^*(X) = M$, $W^*(Y) = N$ and

$$\mathscr{I}(X;P) < \mathscr{G}(M) + \frac{\varepsilon}{2}, \quad \mathscr{I}(Y;Q) < \mathscr{G}(N) + \frac{\varepsilon}{2}. \tag{7.4}$$

By refining if necessary, it may be assumed that P and Q are equivalent since Proposition 2.5 ensures that the estimate (7.4) is unaffected by refinement. Since B is diffuse, choose $E \in \mathscr{P}_B$ equivalent to P and Q and unitaries $u \in M$ and $v \in N$ with $uPu^* = E$ and $vQv^* = E$. Then $uXu^* \cup vYv^*$ is a finite generating set for $M *_B N$ and

$$\mathscr{I}(uXu^* \cup vYv^*;E) = \mathscr{I}(X;P) + \mathscr{I}(Y;Q) < \mathscr{G}(M) + \mathscr{G}(N) + \varepsilon,$$

from which the result follows. □

THEOREM 7.5. *Let M and N be* II_1 *factors. Then*

$$\mathscr{G}(M * N) \leqslant \mathscr{G}(M) + \mathscr{G}(N) + \frac{1}{2}.$$

Proof. Choose masas $A \subset M$ and $B \subset N$. Then $M \cong M *_A A$ and $N \cong N *_B B$, and so

$$M * N \cong (M *_A A) * (B *_B N) \cong (M *_A \mathscr{L}\mathbb{F}_2) *_B N,$$

where $\mathscr{L}\mathbb{F}_2 = A * B$. Now use Lemma 7.4 twice to obtain

$$\mathscr{G}(M * N) \leqslant \mathscr{G}(M *_A \mathscr{L}\mathbb{F}_2) + \mathscr{G}(N) \leqslant \mathscr{G}(M) + \mathscr{G}(N) + \mathscr{G}(\mathscr{L}\mathbb{F}_2).$$

The result follows as Theorem 7.1 gives $\mathscr{G}(\mathscr{L}\mathbb{F}_2) \leqslant 1/2$. □

The remainder of this section examines free products with finite hyperfinite von Neumann algebras. In [13], Jung proves that for a fixed hyperfinite von Neumann algebra Q with a fixed faithful normal tracial state ϕ, Voiculescu's modified free entropy dimension $\delta_0(X)$ is the same for all finite sets X that generate Q. Write $\delta_0(Q)$ for this quantity. The definition of the modified free entropy dimension will be given in the next section, which discusses free entropy dimension in conjunction with the generator invariant. Here only the value of $\delta_0(Q)$ is needed. Following [13], given (Q, ϕ), decompose Q over its centre to obtain

$$Q \cong Q_0 \oplus \left(\bigoplus_{i=1}^{s} \mathbb{M}_{k_i}(\mathbb{C}) \right),$$

where Q_0 is diffuse or $\{0\}$, the sum on the right is either empty, finite or countably infinite and each $k_i \in \mathbb{N}$. The trace ϕ is given by

$$\phi = \alpha_0 \phi_0 \oplus \left(\bigoplus_{i=1}^{s} \alpha_i \mathrm{tr}_{k_i} \right),$$

where

- $\alpha_0 > 0$ and ϕ_0 is a faithful normal trace on Q_0, if $Q_0 \neq \{0\}$;
- $\alpha_0 = 0$ and $\phi_0 = 0$ if $Q_0 = \{0\}$;
- tr_{k_i} is the tracial state on the $k_i \times k_i$ matrices $\mathbb{M}_{k_i}(\mathbb{C})$ and each $\alpha_i > 0$.

Then, from [13],

$$\delta_0(Q) = 1 - \sum_{i=1}^{s} \frac{\alpha_i^2}{k_i^2}. \tag{7.5}$$

Furthermore, as also noted in [13], this quantity agrees with the 'free dimension number' for Q defined in earlier work of the first author [3]. In this work it was shown (Theorem 4.6 of [3]) that if $A = L^\infty[0,1]$ is equipped with the usual trace $\int_0^1 \cdot \, \mathrm{d}t$, then $A * Q \cong \mathscr{L}\mathbb{F}_r$, where $r = \delta_0(Q) + 1$.

THEOREM 7.6. *Let M be a II_1 factor and Q a hyperfinite von Neumann algebra with a fixed faithful normalised trace ϕ. Then*

$$\mathscr{G}(M * Q) \leqslant \mathscr{G}(M) + \frac{1}{2}\delta_0(Q).$$

Proof. Choose a masa $A \subset M$ so that A is isomorphic to $L^\infty[0,1]$ with the trace on A (coming from τ_M) being given by $\int_0^1 \cdot\, \mathrm{d}t$. The discussion preceding the theorem gives

$$M * Q \cong M *_A *(A * Q) \cong M *_A \mathscr{L}\mathbb{F}_r$$

where $r = 1 + \delta_0(Q)$ and A is a masa in $\mathscr{L}\mathbb{F}_r$. By Proposition 7.4 and Theorem 7.1,

$$\mathscr{G}(M * Q) \leqslant \mathscr{G}(M) + \mathscr{G}(\mathscr{L}\mathbb{F}_r) \leqslant \mathscr{G}(M) + \frac{r-1}{2} = \mathscr{G}(M) + \frac{1}{2}\delta_0(Q),$$

exactly as required. □

The next two corollaries are obtained by taking Q to be successively the $n \times n$ matrices with the usual normalised trace and to be $\mathscr{L}\mathbb{Z}_n$, with the group trace. The results follow from calculating $\delta_0(\mathbb{M}_n(\mathbb{C})) = 1 - \frac{1}{n^2}$ and $\delta_0(\mathscr{L}\mathbb{Z}_n) = 1 - \frac{1}{n}$ from Jung's formula (7.5).

COROLLARY 7.7. *Let M be a II_1 factor and $n \geqslant 2$. Then*

$$\mathscr{G}(M * \mathbb{M}_n(\mathbb{C})) \leqslant \mathscr{G}(M) + \frac{1}{2} - \frac{1}{2n^2}.$$

COROLLARY 7.8. *Let M be a II_1 factor and $n \geqslant 2$. Then*

$$\mathscr{G}(M * \mathscr{L}\mathbb{Z}_n) \leqslant \mathscr{G}(M) + \frac{1}{2} - \frac{1}{2n}.$$

8. Free entropy dimension

The objective in this section is to relate the generator invariant to Voiculescu's modified free entropy dimension by proving the inequalities

$$\delta_0(X) \leqslant 1 + 2\mathscr{I}(X), \tag{8.1}$$

when X is a finite generating set in a finite von Neumann algebra M and, under the extra assumption that X consists of self–adjoint elements,

$$\delta_0(X) \leqslant 1 + \mathscr{I}(X). \tag{8.2}$$

For certain sets X of operators, these inequalities give lower bounds on $\mathscr{I}(X)$. These seem to be the only such lower bounds that are currently known. Consider the case of a DT-operator Z, introduced by the first author and Haagerup in [5]. By [2] each such Z, including the quasi–nilpotent DT–operator T, generates $\mathscr{L}\mathbb{F}_2$. The operator Z is constructed by realising $\mathscr{L}\mathbb{F}_2$ as generated by a semicircular element S together with a free copy of $L^\infty[0,1]$, using projections from $L^\infty[0,1]$ to cut out an "upper triangular" part of S, which is the quasi–nilpotent DT–operator T, and then adding an operator from $L^\infty[0,1]$ to get Z. Using projections from $L^\infty[0,1]$, one easily sees that $\mathscr{I}(Z) \leqslant 1/2$. On the other hand, since, by [6], $\delta_0(Z) = 2$, (8.1) gives

$$\mathscr{I}(Z) = 1/2.$$

Similarly, free semi–circular elements h_1, h_2 generating $\mathscr{L}\mathbb{F}_2$ satisfy $\delta_0(h_1,h_2) = 2$ from [23, 24], so that (8.2) gives $\mathscr{I}(h_1,h_2) \geqslant 1$. The reverse inequality follows by taking a sufficiently fine family of projections in either $W^*(h_1)$ or $W^*(h_2)$ — see the proof of Lemma 5.3 — so that

$$\mathscr{I}(h_1,h_2) = 1.$$

The central question in the theory of microstates free entropy dimension is that of invariance: if one has two finite generating sets X_1 and X_2 for the same II_1 factor must $\delta_0(X_1) = \delta_0(X_2)$? Certain conditions on the factor, such as having a Cartan masa ([24]) and non-primeness ([7]), imply that all finite generating sets have free-entropy at most one. Independently Jung [14] and Hadwin and Shen [10] have shown that the existance of a finite generating set X for a II_1 factor with $\delta_0(X) = 1$ and certain additional technical conditions, which are subtly different in each approach, imply that all finite generating sets Y for M have $\delta_0(Y) \leqslant 1$. Before discussing these conditions further and establishing inequalities (8.1) and (8.2), the prevailing notation in this area is outlined. The definition of the modified free-entropy dimension given below is Jung's covering–number reformulation from [12, 15].

Let M be a II_1 factor and continue to write τ for the faithful normal trace on M normalised by $\tau(1) = 1$. For $k \in \mathbb{N}$ write $\mathbb{M}_k(\mathbb{C})$ for the $k\times k$ matrices, equipped with the trace tr_k normalised with $\mathrm{tr}_k(1) = 1$. For a finite subset $X = \{x_1,\dots,x_n\} \subset M$, $\gamma > 0$, $k,m \in \mathbb{N}$ and $R > 0$ define the microstate space $\Gamma_R(X;m,k,\gamma)$ to be the set of all n-tuples $(a_1,\dots,a_n)$ of $k\times k$-matrices whose *-moments approximate those of $(x_1,\dots,x_n)$ up to order m within a tolerance of γ and whose norms are bounded by R. This means that $\|a_i\| \leqslant R$ for all i and

$$\left|\tau(x_{i_1}^{j_1}\dots x_{i_p}^{j_p}) - \mathrm{tr}_k(a_{i_1}^{j_1}\dots a_{i_p}^{j_p})\right| < \gamma,$$

for all $p \leqslant m$, $i_1,\dots,i_p \in \{1,\dots,n\}$ and $j_1,\dots,j_m \in \{1,*\}$. When all the x_j's are self–adjoint, it makes no difference to the definition of $\delta_0(X)$ whether all the a_i's are required to be self–adjoint — see for example the beginning of Section 3 of [6].

Given $\varepsilon > 0$, the covering number $K_\varepsilon(Y)$ of a metric space Y is the minimal cardinality of an ε-net for Y. One easy estimate, used in the sequel, is

$$K_\varepsilon(Y) \leqslant P_{\varepsilon/2}(Y), \tag{8.3}$$

where $P_{\varepsilon/2}(Y)$ is the maximal number of disjoint open $\varepsilon/2$-balls which can be found in Y. In [12], Jung defines, for $m \in \mathbb{N}$, $\gamma,\varepsilon > 0$ and $R > 0$,

$$\mathbb{K}_{\varepsilon,R}(X;m,\gamma) = \limsup_{k\to\infty} k^{-2}\log K_\varepsilon\left(\Gamma_R(X;m,k,\gamma)\right). \tag{8.4}$$

where the metric on $\Gamma_R(X;m,k,\gamma)$ is that obtained from the Euclidian norm on $(\mathbb{M}_k(\mathbb{C}))^n$ given by

$$\|(a_1,\dots,a_n)\|_2^2=\left(\sum_{l=1}^n \mathrm{tr}_k(a_l^*a_l)\right)^{1/2}. \tag{8.5}$$

Then define

$$\mathbb{K}_{\varepsilon,R}(X)=\inf_{m\in\mathbb{N},\gamma>0}\mathbb{K}_{\varepsilon,R}(X;m,\gamma). \tag{8.6}$$

and

$$\mathbb{K}_\varepsilon(X)=\sup_{R>0}K_{\varepsilon,R}(X).$$

Jung shows in Corollary 24 of [12], that the modified free entropy dimension $\delta_0(X)$ is given by

$$\delta_0(X)=\limsup_{\varepsilon\to 0}\frac{\mathbb{K}_\varepsilon(X)}{|\log\varepsilon|}. \tag{8.7}$$

In [14], Jung made a detailed analysis of the rate of convergence in the limit-superior in (8.7). For $\alpha>0$, say that the generating set X of M to be α*-bounded* if, for some $R\geqslant\max_{x\in X}\|x\|$, there exists some constant K and $\varepsilon_0>0$ such that

$$\mathbb{K}_{\varepsilon,R}(X)\leqslant\alpha|\log\varepsilon|+K,$$

for all $0<\varepsilon<\varepsilon_0$. Corollary 1.4 of [14] demonstrates that this is equivalent to the original definition of α-boundedness in [14] and that if X is α-bounded, then $\delta_0(X)\leqslant\alpha$. Furthermore, M is said to be *strongly–*1*–bounded* if it has a generating set X that is 1-bounded and if there exists a self-adjoint element x belonging to the *-algebra generated by X that has finite free entropy. Theorem 3.2 of [14] shows that if M is strongly–1–bounded, then every finite set of generators for M is 1–bounded.

The alternative approach of Hadwin and Shen in [10] is to examine coverings by unitary orbits of balls. For m,k,γ,R as above and $\varepsilon>0$, define

$$\nu_\varepsilon(\Gamma_R(x_1,\dots,x_n;m,k,\gamma))$$

to be the minimum cardinality of a set $\Lambda\subset\Gamma_R(x_1,\dots,x_n;m,k,\gamma)$ such that for every $a=(a_1,\dots,a_n)\in\Gamma_R(x_1,\dots,x_n;m,k,\gamma)$ there is some unitary $u\in\mathbb{M}_k(\mathbb{C})$ and $b\in\Lambda$ with $\|uau^*-b\|_2\leqslant\varepsilon$, i.e. the microstate space $\Gamma_R(x_1,\dots,x_n;m,k,\gamma)$ is covered by $\nu_\varepsilon(\Gamma_R(x_1,\dots,x_n;m,k,\gamma))$ orbit ε-balls. Here uau^* is defined to be the n-tuple $(ua_1u^*,\dots,ua_nu^*)$. Then define

$$\mathscr{K}(x_1,\dots,x_n;R,\varepsilon)=\inf_{m\in\mathbb{N},\gamma>0}\limsup_{k\to\infty}\frac{\log(\nu_\varepsilon(\Gamma_R(x_1,\dots,x_n;m,k,\gamma)))}{-k^2\log\varepsilon}$$

and the *upper free-orbit dimension* $\mathscr{K}_2(x_1,\dots,x_n)$ by

$$\mathscr{K}_2(x_1,\dots,x_n)=\sup_{0<\varepsilon<1}\sup_{R>0}\mathscr{K}(x_1,\dots,x_n;R,\varepsilon).$$

The main theorem of [10] (Section 3, Theorem 1) is that if X is a finite set of generators for a II_1 factor with $\mathscr{K}_2(X)=0$, then every finite set of generators Y for this factor also has $\mathscr{K}_2(Y)=0$. Therefore, it makes sense to say a II_1 factor has zero upper free orbit-dimension when it has a finite generating set X with $\mathscr{K}_2(X)=0$.

The connection between the upper free orbit-dimension and Voiculescu's modified free entropy dimension is immediate from results of Szarek [21] which show that there is an absolute constant $C>0$ such that the groups $\mathscr{U}_k$ of unitary matrices in $\mathbb{M}_k(\mathbb{C})$ equipped with the metric induced from the operator norm satisfy

$$K_\varepsilon(\mathscr{U}(K)) \leqslant \left(\frac{C}{\varepsilon}\right)^{k^2}.$$

Given a finite generating set $X=(x_1,\ldots,x_n)$ for a II_1 factor M, take $S=2(\sum_{i=1}^n \|x_i\|_2^2)^{1/2}$. Provided $m\geqslant 2$ and γ is sufficiently small, the estimate

$$\begin{aligned} K_\varepsilon(\Gamma_R(X;m,k,\gamma)) &\leqslant \nu_{\varepsilon/2}(\Gamma_R(X;m,k,\gamma))\cdot K_{\varepsilon/8S}(\mathscr{U}_k)\\ &\leqslant \left(\frac{8CS}{\varepsilon}\right)^{k^2} \nu_{\varepsilon/2}(\Gamma_R(X;m,k,\gamma)), \end{aligned}$$

follows. Therefore

$$\mathbb{K}_{\varepsilon,R}(X;m,\gamma) \leqslant \log(8CS/\varepsilon) + \limsup_{k\to\infty} \log \nu_{\varepsilon/2}(\Gamma_R(X;m,k,\gamma))$$

and

$$\begin{aligned} \mathbb{K}_{\varepsilon,R}(X) &\leqslant \log(8CS/\varepsilon) + \mathscr{K}(X;R,\varepsilon/2)|\log(\varepsilon/2)|\\ &\leqslant (1+\mathscr{K}(X;R,\varepsilon/2))\,|\log\varepsilon| + \mathscr{K}(X;R,\varepsilon/2)\log(2) + \log(8CS) \end{aligned} \tag{8.8}$$

The free orbit-dimension $\mathscr{K}_1(X)$ of [10] is defined by

$$\mathscr{K}_1(X) = \limsup_{\varepsilon\to 0}\sup_{R>0} \mathscr{K}(X;R,\varepsilon).$$

The relationships between the approaches of [14] and [10] follow from the definition of $\mathscr{K}_1$ and (8.8).

PROPOSITION 8.1. *Let X be a finite set of generators for a II_1 factor M. Then $\delta_0(X)\leqslant 1+\mathscr{K}_1(X)$ and X is α-bounded, for all $\alpha>1+\mathscr{K}_1(X)$.*

PROPOSITION 8.2. *A II_1 factor whose upper free orbit dimension is zero, i.e. has a finite generating set X with $\mathscr{K}_2(X)=0$, is strongly 1-bounded.*

Proof. By hypothesis there is a set of self-adjoint generators X for the II_1 factor with $\mathscr{K}_2(X)=0$. If there is no element of finite entropy in the *-algebra generated by X, adjoin one. This new set of generators Y also has zero upper free orbit-dimension by [10, Section 3, Theorem 1]. The definition of $\mathscr{K}_2$ ensures that

$$\mathbb{K}_{\varepsilon,R}(Y) \leqslant |\log\varepsilon| + \log(8CS),$$

from equation (8.8). That is Y is 1-bounded, and so M is strongly 1-bounded. $\square$

In this section the upper free-orbit dimension is used as it enables us to show, in Corollary 8.7, that no set of generators X for an interpolated free group factor can have $\mathscr{I}(X)=0$. The strongly-1-bounded approach to this would only rule out the existance of such generators X which also have an element of finite entopy in the $*$-algebra they generate. Before proceeding to the main result of this section, recall the following lemma, due to Voiculescu. Formally, a proof can be constructed by copying the ideas of the proof of Proposition 1.6 of [24].

LEMMA 8.3. *Suppose that $X=\{x_1,\dots,x_n\}$ is a finite subset of M with $W^*(X)=M$. Fix $p\in\mathbb{N}$ and pairwise orthogonal projections $e_1,e_2,\dots,e_p$ in M with sum 1 and $\tau(e_i)=p^{-1}$ for each i. Given $\gamma>0$ and $m\in\mathbb{N}$, there exists $\gamma'>0$, $m',k'\in\mathbb{N}$ such that if*

$$(a_1,\dots,a_n)\in\Gamma(X;m',k,\gamma'),$$

for some $k\geqslant k'$ with $p|k$, then there exist pairwise orthogonal projections $(f_1,\dots,f_p)$ in $\mathbb{M}_k(\mathbb{C})$ each with $\mathrm{tr}_k(f_i)=p^{-1}$ (so that $\sum_{i=1}^n f_i=1$) satisfying

$$(a_1,\dots,a_n,f_1,\dots,f_p)\in\Gamma(x_1,\dots,x_n,e_1,\dots,e_p;m,k,\gamma).$$

Now the main result of this section relating $\mathscr{I}(X)$ to $\mathscr{K}_1(X)$ and $\mathscr{K}_2(X)$.

PROPOSITION 8.4. *Let M be a finite von Neumann algebra and $X\subset M$ a finite generating set for M. In general, $\mathscr{K}_1(X)\leqslant 2\mathscr{I}(X)$ and in the case that each $x\in X$ is self-adjoint, $\mathscr{K}_1(X)\leqslant\mathscr{I}(X)$. Furthermore, if $\mathscr{I}(X)=0$, then $\mathscr{K}_2(X)=0$.*

Proof. Let $X=(x_1,\dots,x_n)$ be an n-tuple which generates M. Fix $\varepsilon>0$ and suppose that $c>\mathscr{I}(X)$. Use Lemma 2.6 to find $E=\{e_1,\dots,e_p\}\in\mathscr{P}_{\mathrm{eq}}(M)$ with $\mathscr{I}(X;E)<c$ and each $\tau(e_i)=p^{-1}$. Let $R>0$ and write $S=(\sum_{i=1}^n\|x_i\|_2^2)^{1/2}$. Take $\gamma>0$ and $m\in\mathbb{N}$ with $m\geqslant 6$. Let k',m' and γ' be the constants obtained by applying Lemma 8.3 to X and $(e_1,\dots,e_p)$. Let $k\geqslant k'$ be divisible by p and fix a family of pairwise orthogonal projections $(q_1,\dots,q_p)$ in $\mathbb{M}_k(\mathbb{C})$ with each $\tau_k(q_i)=p^{-1}$.

For each $l=1,\dots,n$ write T_l for the set of pairs (i,j) with $e_ix_le_j\neq 0$. As $\mathscr{I}(I;E)<c$, it follows that

$$\sum_{l=1}^n |T_l|<cp^2. \tag{8.9}$$

Following the approach of [7], define the projection Q from $(\mathbb{M}_k(\mathbb{C}))^n$ into $(\mathbb{M}_k(\mathbb{C}))^n$ by

$$Q(a_1,\dots,a_n)=\left(\sum_{(i,j)\in T_l} q_ia_lq_j\right)_{l=1}^n.$$

When each $x_l=x_l^*$, the sets T_l are invariant under the adjoint operation so in this case Q restricts to give a projection from $(\mathbb{M}_k^{\mathrm{sa}}(\mathbb{C}))^n$ into $(\mathbb{M}_k^{\mathrm{sa}}(\mathbb{C}))^n$. The range $Q((\mathbb{M}_k(\mathbb{C}))^n)$ is a $2\sum_{l=1}^n|T_l|(k/p)^2$–dimensional subspace of $(\mathbb{M}_k(\mathbb{C}))^n\cong\mathbb{R}^{2nk^2}$. Under the additional assumption that $x_l=x_l^*$ for each l, the range $Q((\mathbb{M}_k^{\mathrm{sa}}(\mathbb{C}))^n)$ is a $\sum_{l=1}^n|T_l|(k/p)^2$–dimensional subspace of $(\mathbb{M}_k^{\mathrm{sa}}(\mathbb{C}))^n\cong\mathbb{R}^{nk^2}$.

Let Y be the subset $(Q(\mathbb{M}_k(\mathbb{C})^n))_{2S}$ consisting of all elements $(b_1,\dots,b_n) \in Q(\mathbb{M}_k(\mathbb{C})^n)$ with $\|(b_1,\dots,b_n)\|_2 \leqslant 2S$ if X does not consist of self-adjoint elements, and let $Y = (Q((\mathbb{M}_k^{\mathrm{sa}}(\mathbb{C}))^n))_{2S}$ when each $x_l = x_l^*$. Volume considerations give

$$P_{\varepsilon/2}(Y) \leqslant \left(\frac{2S}{\varepsilon/2}\right)^{2\sum_{l=1}^n |T_l|(k/p)^2},$$

in the first case and

$$P_{\varepsilon/2}(Y) \leqslant \left(\frac{2S}{\varepsilon/2}\right)^{\sum_{l=1}^n |T_l|(k/p)^2},$$

in the second. The simple estimate (8.3) combines with (8.9) to give

$$K_\varepsilon(Y) \leqslant \begin{cases} \left(\frac{4S}{\varepsilon}\right)^{ck^2}, & x_l = x_l^* \text{ for all } l \\ \left(\frac{4S}{\varepsilon}\right)^{2ck^2}, & \text{otherwise.} \end{cases}$$

Given any point $a = (a_1,\dots,a_n) \in \Gamma_R(X;m',k,\gamma)$, let $(f_1,\dots,f_p)$ be the orthogonal projections in $\mathbb{M}_k(\mathbb{C})$ with each $\tau_k(f_i) = p^{-1}$ from Lemma 8.3. Now each x_l can be written $x_l = \sum_{(i,j)\in T_l} e_i x_l e_j$. Since

$$(a_1,\dots,a_n,f_1,\dots,f_p) \in \Gamma_R(x_1,\dots,x_n,e_1,\dots,e_p;m,k,\gamma),$$

it follows that

$$\left\| a_l - \sum_{(i,j)\in T_l} f_i a_l f_j \right\|_2^2 < \gamma,$$

from the standing assumption that $m \geqslant 6$. Take a unitary $u \in \mathbb{M}_k(\mathbb{C})$ such that $uf_iu^* = q_i$ for each $i = 1,\dots,p$. Then $u^*Q(uau^*)u$ is the n-tuple with $\sum_{(i,j)\in T_l} f_i a_l f_j$ in the l-entry. Thus

$$\|a - u^*Q(uau^*)u\|_2^2 \leqslant n\gamma.$$

Provided γ is taken small enough, $\|a\|_2 \leqslant 2S$, so that $Q(uau^*)$ lies in the ball Y. If in addition, $(n\gamma)^{1/2} \leqslant \varepsilon/2$, it follows at most $K_{\varepsilon/2}(Y)$ orbit ε-balls are required to cover $\Gamma_R(x_1,\dots,x_n;m,k,\gamma)$. That is

$$\nu_\varepsilon(\Gamma_R(x_1,\dots,x_n;m,k,\gamma) \leqslant \begin{cases} \left(\frac{8S}{\varepsilon}\right)^{ck^2}, & x_l = x_l^* \text{ for all } l; \\ \left(\frac{8S}{\varepsilon}\right)^{2ck^2}, & \text{otherwise,} \end{cases}$$

for $R > 0$, $m \geqslant 6$, γ sufficiently small and $k \geqslant k'$ which is divisible by p. Thus

$$\begin{aligned} \mathscr{K}(x_1,\dots,x_n;R,\varepsilon) &= \inf_{m\in\mathbb{N},\gamma>0} \limsup_{k\to\infty} \frac{\log \nu_\varepsilon(\Gamma_R(x_1,\dots,x_n;m,k,\gamma))}{-k^2\log\varepsilon} \\ &\leqslant \begin{cases} c\left(1 + \frac{\log(8S)}{|\log\varepsilon|}\right), & x_l = x_l^*, \text{for all } l, \\ 2c\left(1 + \frac{\log(8S)}{|\log\varepsilon|}\right), & \text{otherwise.} \end{cases} \end{aligned}$$

This holds for each $c > \mathscr{I}(X)$ and so we can replace c by $\mathscr{I}(X)$ in the inequality above to obtain

$$\begin{aligned}\mathscr{K}_1(x_1,\ldots,x_n) &= \limsup_{\varepsilon\to 0}\sup_{R>0}\mathscr{K}(x_1,\ldots,x_n;R,\varepsilon)\\ &\leqslant \begin{cases}\mathscr{I}(X), & x_l = x_l^*, \text{ for all } l,\\ 2\mathscr{I}(X), & \text{otherwse.}\end{cases}\end{aligned}$$

in general, and if $\mathscr{I}(X) = 0$, then $\mathscr{K}_2(X) = 0$. □

The first two corollaries below follow immediately.

COROLLARY 8.5. *Let X be a finite set of generators for the II_1 factor (M,τ). Then*

$$\delta_0(X) \leqslant 1 + 2\mathscr{I}(X).$$

COROLLARY 8.6. *Let X be a finite set of self-adjoint generators for the II_1 factor (M,τ). Then*

$$\delta_0(X) \leqslant 1 + \mathscr{I}(X).$$

The next corollary follows from Proposition 8.4 and [10, Section 3, Theorem 1]. It shows in particular that $\mathscr{I}(X) > 0$ for every generating set X of an interpolated free group factor by [23, 24].

COROLLARY 8.7. *Let M be any II_1 factor having a finite generating set Y with $\delta_0(Y) > 1$. If X is any finite generating set for M, then $\mathscr{I}(X) > 0$.*

REFERENCES

[1] R. G. DOUGLAS AND C. PEARCY, *Von Neumann algebras with a single generator*, Michigan Math. J., 16:21–26, 1969.
[2] K. DYKEMA AND U. HAAGERUP, *Invariant subspaces of the quasinilpotent DT-operator*, J. Funct. Anal., 209(2):332–366, 2004.
[3] K. J. DYKEMA, *Free products of hyperfinite von neumann algebras and free dimension*, Duke Math. J., 69(1):97–119, 1993.
[4] K. J. DYKEMA, *Interpolated free group factors*, Pacific J. Math., 163(1):123–135, 1994.
[5] K. J. DYKEMA AND U. HAAGERUP, *DT-operators and decomposability of Voiculescu's circular operator*, Amer. J. Math., 126(1):121–189, 2004.
[6] K. J. DYKEMA, K. JUNG, AND D. SHLYAKHTENKO, *The microstates free entropy dimension of any DT-operator is 2*, Doc. Math., 10:247–261 (electronic), 2005.
[7] L. GE, *Applications of free entropy to finite von neumann algebras*, Amer. J. Math., 119(2):467–485, 1997.
[8] L. GE AND S. POPA, *On some decomposition properties for factors of type* II_1, Duke Math. J., 94(1):79–101, 1998.
[9] L. GE AND J. SHEN, *Generator problem for certain property T factors*, Proc. Natl. Acad. Sci. USA, 99(2):565–567 (electronic), 2002.
[10] D. HADWIN AND J. SHEN, *Free orbit dimension of finite von Neumann algebras*, J. Funct. Anal., 249(1):75–91, 2007.
[11] V. F. R. JONES, *Index for subfactors*, Invent. Math., 72(1):1–25, 1983.
[12] K. JUNG, *A free entropy dimension lemma*, Pacific J. Math., 211(2):265–271, 2003.
[13] K. JUNG, *The free entropy dimension of hyperfinite von Neumann algebras*, Trans. Amer. Math. Soc., 355(12):5054–5089, 2003.
[14] K. JUNG, *Strongly 1-bounded von neumann algebras*, Geom. Func. Anal., 17 (4):1180–1200, 2007.
[15] K. JUNG, *A hyperfinite inequality for free entropy dimension*, Proc. Amer. Math. Soc., 134(7):2099–2108 (electronic), 2006.

[16] C. PEARCY, *W^* -algebras with a single generator*, Proc. Amer. Math. Soc., 13:831–832, 1962.
[17] F. RĂDULESCU, *The fundamental group of the von Neumann algebra of a free group with infinitely many generators is $\mathbb{R}_+$*, J. Amer. Math. Soc., 5(3):517–532, 1992.
[18] F. RĂDULESCU, *Random matrices, amalgamated free products and subfactors of the von Neumann algebra of a free group, of noninteger index*, Invent. Math., 115(2):347–389, 1994.
[19] J. SHEN, *Singly generated II_1 factors*, J. Operator Theory, to appear. arXiv:math.OA/0511327, 2005.
[20] A. M. SINCLAIR AND R. R. SMITH, *Finite von Neumann algebras and masas*, Volume 351 of London Mathematical Society Lecture Notes Series, Cambridge University Press, Cambridge, 2008.
[21] S. J. SZAREK, *Nets of Grassmann manifold and orthogonal group*, In Proceedings of research workshop on Banach space theory (Iowa City, Iowa, 1981), pages 169–185, Iowa City, IA, 1982. Univ. Iowa.
[22] D. M. TOPPING, *Lectures on von Neumann algebras*, Number 36 in Van Nostrand Reinhol Mathematical Studies. Van Nostrand, London, 1971.
[23] D. VOICULESCU, *The analogues of entropy and of Fisher's information measure in free probability theory, II*, Invent. Math., 118(3):411–440, 1994.
[24] D. VOICULESCU, *The analogues of entropy and of Fisher's information measure in free probability theory, III, The absence of Cartan subalgebras*, Geom. Funct. Anal., 6(1):172–199, 1996.
[25] D. VOICULESCU, K. J. DYKEMA, AND A. NICA, *Free Random Variablees*, volume 1 of CRM Monograph, American Mathematical Society, Providence, 1992.
[26] W. WOGEN, *On generators for von Neumann algebras*, Bull. Amer. Math. Soc., 75:95–99, 1969.

(Received November 27, 2007)

Ken Dykema
Department of Mathematics
Texas A&M University
College Station, TX 77843
USA
e-mail: kdyekma@math.tamu.edu

Allan Sinclair
School of Mathematics
University of Edinburgh
Edinburgh EH9 3JZ
UK
e-mail: a.sinclair@ed.ac.uk

Roger Smith
Department of Mathematics
Texas A&M University
College Station, TX 77843
USA
e-mail: rsmith@math.tamu.edu

Stuart White
Department of Mathematics
University of Glasgow
University Gardens
Glasgow G12 8QW
UK
e-mail: s.white@maths.gla.ac.uk