

Claire Anantharaman-Delaroche

On ergodic theorems for free group actions on noncommutative spaces

Received: 18 January 2004 / Revised version: 28 April 2006 /
Published online: 24 April 2006 – © Springer-Verlag 2006

Abstract. We extend in a noncommutative setting the individual ergodic theorem of Nevo and Stein concerning measure preserving actions of free groups and averages on spheres s_{2n} of even radius. Here we study state preserving actions of free groups on a von Neumann algebra A and the behaviour of $(s_{2n}(x))$ for x in noncommutative spaces $L^p(A)$. For the Cesàro means $\frac{1}{n} \sum_{k=0}^{n-1} s_k$ and $p = +\infty$, this problem was solved by Walker. Our approach is based on ideas of Bufetov. We prove a noncommutative version of Rota “Alternierende Verfahren” theorem. To this end, we introduce specific dilations of the powers of some noncommutative Markov operators.

1. Introduction

Ergodic theorems for measure preserving group actions have a long history, going back to Birkhoff and von Neumann around 1930. Until some ten years ago they mainly concerned actions of amenable groups. Initiated by Arnold and Krylov [2] and carried on by Guivarc’h [15], Grigorchuk [14] and Nevo [31] for free group actions, the ergodic theory of actions of non amenable groups has greatly progressed since, due to the works of Nevo, Nevo-Stein, and Margulis-Nevo-Stein. In particular, these authors have developed powerful techniques well adapted to actions of semi-simple groups.

It is interesting, and not straightforward, to study the analogues of these results in noncommutative (or quantum) probability theory. The setting is a $(W^*$ -)noncommutative probability space, that is, in this paper, a pair (A, φ) where A is a von Neumann algebra and φ a faithful normal state on A . The first result obtained in this framework is the noncommutative individual ergodic theorem due to Lance [28]. He proved that, for any automorphism σ of A which preserves φ , the averages

$$c_n(x) = \frac{1}{n} \sum_{k=0}^{n-1} \sigma^k(x)$$

converge almost uniformly to a σ -invariant element in A . The almost uniform convergence, which will be defined below (see definition 2.8), is the noncommutative

C. Anantharaman-Delaroche: Département de Mathématiques, Université d’Orléans, B. P. 6759, 45067 Orléans Cedex 2, France.

e-mail: claire.anantharaman@univ-orleans.fr

Mathematics Subject Classifications (2000): Primary 46L53; 46L55; Secondary 46L50

Key words or phrases: Noncommutative ergodic theorems – Free group actions.

analogue of almost everywhere convergence. Shortly after, the case of amenable group actions in the noncommutative framework was studied by Conze and Dang Ngoc [10]. Lance's result was further extended by Kümmerer [27] (replacing σ by any normal positive map such that $\sigma(1) \leq 1$ and $\varphi \circ \sigma \leq \varphi$) and Yeadon [46] who studied the convergence of $(c_n(x))$ when x belongs to the L^1 -space $L^1(A, \tau)$, in case τ is a faithful normal semi-finite trace. For more informations on the state of subject before 1990 we refer to the above mentioned papers and to the books of Jajte [16], [17]. The most recent and important developments in that direction are due to Junge and Xu [19, 20]. As a consequence of their maximal ergodic inequalities in noncommutative L^p -spaces, they generalized in particular Yeadon's theorem to every space $L^p(A, \tau)$, for $p \geq 1$.

The case of free group actions in quantum probability was considered by Walker in [45]. In order to state his result, as well as the results of Nevo [31] and Nevo-Stein [32] relative to measure preserving free group actions we need to introduce some notations. We shall denote by \mathbb{F}_d the free group with d generators g_1, \dots, g_d and by $|w|$ the length of $w \in \mathbb{F}_d$, that is the smallest number of generators, together with their inverses, needed to write the word w . Given d automorphisms $\sigma_1, \dots, \sigma_d$ of a von Neumann algebra A , we consider the corresponding homomorphism $w \mapsto \sigma_w$ from \mathbb{F}_d into the group $\text{Aut}(A)$ of automorphisms of A defined by assigning σ_i to g_i , $1 \leq i \leq d$. The averaging operator s_n on the sphere \mathcal{S}_n of radius n is defined as

$$s_n(x) = \frac{1}{\#\mathcal{S}_n} \sum_{w \in \mathcal{S}_n} \sigma_w(x),$$

for $x \in A$, where $\mathcal{S}_n = \{w \in \mathbb{F}_d : |w| = n\}$. Let us recall the Nevo-Stein theorem.

Theorem (Nevo-Stein, [31, 32]). *Let $A = L^\infty(X, \mu)$ be the commutative von Neumann algebra associated with a probability space (X, μ) . Let σ_i , $1 \leq i \leq d$, be automorphisms of A induced by measure preserving transformations of X . Let $f \in L^p(X, \mu)$. Then*

- (i) *If $p > 1$, the sequence $(s_{2n}(f))$ converges almost everywhere and in $L^p(X, \mu)$ to the conditional expectation $\mathbb{E}(f|\mathcal{I}_2)$ with respect to the σ -field \mathcal{I}_2 of $\mathbb{F}_d^{(2)}$ -invariant measurable subsets, where $\mathbb{F}_d^{(2)}$ is the subgroup of even length words.*
- (ii) *If $p \geq 1$, the sequence $\frac{1}{n} \sum_{k=0}^{n-1} s_k(f)$ converges almost everywhere and in $L^p(X, \mu)$ to the conditional expectation $\mathbb{E}(f|\mathcal{I})$ with respect to the σ -field \mathcal{I} of \mathbb{F}_d -invariant measurable subsets.*

In the noncommutative setting, Walker obtained the following generalization of a part of the Nevo-Stein theorem.

Theorem (Walker, [45]). *Let σ_i , $1 \leq i \leq d$, be automorphisms of a von Neumann algebra A which leave a faithful normal state φ invariant. For $x \in A$, the sequence $\frac{1}{n} \sum_{k=0}^{n-1} s_k(x)$ converges almost uniformly to an element $\hat{x} \in A$.*

The proof given by Walker uses the result of Nevo and Stein showing that the Cesàro means $\frac{1}{n} \sum_{k=0}^{n-1} s_k$ are dominated by Cesàro averages of powers of the only

contraction s_1 . In the commutative case, Nevo and Stein could conclude with the help of the classical Hopf-Dunford-Schwartz maximal inequality. In the noncommutative case, this inequality has to be replaced by a maximal inequality due to Goldstein. In both cases, delicate spectral estimations are also needed.

For a recent version of Walker's result when $x \in L^1(A, \tau)$, where τ is a normal faithful semi-finite trace on A , we refer to [8]. In this paper the approach of the problem combines the one of Walker with an appropriate noncommutative Banach principle.

Bufetov has proposed in [7] a totally different proof of the Nevo-Stein theorem. In addition to being very simple, another advantage of his method is that it allows to extend part (i) of the Nevo-Stein theorem to functions f in the class $L \log L$.

The aim of this paper is to show how the method of Bufetov can be adapted to quantum probability theory. Combined with recent noncommutative martingale convergence results due to Junge [18] and to Defant and Junge [12], it gives the following result.

Theorem (Noncommutative Nevo-Stein ergodic theorem). *Let σ_i , $1 \leq i \leq d$, be automorphisms of a von Neumann algebra A which leave a faithful normal state φ invariant. Let $x \in L^p(A)$ with $p \in]1, +\infty[$.*

- (i) *the sequence $s_{2n}(x)$ converges bilaterally almost surely to the conditional expectation of x with respect to the $\mathbb{F}_d^{(2)}$ -invariant elements in $L^p(A, \varphi)$.*
- (ii) *the sequence $\frac{1}{n} \sum_{k=0}^{n-1} s_k(x)$ converges bilaterally almost surely to the conditional expectation of x with respect to the \mathbb{F}_d -invariant elements in $L^p(A, \varphi)$.*

Moreover, in both cases, the convergence holds almost surely whenever $p \in [2, +\infty[$ and almost uniformly when $p = +\infty$.

The various notions of convergence appearing in this statement are analogues of almost everywhere convergence. They are defined in 2.8.

This paper is organized as follows. In the next section, we shall begin by explaining the notions of noncommutative probability theory used to state and prove the above theorem. The reader only interested in a very short proof of the result of Walker may skip these technical preliminaries and pass directly to section 3. In theorem 3.1 we show that the techniques of Bufetov developed in [6] apply immediately to extend the theorem of Walker. The idea is to relate the problem to the study of an appropriate Markov operator $P : A^{2d} \rightarrow A^{2d}$ and to use available ergodic theorems for the sequence $\frac{1}{n} \sum_{k=0}^{n-1} P^k$ of averaging operators.

Part (i) of the noncommutative Nevo-Stein theorem requires a more involved analysis. In section 4 we first construct a suitable "noncommutative Markov chain" attached to the Markov operator P (see theorem 4.1) and we prove (corollary 4.2) a noncommutative version of Rota "Alternierende Verfahren" theorem, that is a convergence result for the sequence $P^n \circ (P^*)^n$ where P^* is an adjoint of P (as defined in section 2). Following Bufetov, we show in section 5 how this last result implies the noncommutative Nevo-Stein theorem.

The above operator P is factorizable in the sense of definition 6.2. Roughly speaking, this means that $P = j_0^* \circ j_1$, where j_0, j_1 are homomorphisms preserving

given states in a strong sense (see definition 2.6). Note that in the commutative case, any measure preserving Markov operator is factorizable. In the last section, we show how a Daniell-Kolmogorov type extension of the classical construction of the Markov chain associated with a transition probability and an initial distribution can be carried out for every factorizable Markov operator. Such constructions already appeared in the theory of quantum stochastic processes (see [38] and [3] for instance), but here we insist on having, in addition, a good behaviour with respect to P^* (see formula (6.3)). In particular, the noncommutative Rota theorem holds for such an operator P . However, since the construction involves amalgamated free products of von Neumann algebras, we have chosen to provide in section 4 a much simpler construction for the concrete Markov operator used in the proof of the noncommutative Nevo-Stein theorem.

2. Preliminaries

Let us briefly summarize the main concepts and results needed in this paper.

2.1. Tomita-Takesaki modular theory and noncommutative L^p -spaces

We refer to [42], [24], [5, Section 2.5] for general backgrounds on the modular theory of von Neumann algebras and to [23], [26], [43] for details on noncommutative L^p -spaces.

When τ is a normal faithful semi-finite trace on a von Neumann algebra A , the spaces $L^p(A, \tau)$ are easy to introduce and well understood (see [39], and [29] for a short exposition).

Let us concentrate on the case of a noncommutative probability space (A, φ) where φ is a normal faithful state. We denote by $L^2(A, \varphi)$ the completion of A with respect to the scalar product $\langle a_1, a_2 \rangle = \varphi(a_2^* a_1)$ and by ξ_φ the unit of A viewed as a vector of $L^2(A, \varphi)$. We identify A with its GNS-representation in $L^2(A, \varphi)$. In particular, the inclusion $A \subset L^2(A, \varphi)$ is given by the map $a \mapsto a\xi_\varphi$. We denote by S_φ the closure of the operator $a\xi_\varphi \mapsto a^*\xi_\varphi$. As usual, its polar decomposition is written $S_\varphi = J_\varphi \Delta_\varphi$. Recall that J_φ is an anti-unitary involution and that $\Delta_\varphi^{it} A \Delta_\varphi^{-it} = A$ for all $t \in \mathbb{R}$. The modular automorphism group associated with (A, φ) is the one-parameter automorphism group of A defined by $\sigma_t^\varphi(a) = \Delta_\varphi^{it} a \Delta_\varphi^{-it}$ for $a \in A$ and $t \in \mathbb{R}$.

There are several ways to introduce noncommutative L^p -spaces. We follow the construction of Haagerup [23]. For simplicity, let us set $\sigma_t = \sigma_t^\varphi$ for $t \in \mathbb{R}$. Given a concrete representation of A on a Hilbert space H (for instance the GNS-representation in $L^2(A, \varphi)$), recall that the crossed product $A \rtimes_\sigma \mathbb{R}$ is the von Neumann algebra of operators acting on $L^2(\mathbb{R}, H)$ generated by $\pi(a)$, $a \in A$, and $\lambda(s)$, $s \in \mathbb{R}$, where for $\xi \in L^2(\mathbb{R}, H)$ and $t \in \mathbb{R}$,

$$\pi(a)(\xi)(t) = \sigma_{-t}(a)\xi(t) \quad \text{and} \quad \lambda(s)(\xi)(t) = \xi(t - s).$$

This crossed product does not depend on the choice of H . Since π is a normal faithful representation of A on $L^2(\mathbb{R}, H)$, one identifies A with $\pi(A)$. There is a

one parameter automorphism group $t \mapsto \hat{\sigma}_t$ of $A \rtimes_{\sigma} \mathbb{R}$, implemented by the unitary representation $t \mapsto W(t)$ of \mathbb{R} on $L^2(\mathbb{R}, H)$, where $W(t)(\xi)(s) = e^{-its} \xi(s)$ for $\xi \in L^2(\mathbb{R}, H)$ and $s, t \in \mathbb{R}$. One has

$$\hat{\sigma}_t(\lambda(s)) = W(t)\lambda(s)W(t)^* = e^{-ist}\lambda(s) \quad \text{for } s, t \in \mathbb{R},$$

and

$$A = \{x \in A \rtimes_{\sigma} \mathbb{R} : \hat{\sigma}_t(x) = x, \forall t \in \mathbb{R}\}. \tag{2.1}$$

The crossed product $A \rtimes_{\sigma} \mathbb{R}$ has a canonical normal semi-finite faithful trace τ satisfying $\tau \circ \hat{\sigma}_t = e^{-t}\tau$ for all $t \in \mathbb{R}$. We can therefore introduce the topological $*$ -algebra $\mathcal{M}(A \rtimes_{\sigma} \mathbb{R}, \tau)$ formed of the closed densely defined operators on $L^2(\mathbb{R}, H)$, affiliated with $A \rtimes_{\sigma} \mathbb{R}$, that are measurable with respect to τ [42, Chapter IX]. This algebra is the substitute for the algebra of measurable functions in the commutative case. Following the point of view of Haagerup, for $p \in [1, +\infty]$ we define the noncommutative L^p -space $L^p(A, \varphi)$ as

$$L^p(A, \varphi) = \{x \in \mathcal{M}(A \rtimes_{\sigma} \mathbb{R}, \tau) : \hat{\sigma}_t(x) = e^{-t/p}x, \forall t \in \mathbb{R}\}.$$

It is an ordered Banach space. We shall not describe its norm $\|\cdot\|_p$ here (see [23]). Its positive cone $L^p(A, \varphi)_+$ is the intersection of $L^p(A, \varphi)$ with the cone formed by the positive measurable operators. Since $L^p(A, \varphi)$ does not depend on φ , up to order preserving isometry, we shall write $L^p(A)$ for its abstract version. These L^p -spaces behave as their commutative analogue with respect to duality.

Note that we have given above two definitions of $L^2(A, \varphi)$. This is not confusing since, due to the unicity of the standard form of a von Neumann algebra, the two spaces can be identified in a natural way.

Observe also that, as a consequence of (2.1), we have $L^\infty(A, \varphi) = A$. On the other hand, $L^1(A, \varphi)$ is identified with the predual A_* of A in the following way: any normal positive linear form ω on A induces a dual weight $\hat{\omega}$ on $A \rtimes_{\sigma} \mathbb{R}$ whose Radon-Nikodym derivative h_ω with respect to τ belongs to $L^1(A, \varphi)$ and this map $\omega \mapsto h_\omega$ extends to an order preserving isometry between A_* and $L^1(A, \varphi)$.

As concrete subspaces of $\mathcal{M}(A \rtimes_{\sigma} \mathbb{R}, \tau)$, the L^p -spaces satisfy the odd relation

$$L^p(A, \varphi) \cap L^q(A, \varphi) = \{0\} \quad \text{if } p \neq q.$$

However, one can define nice positive embeddings as follows. Let us denote by $D = D_\varphi (= h_\varphi)$ the Radon-Nikodym derivative of the dual weight $\hat{\varphi}$ with respect to τ .

Lemma 2.1 (Theorem 1.7, [13]). *Let $p \in [1, +\infty]$ and let p' be such $1/p + 1/p' = 1$.*

- (i) $\iota_p : x \mapsto D^{\frac{1}{2p}} x D^{\frac{1}{2p}}$ is an embedding from A_+ into $L^p(A, \varphi)_+$ with dense range;
- (ii) $\kappa_p : x \mapsto D^{\frac{1}{2p'}} x D^{\frac{1}{2p'}}$ is an embedding from $L^p(A, \varphi)_+$ into $L^1(A, \varphi)_+$ with dense range.

2.2. Markov operators in noncommutative probability theory

In the commutative case, there are several variant of the notion of Markov operator. We shall adopt a definition well suited to be extended to the noncommutative setting.

Definition 2.2 ([25]). *A Markov operator on a probability space (X, μ) is a positive unital normal operator Q from $L^\infty(X, \mu)$ into itself. We say that μ is Q -stationary, or that Q is μ -preserving if $\int_X Q(f)d\mu = \int_X f d\mu$ for every $f \in L^\infty(X, \mu)$.*

Remark 2.3. Usually, Markov operators are defined to be positive contractions from $L^1(X, \mu)$ into itself, preserving the constant function 1 (see [30, page 178]). For μ -preserving operators the two definitions coincide. Indeed, let Q be a μ -preserving operator in the sense of definition 2.2. Then the (predual) operator $Q_* : L^1(X, \mu) \rightarrow L^1(X, \mu)$ satisfies $Q_*(1) = 1$ and therefore preserves the subspace $L^\infty(X, \mu)$. Hence, the dual of $Q_*|_{L^\infty(X, \mu)}$ gives a unique extension of Q to a positive contraction of $L^1(X, \mu)$.

A normal unital completely positive map from a noncommutative probability space (A, φ) into another one (B, ψ) will usually be called a *Markov operator*. As in the commutative case, Q can be extended to L^p -spaces if $\psi \circ Q = \varphi$. More generally we have the following result:

Lemma 2.4 ([13],[22]). *Let Q be a normal positive map from (A, φ) into (B, ψ) with $Q(1) \leq 1$ and $\psi \circ Q \leq \varphi$. The map $Q_{(p)} : D_\varphi^{\frac{1}{2p}} A D_\varphi^{\frac{1}{2p}} \rightarrow D_\psi^{\frac{1}{2p}} B D_\psi^{\frac{1}{2p}}$, defined by*

$$Q_{(p)}(D_\varphi^{\frac{1}{2p}} a D_\varphi^{\frac{1}{2p}}) = D_\psi^{\frac{1}{2p}} Q(a) D_\psi^{\frac{1}{2p}}$$

for $a \in A$, extends to a positive contraction $Q_{(p)}$ from $L^p(A, \varphi)$ onto $L^p(B, \psi)$.

Hereafter, we shall drop the subscript p and therefore write Q instead of $Q_{(p)}$.

The case of conditional expectations will be especially useful hereafter. Given (A, φ) as above, we shall say that a von Neumann subalgebra A_1 of A is φ -invariant if it is invariant under the modular automorphism group σ_t^φ of φ (i. e. $\sigma_t^\varphi(A_1) \subset A_1$ for every $t \in \mathbb{R}$). This condition is equivalent (see [41]) to the existence of a (unique) normal conditional expectation $\mathbb{E} : A \rightarrow A_1$ such that $\varphi \circ \mathbb{E} = \varphi$. Moreover, in this situation, we have $\mathbb{E} \circ \sigma_t^\varphi = \sigma_t^\varphi \circ \mathbb{E}$ for $t \in \mathbb{R}$. If φ_1 denotes the restriction of φ to A_1 , the space $L^p(A_1, \varphi_1)$ is naturally embedded as a Banach subspace of $L^p(A, \varphi)$ and \mathbb{E} extends to a contractive projection from $L^p(A, \varphi)$ onto $L^p(A_1, \varphi_1)$, still denoted \mathbb{E} (see lemma 2.4 or [21, Section 2]). For $x \in L^p(A, \varphi)$, we shall say that $\mathbb{E}(x)$ is the *conditional expectation of x with respect to A_1* .

In the commutative case, given a μ -preserving Markov operator Q there is a unique Markov operator on (X, μ) , that we shall denote Q^* , such that $\int_X Q^*(f)g d\mu = \int_X f Q(g) d\mu$ for every $f, g \in L^\infty(X, \mu)$ (take $Q^* = Q_*|_{L^\infty(X, \mu)}$ in remark 2.3

above). In general the situation is more subtle, and we have to take the modular automorphism groups of the noncommutative probability spaces into account. Recall that for any normal unital completely positive map Q from (A, φ) to (B, ψ) such that $\psi \circ Q = \varphi$, there always exists a unital completely positive map $Q^* : B \rightarrow A$ with

$$\varphi(Q^*(b)\sigma_{-i/2}^\varphi(a)) = \psi(\sigma_{i/2}^\psi(b)Q(a)) \tag{2.2}$$

for every σ^φ -analytic element $a \in A$ and every σ^ψ -analytic element $b \in B$ (see [1], [9]). There exists Q^* such that

$$\varphi(Q^*(b)a) = \psi(bQ(a)) \tag{2.3}$$

for all $a \in A, b \in B$ if and only if Q intertwines the modular automorphism groups of φ and ψ . We give below a proof of this result, for the reader’s convenience.

Lemma 2.5 (Proposition 6.1, [1]).

Let Q be a completely positive normal unital map from (A, φ) into (B, ψ) . The two following conditions are equivalent:

- (i) there exists a normal unital completely positive map $R : (B, \psi) \rightarrow (A, \varphi)$ such that $\varphi(R(b)a) = \psi(bQ(a))$ for every $a \in A$ and $b \in B$;
- (ii) $\psi \circ Q = \varphi$ and $\sigma_t^\psi \circ Q = Q \circ \sigma_t^\varphi$ for every $t \in \mathbb{R}$.

Proof. Note first that when Q is a unital completely positive map with $\psi \circ Q = \varphi$, there exists a unique contraction $V : L^2(A, \varphi) \rightarrow L^2(B, \psi)$ such that $V(a\xi_\varphi) = Q(a)\xi_\psi$ for $a \in A$. Moreover, we have $V S_\varphi \subset S_\psi V$.

Assume first the existence of R as in (i). In particular we have $\psi \circ Q = \varphi$ and $\varphi \circ R = \psi$. It is easily checked that the adjoint V^* of V is the operator constructed similarly from R . We have $V^* S_\psi \subset S_\varphi V^*$, so that $V S_\varphi^* S_\psi \subset S_\psi^* S_\varphi V$ and therefore $V \Delta_\psi^{it} = \Delta_\varphi^{it} V$ for every $t \in \mathbb{R}$. We can conclude that $\sigma_t^\psi \circ Q = Q \circ \sigma_t^\varphi$ since

$$\sigma_t^\psi \circ Q(a)\xi_\psi = \Delta_\psi^{it} V(a\xi_\varphi) = V \Delta_\varphi^{it}(a\xi_\varphi) = Q \circ \sigma_t^\varphi(a)\xi_\psi$$

for $a \in A$ and $t \in \mathbb{R}$.

Let us now prove that (ii) \Rightarrow (i). Since $\sigma_t^\psi \circ Q = Q \circ \sigma_t^\varphi$ we get $V \Delta_\psi^{it} = \Delta_\varphi^{it} V$ for every $t \in \mathbb{R}$ and therefore $J_\psi V = V J_\varphi$. Obviously we have $V(A_+ \xi_\varphi) \subset B_+ \xi_\psi$. For $b \in B_+$ and $a' = J_\varphi a J_\varphi \in A'_+$, observe that

$$\langle V^* b \xi_\psi, a' \xi_\varphi \rangle = \langle b \xi_\psi, J_\psi V a \xi_\varphi \rangle \geq 0.$$

It follows that $V^*(B_+ \xi_\psi) \subset \overline{A_+ \xi_\varphi}$ since

$$\overline{A_+ \xi_\varphi} = \{\eta \in L^2(A, \varphi) : \langle \eta, a' \xi_\varphi \rangle \geq 0, \forall a' \in A'_+\}$$

(see [5, Proposition 2.5.27]).

Since $\|V^*\| = 1$ and $\langle V^* \xi_\psi, \xi_\varphi \rangle = 1$, we see that $V^* \xi_\psi = \xi_\varphi$. It follows from [5, Lemma 3.2.19] that $V^*(B_+ \xi_\psi) \subset A_+ \xi_\varphi$.

For $b \in B_+$, let us define $R(b)$ as the unique element $a \in A_+$ such that $V^* b \xi_\psi = a \xi_\varphi$. Then it is easily checked that R fulfils the conditions of (i). \square

Definition 2.6. Let Q be a completely positive normal unital map from (A, φ) into (B, ψ) . If the equivalent conditions of the previous lemma are satisfied, we say that the pair (ψ, φ) is stationary with respect to Q or Q -stationary. We shall also say that Q is a (ψ, φ) -preserving Markov operator. The operator Q^* will be called the adjoint of Q with respect to (ψ, φ) .

When $(A, \varphi) = (B, \psi)$, we say that φ is stationary with respect to Q or Q -stationary, or that Q is a φ -preserving Markov operator.

We insist on the fact that Q -stationarity is strictly stronger (in general) than the equality $\psi \circ Q = \varphi$.

- Examples 2.7.* (a) *Finite von Neumann algebras.* The modular theory is trivial if φ and ψ are chosen to be tracial normal faithful states. Therefore, the maps Q that we consider are normal completely positive unital maps with $\psi \circ Q = \varphi$.
 (b) *Conditional expectations.* As already observed, normal conditional expectations \mathbb{E} with $\varphi \circ \mathbb{E} = \varphi$ automatically commute with σ^φ .
 (c) *Homomorphisms.* Let Q be a unital normal $*$ -homomorphism from A into B such that $\psi \circ Q = \varphi$. Then Q satisfies the equivalent conditions of lemma 2.5 if and only if $Q(A)$ is invariant under the modular automorphism group of ψ . Moreover in this case, for $b \in B$ we have $Q^*(b) = Q^{-1}(\mathbb{E}(b))$, where \mathbb{E} is the ψ -preserving conditional expectation from B onto $Q(A)$.

2.3. Almost sure convergence in noncommutative probability theory

Finally, let us introduce noncommutative substitutes for almost sure convergence (see [16], [17] and [12]).

Definition 2.8. Let A be a von Neumann algebra with a faithful normal state φ .

- (a) We say that a sequence (x_n) of elements of A converges to 0 almost uniformly (resp. bilaterally almost uniformly) if for every $\varepsilon > 0$ there is a projection $e \in A$ with $\varphi(1 - e) \leq \varepsilon$ and $\lim_{n \rightarrow \infty} \|x_n e\|_\infty = 0$ (resp. $\lim_{n \rightarrow \infty} \|e x_n\|_\infty = 0$).
 (b) Let $p \in [1, +\infty]$. We say that a sequence (x_n) of elements of $L^p(A, \varphi)$ converges to 0 almost surely if for every $\varepsilon > 0$ there is a projection $e \in A$ and a family $(a_{n,k})$ in A such that

$$\varphi(1 - e) \leq \varepsilon, \quad x_n = \sum_k (a_{n,k} D^{1/p}), \quad \text{and} \quad \lim_{n \rightarrow \infty} \left\| \sum_k (a_{n,k} e) \right\|_\infty = 0,$$

where the two series converge in $L^p(A, \varphi)$ and A respectively.

- (c) Let $p \in [1, +\infty]$. We say that a sequence (x_n) of elements of $L^p(A, \varphi)$ converges to 0 bilaterally almost surely if for every $\varepsilon > 0$ there is a projection $e \in A$ and a family $(a_{n,k})$ in A such that

$$\varphi(1 - e) \leq \varepsilon, \quad x_n = \sum_k (D^{1/2p} a_{n,k} D^{1/2p}), \quad \text{and} \quad \lim_{n \rightarrow \infty} \left\| \sum_k (e a_{n,k} e) \right\|_\infty = 0,$$

where the two series converge in $L^p(A, \varphi)$ and A respectively.

When $A = L^\infty(X, m)$, with (X, m) a probability measure, all these notions of convergence coincide, via Egorov’s theorem, with the almost everywhere point-wise convergence. When A is a semi-finite von Neumann algebra equipped with a normal faithful semi-finite trace τ , the notions of almost uniform convergence and bilateral almost uniform convergence can still be introduced as in definition 2.8 (a) for every sequence in $L^p(A, \tau)$ and all $p \in [1, +\infty]$.

3. Ergodic theorems for some Cesàro means

Let us begin by introducing some notations. Let $[p(ij)]$ be a stochastic matrix, whose rows and columns are indexed by a finite set I and let $(p(i))_{i \in I}$ be a stationary distribution, that is a probability measure on I such that $p(j) = \sum_{i \in I} p(i)p(ij)$ for $j \in I$. We assume that $p(i) > 0$ for all $i \in I$.

We shall usually view elements $w \in I^n$ as words of length n in the alphabet I and write $w = w_0w_1 \cdots w_{n-1}$ where w_k is the $(k + 1) - th$ component of w . For $w \in I^n, n \geq 2$, we set

$$p_{n-1}(w) = p(w_0)p(w_0w_1) \cdots p(w_{n-2}w_{n-1})$$

and

$$I(n) = \{w \in I^n : p_{n-1}(w) \neq 0\}.$$

We set $p_0(i) = p(i)$ and $I(1) = I$.

Let L be a linear space and consider, for $i \in I$, linear operators $P_i : L \rightarrow L$. For $n \geq 1$ and $w \in I(n)$, we set $P_w = P_{w_0} \circ \cdots \circ P_{w_{n-1}}$. We define the operators s_n and c_n by the formulas

$$s_n = \sum_{w \in I(n)} p_{n-1}(w)P_w \quad \text{if } n \geq 1, \quad s_0 = \text{Id}_A,$$

$$c_n = \frac{1}{n} \sum_{k=0}^{n-1} s_k \quad \text{if } n \geq 1.$$

For us L will be a noncommutative L^p -space and we shall study the convergence of the sequences $(s_n(x))$ and $(c_n(x))$, where $x \in L$.

Following the idea of Bufetov [6], we assign to our data the operator $P : L^I \rightarrow L^I$ such that

$$P(x)_i = \sum_{j \in I} p(ij)P_i(x_j), \quad i \in I,$$

for all $x = (x_i)_{i \in I}$.

Given $x \in L$, we denote by \tilde{x} the element of L^I with $\tilde{x}_i = x$ for all i . The main observation, easily proved by induction, is that, for $x \in L$ and $n \geq 1$,

$$P^n(\tilde{x})_i = \frac{1}{p(i)} \sum_{\substack{w \in I(n) \\ w_0=i}} p_{n-1}(w)P_w(x), \tag{3.1}$$

so that

$$s_n(x) = \sum_{i \in I} p(i) P^n(\tilde{x})_i \tag{3.2}$$

and

$$c_n(x) = \sum_{i \in I} p(i) \left(\frac{1}{n} \sum_{k=0}^{n-1} P^k(\tilde{x})_i \right). \tag{3.3}$$

In this section, we shall concentrate on $(c_n(x))$ which is much easier to deal with.

Theorem 3.1. *Let $[p(ij)]$ and $(p(i))_{i \in I}$ be as above and (A, φ) be a noncommutative probability space. Let $P_i : A \rightarrow A$ be normal positive contractions such that*

$$\varphi \circ P_i \leq \varphi \ (i \in I). \text{ For every } x \in A, \text{ the sequence } c_n(x) = \frac{1}{n} \sum_{k=0}^{n-1} \left(\sum_{w \in I(n)} p_{n-1}(w) \right.$$

$\left. P_w(x) \right)$ converges almost uniformly to an element $\hat{x} \in A$.

Proof. We use the notations introduced above where L is here taken to be A . The space A^I is a von Neumann algebra, that we denote by B , and we equip B with the normal faithful state ϕ defined by

$$\phi(b) = \sum_{i \in I} p(i) \varphi(b_i)$$

where $b = (b_i)_{i \in I}$.

Then $P : B \rightarrow B$ defined as above is a normal positive contraction such that $\phi \circ P \leq \phi$. The ergodic theorem of Kümmerer [27] implies that for every $b \in B$

the sequence $\left(\frac{1}{n} \sum_{k=0}^{n-1} P^k(b) \right)$ converges almost uniformly in B to a P -invariant element. It follows from (3.3) that for every $x \in A$ that sequence $c_n(x)$ converges almost uniformly in A . □

Remark 3.2. The situation considered by Walker is the following particular case:

– I is a set of $2d$ elements, that we write

$$I = \{-d, -d + 1, \dots, -1, 1, \dots, d - 1, d\};$$

- for $1 \leq i \leq d$, P_i is an automorphism σ_i of A such that $\varphi \circ \sigma_i = \varphi$ and we set $\sigma_{-i} = \sigma_i^{-1}$;
- $[p(ij)]$ is the stochastic matrix defined by

$$p(ij) = \begin{cases} 0 & \text{if } i = -j, \\ \frac{1}{2d-1} & \text{otherwise.} \end{cases}$$

– $(p(i))_{i \in I}$ is the uniform probability measure on I , i.e. $p(i) = 1/2d$ for all $i \in I$.

Obviously, in this case, $I(n)$ is the sphere \mathcal{S}_n formed by the words of length n .

The identification of the limit in theorem 3.1 is a more difficult problem. However, in the situation of the previous remark, it is easy to show that the limit \hat{x} is the conditional expectation of x on the subalgebra of fixed points under the action of \mathbb{F}_d (see the proof of the Nevo-Stein theorem in section 4).

Similarly, using the noncommutative individual Dunford-Schwartz ergodic theorem obtained recently by Junge and Xu [19, Corollaire 5], [20, Corollary 6.4], we get:

Theorem 3.3. *Let A be a semifinite von Neumann algebra equipped with a normal semi-finite faithful trace τ . Let $P_i : A \rightarrow A, i \in I$, be normal positive contractions such that $\tau \circ P_i \leq \tau$. We are given $[p(ij)]$ and $(p(i))_{i \in I}$ as in theorem 3.1. Then, for every $x \in L^p(A)$ and $p \in [1, +\infty[$, the sequence $(c_n(x))$ converges bilaterally almost uniformly to an element of $L^p(A)$. If $p \geq 2$, the convergence also holds almost uniformly. It also holds in L^p -norm for $p \in]1, +\infty[$.*

As said in the introduction, the above theorem has been proved by Chilin, Litvinov and Skalski for $p = 1$, in a different way (see [8]).

The study of the sequence (s_n) is much more delicate and requires additional assumptions. We shall only consider the situation described in remark 3.2. We need a generalization of the Rota ‘‘Alternierende Verfahren’’ theorem that we shall examine in the next section.

4. Noncommutative Rota theorem in a simple case

Let (A, φ) be a noncommutative probability space. We are given normal unital endomorphisms σ_i of A where i belongs to a finite set I of indices. We assume that $\varphi \circ \sigma_i = \varphi$ and $\sigma_i \circ \sigma_t^\varphi = \sigma_t^\varphi \circ \sigma_i$ for $i \in I$ and $t \in \mathbb{R}$. We are also given, as before, a stochastic matrix $[p(ij)]$ and a stationary distribution $(p(i))_{i \in I}$ with $p(i) > 0$ for all $i \in I$

We set $B = A^I$ and ϕ will be the state defined by

$$\phi(b) = \sum_{i \in I} p(i)\varphi(b_i).$$

We still introduce the Markov operator P from B into B defined by

$$P(b)_i = \sum_{j \in I} p(ij)\sigma_i(b_j), \tag{4.1}$$

for all $b \in B$ and $i \in I$. Each σ_i has an adjoint σ_i^* with respect to φ , and therefore P is a ϕ -preserving Markov operator. We have, for $b \in B$ and $i \in I$,

$$P^*(b)_i = \frac{1}{p(i)} \sum_{j \in I} p(j)p(ji)\sigma_j^*(b_j).$$

We need the following dilation result.

Theorem 4.1 (Noncommutative Kolmogorov construction). *There exist*

- a von Neumann algebra \mathcal{B} ,
- a normal faithful state Φ on \mathcal{B} ,
- a normal unital endomorphism $\beta : \mathcal{B} \rightarrow \mathcal{B}$ with $\Phi \circ \beta = \Phi$ and $\sigma_t^\Phi \circ \beta = \beta \circ \sigma_t^\Phi$ for $t \in \mathbb{R}$,
- a normal unital homomorphism $J_0 : \mathcal{B} \rightarrow \mathcal{B}$ with $\Phi \circ J_0 = \phi$ and $\sigma_t^\Phi \circ J_0 = J_0 \circ \sigma_t^\phi$ for $t \in \mathbb{R}$,

such that, if we set $J_n = \beta^n \circ J_0$ for $n \geq 0$, and if $\mathcal{B}_{[n]}$ and $\mathcal{B}_{\leq n}$ denotes the von Neumann subalgebras of \mathcal{B} generated by $\bigcup_{k \leq n} J_k(B)$ and $\bigcup_{k \geq n} J_k(B)$ respectively, then

- (i) the algebras $\mathcal{B}_{[n]}$ and $\mathcal{B}_{\leq n}$ are Φ -invariant;
- (ii) if $\mathbb{E}_{[n]}$ and $\mathbb{E}_{\leq n}$ are the corresponding Φ -preserving conditional expectations, for $n \in \mathbb{N}$ and $q \geq n$, we have

$$\mathbb{E}_{[n]} \circ J_q = J_n \circ P^{q-n}, \tag{4.2}$$

$$\mathbb{E}_{\leq n} \circ J_0 = J_n \circ (P^*)^n. \tag{4.3}$$

Let us show first how the following corollary can be deduced from this theorem.

Corollary 4.2 (Noncommutative Rota theorem). *For $p \in]1, +\infty]$ and $x \in L^p(B)$, the sequence $(P^n \circ (P^*)^n(x))$ converges bilaterally almost surely. For $p \in [2, +\infty]$ the convergence also holds almost surely.*

Proof. We have $\mathbb{E}_{[0]} \circ \mathbb{E}_{\leq n} \circ J_0(x) = J_0 \circ P^n \circ (P^*)^n(x)$ for $x \in B$ and, using lemma 2.4, this formula extends to $L^p(B)$, $p > 1$. Therefore the result is a consequence of the lemma below. □

Lemma 4.3. *Let (\mathcal{B}, Φ) be a noncommutative probability space and let (B_n) be a decreasing sequence of Φ -invariant von Neumann subalgebras. We set $B_\infty = \bigcap B_n$ and we denote by \mathbb{E}_n the Φ -preserving conditional expectation from \mathcal{B} onto B_n , $0 \leq n \leq \infty$. Let $Q : \mathcal{B} \rightarrow \mathcal{B}$ be a normal completely positive contraction. Then*

- (a) for $x \in \mathcal{B}$ (i.e. case $p = +\infty$), the sequence $(Q \circ \mathbb{E}_n(x))$ converges to $Q \circ \mathbb{E}_\infty(x)$ almost uniformly;
- (b) for $p \in [2, +\infty]$ and $x \in L^p(\mathcal{B}, \Phi)$, the sequence $(Q \circ \mathbb{E}_n(x))$ converges to $Q \circ \mathbb{E}_\infty(x)$ almost surely;
- (c) for $p \in]1, +\infty]$ and $x \in L^p(\mathcal{B}, \Phi)$, the sequence $(Q \circ \mathbb{E}_n(x))$ converges to $Q \circ \mathbb{E}_\infty(x)$ bilaterally almost surely.

Proof. Replacing x by $x - \mathbb{E}_\infty(x)$, we may assume that $\mathbb{E}_\infty(x) = 0$.

The convergence of $(\mathbb{E}_n(x))$ in the appropriate sense is a result due to Dang-Ngoc [11, Theorem 4] for $p = +\infty$ and to Defant and Junge [12] in the two other cases. The main problem, already appearing in the commutative case, is to show that $(Q \circ \mathbb{E}_n(x))$ still converges in the same sense. The case $p = +\infty$ is immediately solved because Q is continuous on the norm bounded subsets of \mathcal{B} when \mathcal{B} is endowed with the topology of almost uniform convergence (see for instance [37, Proposition 1]).

Let us assume that $p \in]1, +\infty[$ and let us first recall the proof in the commutative case. Thanks to Chebichev inequality, to show that $(Q \circ \mathbb{E}_n(x))$ goes to 0 almost surely, it is enough to show that

$$\lim_m \left\| \sup_{n \geq m} |Q \circ \mathbb{E}_n(x)| \right\|_p = 0. \tag{4.4}$$

But, since Q is a positive contraction, we have

$$\left\| \sup_{n \geq m} |Q \circ \mathbb{E}_n(x)| \right\|_p \leq \left\| Q \left(\sup_{n \geq m} |\mathbb{E}_n(x)| \right) \right\|_p \leq \left\| \sup_{n \geq m} |\mathbb{E}_n(x)| \right\|_p. \tag{4.5}$$

Therefore, it suffices to prove that $\lim_m \left\| \sup_{n \geq m} |\mathbb{E}_n(x)| \right\|_p = 0$. The classical Doob maximal inequality gives, for all $m \in \mathbb{N}$,

$$\left\| \sup_n |\mathbb{E}_n(\mathbb{E}_m(x))| \right\|_p \leq \frac{p}{p-1} \|\mathbb{E}_m(x)\|_p. \tag{4.6}$$

Now, observe that for $n \geq m$, we have $\mathbb{E}_n \circ \mathbb{E}_m = \mathbb{E}_n$, and therefore

$$\left\| \sup_{n \geq m} |\mathbb{E}_n(x)| \right\|_p = \left\| \sup_{n \geq m} |\mathbb{E}_n(\mathbb{E}_m(x))| \right\|_p \leq \frac{p}{p-1} \|\mathbb{E}_m(x)\|_p.$$

Since $(\mathbb{E}_m(x))$ goes to 0 in L^p -norm, this concludes the proof in the commutative case.

When \mathcal{B} is not commutative, the proof follows the same pattern. However, it is not a straightforward generalization and several major difficulties arise. Given a sequence (x_n) in $L^p(\mathcal{B}, \Phi)$ a first problem is to give a meaning to $\sup_n |x_n|$. To this purpose, Junge has introduced in [18] the space $L^p(\mathcal{B}, \ell^\infty)$, derived from Pisier’s theory of vector valued noncommutative L^p -spaces [34]. It is defined as the space of sequences (x_n) in $L^p(\mathcal{B}, \Phi)$ such that there exist $a, b \in L^{2p}(\mathcal{B}, \Phi)$ and $(y_n) \in \ell^\infty(\mathcal{B})$ with $x_n = ay_nb$ for all n . Equipped with the norm

$$\|(x_n)\|_{L^p(\mathcal{B}, \ell^\infty)} = \inf \{ \|a\|_{2p} \sup_n \|y_n\|_\infty \|b\|_{2p} \},$$

where the infimum runs over all the possible decompositions, $L^p(\mathcal{B}, \ell^\infty)$ is a Banach space. As explained for instance in [20], $\|(x_n)\|_{L^p(\mathcal{B}, \ell^\infty)}$ can be viewed as a noncommutative analogue of $\|\sup_n |x_n|\|_p$. The Doob-Junge inequality for noncommutative martingales (see [18]) is expressed as the existence of a constant c_p such that for all $x \in L^p(\mathcal{B}, \Phi)$,

$$\|(\mathbb{E}_n(x))\|_{L^p(\mathcal{B}, \ell^\infty)} \leq c_p \|x\|_p.$$

We also need the two following facts, proved by Defant and Junge in [12] :

- a sequence (x_n) in $L^p(\mathcal{B}, \Phi)$ goes to zero bilaterally almost uniformly if and only if $\lim_m \|(x_n)_{n \geq m}\|_{L^p(\mathcal{B}, \ell^\infty)} = 0$;
- given Q as in the statement of the lemma, we have, for $m \in \mathbb{N}$,

$$\|(Q(x_n))_{n \geq m}\|_{L^p(\mathcal{B}, \ell^\infty)} \leq \|(x_n)_{n \geq m}\|_{L^p(\mathcal{B}, \ell^\infty)}.$$

Having at hand the noncommutative analogues of (4.5) and (4.6), the proof of statement (c) proceeds exactly as in the commutative case. Statement (b) can be obtained similarly. The space $L^p(\mathcal{B}, \ell^\infty)$ has to be replaced by the right hand sided space $L^p(\mathcal{B}, \ell_c^\infty)$, defined as the space of sequences (x_n) in $L^p(\mathcal{B}, \Phi)$ such that there exist $b \in L^p(\mathcal{B}, \Phi)$ and $(y_n) \in \ell^\infty(\mathcal{B})$ with $x_n = y_n b$ for all n . This space is endowed with the norm

$$\|(x_n)\|_{L^p(\mathcal{B}, \ell_c^\infty)} = \inf \{ \sup_n \|y_n\|_\infty \|b\|_p \},$$

where the infimum runs over all the possible decompositions. Now the Doob-Junge one-sided maximal inequality reads as

$$\|(\mathbb{E}_n(x))\|_{L^p(\mathcal{B}, \ell_c^\infty)} \leq \sqrt{c \frac{p}{2}} \|x\|_p.$$

□

Proof of theorem 4.1. We keep some notations introduced in the previous section. As already said, we shall usually view elements $w \in I^n$ as words of length n in the alphabet I and write $w = w_0 w_1 \cdots w_{n-1}$ where w_k is the $(k + 1) - th$ component of w . For $n \geq 1$, we set $I(n) = \{w \in I^n : p_{n-1}(w) \neq 0\}$, with p_n defined in section 3.

We denote by B_n the product von Neumann algebra formed by the maps $b : w \mapsto b_w$ from $I(n + 1)$ into A . This von Neumann algebra will be equipped with the normal faithful state ϕ_n defined by

$$\phi_n(b) = \sum_{w \in I(n+1)} p_n(w) \phi(b_w).$$

Its modular automorphism group $t \mapsto \sigma_t^{\phi_n}$ satisfies

$$\sigma_t^{\phi_n}(b)_w = \sigma_t^\varphi(b_w)$$

for $b \in B_n$ and $w \in I(n + 1)$.

Note that $B_0 = B$ and $\phi_0 = \phi$. We shall need the two following unital injective homomorphisms α_n and β_n from B_{n-1} into B_n , defined, for $n \geq 1$ and $b \in B_{n-1}$ by

$$\alpha_n(b)_w = b_{w_0 \cdots w_{n-1}} \tag{4.7}$$

$$\beta_n(b)_w = \sigma_{w_0}(b_{w_1 \cdots w_n}). \tag{4.8}$$

They satisfy the following relations :

$$\begin{aligned} \alpha_{n+1} \circ \beta_n &= \beta_{n+1} \circ \alpha_n \\ \phi_n \circ \alpha_n &= \phi_{n-1} = \phi_n \circ \beta_n. \end{aligned}$$

Obviously, α_n and β_n intertwine the modular automorphism groups:

$$\alpha_n \circ \sigma_t^{\phi_{n-1}} = \sigma_t^{\phi_n} \circ \alpha_n \quad \text{and} \quad \beta_n \circ \sigma_t^{\phi_{n-1}} = \sigma_t^{\phi_n} \circ \beta_n \tag{4.9}$$

for all $t \in \mathbb{R}$.

It follows that the algebras $\alpha_n(B_{n-1})$ and $\beta_n(B_{n-1})$ are ϕ_n -invariant.

We denote by $(\tilde{\mathcal{B}}, \tilde{\Phi})$ the inductive limit of (B_n, ϕ_n) , with respect to the morphisms $\alpha_n : (B_{n-1}, \phi_{n-1}) \rightarrow (B_n, \phi_n)$, and by λ_n the canonical injective homomorphism from B_n into $\tilde{\mathcal{B}}$. By definition, we have

$$\lambda_{n+1} \circ \alpha_{n+1} = \lambda_n, \tag{4.10}$$

and, using the first equality in (4.9), we see that λ_n is $(\tilde{\Phi}, \phi_n)$ -preserving.

Let β be the injective endomorphism of $\tilde{\mathcal{B}}$ characterized by

$$\lambda_{n+1} \circ \beta_{n+1} = \beta \circ \lambda_n \tag{4.11}$$

for every $n \geq 0$. Note that β is $\tilde{\Phi}$ -preserving.

Next, we set $J_0 = \lambda_0$ and $J_n = \beta^n \circ J_0$ for $n \geq 0$. Then \mathcal{B} is defined to be the von Neumann subalgebra of $\tilde{\mathcal{B}}$ generated by $\bigcup_{n \geq 0} J_n(B)$ and we equip this algebra with the restriction Φ of the state $\tilde{\Phi}$. Finally we denote by $\mathcal{B}_{[n]}$ and $\mathcal{B}_{[n]}$ the von Neumann subalgebras generated by $\bigcup_{k \leq n} J_k(B)$ and $\bigcup_{k \geq n} J_k(B)$ respectively. Obviously, we have $\beta(\mathcal{B}) \subset \mathcal{B}$, and we shall still denote by β the restriction endomorphism $\beta|_{\mathcal{B}} : \mathcal{B} \rightarrow \mathcal{B}$. Since $\sigma_t^\Phi \circ J_n = J_n \circ \sigma_t^\beta$ for all $n \in \mathbb{N}$, we see that the algebras $J_n(B)$ are Φ -invariant, as well as the algebras $\mathcal{B}_{[n]}$ and $\mathcal{B}_{[n]}$.

Computations hereafter are straightforward. They use the following observation: $J_q(b)$, when viewed, for $r \geq q$, as an element of $B_{[r]} \subset \lambda_r(B_r)$ (and therefore, after identification, as an element of B_r) is written

$$J_q(b) = \alpha_r \circ \dots \circ \alpha_{q+1} \circ \beta_q \circ \dots \circ \beta_1(b),$$

hence

$$J_q(b)_{w_0 \dots w_r} = \sigma_{w_0 \dots w_{q-1}}(b_{w_q}). \tag{4.12}$$

Let us first prove that $\mathbb{E}_{[n]} \circ J_q = J_n \circ P^{q-n}$ for $q \geq n$. We have to show that for $c \in B_{[n]}$ and $b \in B$, then

$$\Phi(cJ_q(b)) = \Phi(cJ_n \circ P^{q-n}(b)).$$

We may take c of the form

$$c = J_{k_1}(c^1)J_{k_2}(c^2) \dots J_{k_\ell}(c^\ell)$$

where $c^i \in B$ and $0 \leq k_i \leq n$ for $1 \leq i \leq \ell$. We work in B_q . Thanks to formula (4.12) we get

$$\begin{aligned} \Phi(cJ_q(b)) &= \phi_q(cJ_q(b)) \\ &= \sum_{w \in I(q+1)} p_q(w)\varphi(C(w)\sigma_{w_0 \dots w_{q-1}}(b_{w_q})), \end{aligned}$$

where we have set $C(w) = \sigma_{w_0 \dots w_{k_1-1}}(c^1_{w_{k_1}}) \dots \sigma_{w_0 \dots w_{k_\ell-1}}(c^\ell_{w_{k_\ell}})$.

On the other hand, we have

$$\phi_q(cJ_n \circ P^{q-n}(b)) = \sum_{w \in I(n+1)} p_n(w)\varphi\left(C(w)\sigma_{w_0 \dots w_{n-1}}(P^{q-n}(b)_{w_n})\right), \tag{4.13}$$

and by iteration of formula (4.1),

$$P^{q-n}(b)_{w_n} = \sum_{\substack{v \in I(q-n+1) \\ v_0 = w_n}} p(v_0 v_1) \cdots p(v_{q-n-1} v_{q-n}) \sigma_{v_0 \dots v_{q-n-1}}(b_{v_{q-n}}). \tag{4.14}$$

Writing the above v as $v = w_n \cdots w_q$ and replacing $P^{q-n}(b)_{w_n}$ by its expression (4.14) in (4.13), we immediately get that

$$\phi_q(cJ_q(b)) = \phi_q(cJ_n \circ P^{q-n}(b)).$$

The proof of the relation $\mathbb{E}_{J_n \circ J_0} = J_n \circ (P^*)^n$ is similar. We have to check that

$$\Phi(cJ_0(b)) = \Phi(cJ_n \circ (P^*)^n(b))$$

for every $c \in \mathcal{B}_{J_n}$ and $b \in B$. It is enough to take $c \in \left(\bigcup_{n \leq k \leq r} J_k(B)\right)''$ with $r \geq n$.

We may choose c of the form

$$c = J_{k_1}(c^1) \cdots J_{k_\ell}(c^\ell)$$

where $c^i \in B$ and $n \leq k_i \leq r$ for $1 \leq i \leq \ell$. Again, we get

$$\Phi(cJ_0(b)) = \sum_{w \in I(r+1)} p_r(w)\varphi(C(w)b_{w_0}) \tag{4.15}$$

and

$$\Phi(cJ_n \circ (P^*)^n(b)) = \sum_{w \in I(r+1)} p_r(w)\varphi\left(C(w)\sigma_{w_0 \dots w_{n-1}}((P^*)^n(b)_{w_n})\right),$$

where $C(w) = \sigma_{w_0 \dots w_{k_1-1}}(c^1_{w_{k_1}}) \cdots \sigma_{w_0 \dots w_{k_\ell-1}}(c^\ell_{w_{k_\ell}})$.

Using the invariance of φ by the endomorphisms σ_i , we obtain

$$\Phi(cJ_n \circ (P^*)^n(b)) = \sum_{w \in I(r+1)} p_r(w)\varphi(C'(w_n \cdots w_r)(P^*)^n(b)_{w_n})$$

with $C'(w_n \cdots w_r) = \sigma_{w_n \dots w_{k_1-1}}(c^1_{w_{k_1}}) \cdots \sigma_{w_n \dots w_{k_\ell-1}}(c^\ell_{w_{k_\ell}})$. If we begin by summing on the n first letters of w , we get that $\Phi(cJ_n \circ (P^*)^n(b))$ is equal to

$$\sum_{w_n \cdots w_r \in I(r-n+1)} p_{r-n}(w_n \cdots w_r)\varphi(C'(w_n \cdots w_r)(P^*)^n(b)_{w_n}). \tag{4.16}$$

Furthermore we have

$$(P^*)^n(b)_{w_n} = \frac{1}{p(w_n)} \sum_{\substack{v \in I(n+1) \\ v_n = w_n}} p_n(v) \sigma_{v_{n-1}}^* \circ \dots \circ \sigma_{v_0}^*(b_{v_0}). \tag{4.17}$$

In the expression (4.16) of $\Phi(cJ_n \circ (P^*)^n(b))$ we replace $(P^*)^n(b)_{w_n}$ by the its expression (4.17). It follows from the definition of σ_i^* that $\Phi(cJ_n \circ (P^*)^n(b))$ is equal to

$$\sum_{w \in I(r+1)} p_r(w) \varphi(\sigma_{w_0 \dots w_{k_1-1}}(c_{w_{k_1}}^1) \dots \sigma_{w_0 \dots w_{k_\ell-1}}(c_{w_{k_\ell}}^\ell) b_{w_0}).$$

Comparing with the expression found in (4.15), this concludes the proof of theorem 4.1 □

Remark 4.4. Condition (4.2) is sometimes called the *Markov property* in probability theory. In the classical commutative case, we have, furthermore, the following Markov (or covariance) property (see [35, Chapter 1]):

$$\forall n, q \in \mathbb{N}, \quad \beta^n \circ \mathbb{E}_{q1} = \mathbb{E}_{n+q1} \circ \beta^n. \tag{4.18}$$

Here also, in the setting of theorem 4.1 it can be proved by an easy explicit computation of \mathbb{E}_n that the Markov property (4.18) holds.

5. Proof of the noncommutative Nevo-Stein theorem

Proof. We shall apply the previous section in the particular situation described in remark 3.2. Recall that

$$I = \{-d, -d + 1, \dots, -1, 1, \dots, d - 1, d\},$$

that $[p(ij)]$ is the matrix with $p(ij) = 0$ if $i = -j$ and $p(ij) = \frac{1}{2d - 1}$ otherwise, and that $(p(i))_{i \in I}$ is the uniform probability measure on I , i.e. $p(i) = 1/2d$ for all $i \in I$. Furthermore, $\sigma_i, 1 \leq i \leq d$, is a φ -preserving automorphism of A and $\sigma_{-i} = \sigma_i^{-1}$.

We follow the method used by Bufetov in the commutative case [7]. We introduce $B = A^{2d} = \{(b_i)_{i \in I} : b_i \in A\}$, the state ϕ on B and the completely positive map P as in the previous section, that is

$$P(b)_i = \frac{1}{2d - 1} \sum_{\substack{j \in I \\ j \neq -i}} \sigma_i(b_j).$$

For $x \in A$ we denote by \tilde{x} the element of B such that $\tilde{x}_i = x$ for $i \in I$. Formula (3.1) becomes

$$P^n(\tilde{x})_i = \frac{1}{(2d - 1)^{n-1}} \sum_{\substack{w \in S_n \\ w_0 = i}} \sigma_w(x)$$

for $n \geq 1$, so that

$$s_n(x) = \frac{1}{\#\mathcal{S}_n} \sum_{w \in \mathcal{S}_n} \sigma_w(x) = \frac{1}{2d} \sum_{i \in I} P^n(\tilde{x})_i. \tag{5.1}$$

Let us recall that, for $b \in B$, we have

$$P^*(b)_i = \frac{1}{2d-1} \sum_{\substack{j \in I \\ j \neq i}} \sigma_{-j}(b_j) = \frac{1}{2d-1} \sum_{\substack{j \in I \\ j \neq i}} \sigma_j(b_{-j})$$

since $\sigma_j^* = \sigma_{-j}$ in this setting.

The operators P and P^* are related in the following way. Denoting by U the symmetry of B such that $U(b)_i = \sigma_i(b_{-i})$, we immediately see that $UP^*U = P$. We have

$$P^*P(b)_i = \frac{1}{(2d-1)^2} \sum_{\substack{j \in I \\ j \neq i}} \sum_{\substack{k \in I \\ k \neq j}} b_k$$

and therefore

$$P^*P = \frac{2d-2}{2d-1}UP + \frac{1}{2d-1}Id_B.$$

By induction we get

$$(P^*)^n P^n = \frac{2d-2}{2d-1}UP^{2n-1} + \frac{1}{2d-1}(P^*)^{n-1}P^{n-1}.$$

It follows that

$$\begin{aligned} P^{2n-1} &= \frac{2d-1}{2d-2}U(P^*)^n P^n - \frac{1}{2d-2}U(P^*)^{n-1}P^{n-1} \\ &= \frac{2d-1}{2d-2}P^n(P^*)^n U - \frac{1}{2d-2}P^{n-1}(P^*)^{n-1}U. \end{aligned}$$

Now, using corollary 4.2, we see that for every $b \in B$, the sequences $(P^{2n-1}(b))$ and $(P^{2n}(b))$ converge almost uniformly. As a consequence of the relation (5.1), we get, for $x \in A$, the almost uniform convergence of the sequences $(s_{2n-1}(x))$, $(s_{2n}(x))$ and therefore of $(\frac{1}{n} \sum_{k=0}^{n-1} s_k(x))$.

Let $\mathbb{E}^{(2)}$ be the φ -invariant conditional expectation from A onto the subalgebra $A^{(2)}$ of $\mathbb{F}_d^{(2)}$ -invariant elements in A and let us show that $\mathbb{E}^{(2)}(x)$ is the limit \hat{x} of $(s_{2n}(x))$. Since the automorphisms σ_i are φ -preserving, we have $\varphi(yx) = \varphi(ys_{2n}(x))$ for every $y \in A^{(2)}$. Moreover, the bounded sequence $(s_{2n}(x))$, which converges to \hat{x} almost uniformly, also converges to \hat{x} in the strong topology (see [16, Theorem 1.1.3] for instance). Therefore we have, for $y \in A^{(2)}$,

$$\varphi(y\mathbb{E}^{(2)}(x)) = \varphi(yx) = \lim_{n \rightarrow +\infty} \varphi(ys_{2n}(x)) = \varphi(y\hat{x}).$$

To conclude that $\hat{x} = \mathbb{E}^{(2)}(x)$ it remains to check that \hat{x} is $\mathbb{F}_d^{(2)}$ -invariant. We shall use the following identity:

$$s_1 \circ s_n = \frac{2d - 1}{2d} s_{n+1} + \frac{1}{2d} s_{n-1}. \tag{5.2}$$

This relation is an immediate consequence of the fact that multiplying all words of length 1 by all words of length n , each word of length $n + 1$ is obtained once and each word of length $n - 1$ is obtained $2d - 1$ times. It follows from (5.2) that

$$s_1^2 \circ s_{2n} = \left(\frac{2d - 1}{2d}\right)^2 s_{2n+2} + 2\frac{2d - 1}{(2d)^2} s_{2n} + \frac{1}{(2d)^2} s_{2n-2},$$

from which we get $s_1^2(\hat{x}) = \hat{x}$.

If we view the elements of A as vectors in the Hilbert space $L^2(A, \varphi)$ of the GNS representation associated with φ , a strict convexity argument gives that $\sigma_w(\hat{x}) = \hat{x}$ for every word of length 2. Therefore, \hat{x} is $\mathbb{F}_d^{(2)}$ -invariant.

Similarly, we show that the limit of $\frac{1}{n} \sum_{k=0}^{n-1} s_k(x)$ is the conditional expectation of x with respect to the subalgebra of \mathbb{F}_d -invariant elements.

The proof is similar for any $p > 1$, still using corollary 4.2. □

6. About the noncommutative Rota theorem

As said before, our corollary 4.2 is a noncommutative version (for a particular Markov operator) of the following theorem of G.-C. Rota [36]:

Theorem 6.1 ([36]). *Let P be a μ -preserving Markov operator on the probability measure space (X, μ) . Then for every $p > 1$ and $f \in L^p(X, \mu)$, the sequence $(P^n(P^*)^n(f))$ converges almost everywhere.*

This result was established by Rota from the commutative analogue of theorem 4.1 (see [36]). In fact, Rota’s theorem is still true whenever $f \in L\text{Log}L(X, \mu)$.

The extension of Rota’s theorem to the noncommutative setting (with a minimum of hypotheses on P) is an interesting open problem. It can be studied using the general definition of P^* introduced in (2.2). As a first step, we only consider the case of a state preserving Markov operator, so that P^* satisfies (2.3).

As already explained in Section 4, it is crucial to determine for which operators P theorem 4.1 holds, and we shall examine hereafter this problem.

Let (B, ϕ) be a noncommutative probability space and $P : B \rightarrow B$ be a ϕ -preserving Markov operator. Obviously, if theorem 4.1 holds, then P must be factorizable in the following sense.

Definition 6.2. *Let $P : (A_1, \varphi_1) \rightarrow (A_0, \varphi_0)$ be a (φ_0, φ_1) -preserving Markov operator. We say that P is factorizable if there exists a noncommutative probability space (C, ψ) and two normal unital homomorphisms $j_0 : (A_0, \varphi_0) \rightarrow (C, \psi)$, $j_1 : (A_1, \varphi_1) \rightarrow (C, \psi)$ such that (ψ, φ_0) and (ψ, φ_1) are respectively j_0 - and j_1 -stationary and $P = j_0^* \circ j_1$.*

Indeed, if theorem 4.1 holds, take $(C, \psi) = (\mathcal{B}, \Phi)$ and $j_i = J_i, i = 0, 1$. Since $J_0^* = J_0^{-1} \circ \mathbb{E}_{01}$, the equality $\mathbb{E}_{01} \circ J_1 = J_0 \circ P$ gives $P = J_0^* \circ j_1$.

Examples 6.3. (a) If A_0, A_1 are commutative abelian von Neumann algebras, every unital positive map $P : A_1 \rightarrow A_0$ is factorizable. Indeed, let us endow the algebraic tensor product $A_0 \odot A_1$ with the inner product

$$\langle f_0 \otimes f_1, g_0 \otimes g_1 \rangle = \varphi_0(g_0^* f_0 P(g_1^* f_1))$$

and denote by H_P the Hilbert space obtained by separation and completion. Let j_0 (resp. j_1) be the normal representation of A_0 (resp. A_1) defined by

$$j_0(g_0)(f_0 \otimes f_1) = g_0 f_0 \otimes f_1 \quad \text{for } g_0, f_0 \in A_0, f_1 \in A_1$$

(resp.

$$j_1(g_1)(f_0 \otimes f_1) = f_0 \otimes g_1 f_1 \quad \text{for } f_0 \in A_0, g_1, f_1 \in A_1).$$

We denote by C the von Neumann algebra generated by $j_0(A_0)$ and $j_1(A_1)$ and by ψ the state $c \mapsto \langle c \xi_P, \xi_P \rangle$ where ξ_P is the class of $1 \otimes 1$ in H_P . In particular, we have $\psi(j_0(f_0)j_1(f_1)) = \varphi(f_0 P(f_1))$. We easily check that $P = J_0^* \circ j_1$.

The reader familiar with Connes' notion of correspondence will recognize in H_P the Hilbert space of the correspondence associated with P . This construction is also well known in probability theory. In this framework, to (A_i, φ_i) is associated a probability space (X_i, μ_i) such that $A_i = L^\infty(X_i, \mu_i)$ and $\varphi_i(f) = \int_{X_i} f d\mu_i$ for $f \in A_i, i = 0, 1$. Let p be the probability transition associated with P , i.e. $P(f)(x) = \int_{X_1} f(y)p(x, dy)$ for $f \in L^\infty(X_1, \mu_1)$. Then $C = L^\infty(X_0 \times X_1, \nu)$ where ν is the probability measure defined by

$$\int_{X_0 \times X_1} f d\nu = \int_{X_0} \left(\int_{X_1} f(x_0, x_1)p(x_0, dx_1) \right) d\mu_0(x_0).$$

Here, ψ is the state associated with ν and the embeddings j_0, j_1 are the obvious ones.

- (b) Let us come back to the setting $P : (B, \phi) \rightarrow (B, \phi)$ considered in section 4. We take for (C, ψ) the algebra $B_1 = \{b : I(2) \rightarrow A\}$ equipped with the state ϕ_1 . Recall that $\phi_1(b) = \sum_{w \in I(2)} p(w_0)p(w_0 w_1)\varphi(b_{w_0 w_1})$ for $b \in B_1$. Let j_0 be the map $\alpha_1 : B \rightarrow B_1$ defined by $\alpha_1(b)_{w_0 w_1} = b_{w_0}$ and let j_1 be the map $\beta_1 : B \rightarrow B_1$ defined by $\beta_1(b)_{w_0 w_1} = \sigma_{w_0}(b_{w_1})$. Then it is easily checked that $P = \alpha_1^* \circ \beta_1$.
- (c) More generally, let us replace the endomorphisms σ_i of A by any factorizable φ -preserving completely positive map P_i from A to A for $i \in I$ and let us keep the notations of (b) for (B, ϕ) . Then the map $P : (B, \phi) \rightarrow (B, \phi)$ defined by

$$P(b)_i = \sum_{j \in I} p(ij)P_i(b_j)$$

is still factorizable. Indeed, let us write $P_i = \alpha_i^* \circ \beta_i$ for α_i, β_i appropriate maps from (A, φ) into a noncommutative probability space (C_i, ψ_i) . We define C to be the von Neumann algebra formed of the maps $b : w_0 w_1 \in I(2) \mapsto b_{w_0 w_1} \in C_{w_0}$. We endow C with the state

$$\psi : b \mapsto \sum_{w \in I(2)} p(w_0) p(w_0 w_1) \psi_{w_0}(b_{w_0 w_1}).$$

Moreover we define α, β from $B = A^I$ into C by

$$\alpha(b)_{w_0 w_1} = \alpha_{w_0}(b_{w_0}), \quad \beta(b)_{w_0 w_1} = \beta_{w_0}(b_{w_1}).$$

Then we have $P = \alpha^* \circ \beta$.

(d) Another useful example is given by the following lemma.

Before stating it, we make the convention hereafter, for simplicity of language, that all our maps between noncommutative probability spaces will (implicitly) be preserving Markov operators with respect to the given states.

Lemma 6.4. *Let $(B, \varphi), (A_i, \varphi_i) i = 0, 1$, be non commutative probability spaces. We assume that for $i = 0, 1, A_i$ contains a copy of B as a von Neumann subalgebra and that there is a conditional expectation $\mathbb{E}_i : A_i \rightarrow B$ with $\varphi_i = \varphi \circ \mathbb{E}_i$. Then $P = i_0 \circ i_1^*$ (where $i_k, k = 1, 0$, is the canonical embedding of B into A_k) is factorizable. More precisely, one can find (C, ψ) and homomorphisms j_0, j_1 from $(A_0, \varphi_0), (A_1, \varphi_1)$ respectively, into (C, ψ) such that $j_0 \circ i_0 = j_1 \circ i_1$ and $i_0 \circ i_1^* = j_0^* \circ j_1$.*

Proof. We define (C, ψ) to be the reduced amalgamated free product of (A_1, φ_1) by (A_0, φ_0) over B and $j_0 : A_0 \rightarrow C, j_1 : A_1 \rightarrow C$ are the canonical embeddings. This construction is explained for instance in [4]. We give below an equivalent construction, in terms of self-dual Hilbert modules (see [33] for that notion). It is the exact analogue of the construction exposed by Voiculescu in [44, Section 5], except that self-dual completions replace norm completions in the definitions (e. g. for relative tensor products). For $i = 0, 1$, we denote by H_i the self-dual right Hilbert B -module obtained by completion of A_i with respect to the B -valued inner product $\langle x, y \rangle = \mathbb{E}_i(x^* y)$ for $x, y \in A_i$ and $x \mapsto \hat{x}$ is the canonical embedding of A_i into H_i . Recall that left multiplication on A_i yields a unital normal injective homomorphism π_i from A_i into $\mathcal{B}_B(H_i)$ (the von Neumann algebra of B -linear operators having an adjoint). In particular H_i is a bimodule over B . We set $\xi_i = \widehat{1}_{A_i}$ and decompose H_i as the orthogonal direct sum of self-dual Hilbert B -modules $H_i = \xi_i B \oplus \overset{\circ}{H}_i$. As in [44] we define $(H, \xi) = (H_0, \xi_0) * (H_1, \xi_1)$ by

$$H = \xi B \oplus_{n \geq 1} \bigoplus_{\substack{\iota_1, \dots, \iota_n \in \{0, 1\} \\ \iota_1 \neq \iota_2 \neq \iota_3 \neq \dots \neq \iota_{n-1} \neq \iota_n}} \overset{\circ}{H}_{\iota_1} \otimes_B \dots \otimes_B \overset{\circ}{H}_{\iota_n}.$$

We denote by j_i the normal representation of A_i into $\mathcal{B}_B(H)$ introduced by Voiculescu (written λ_i in [44]) and $C = A_0 *_B A_1$ is the von Neumann subalgebra

of $\mathcal{B}_B(H)$ generated by $j_0(A_0)$ and $j_1(A_1)$. We shall freely identify the elements $a \in A_i$ and $j_i(a) \in C, i = 0, 1$.

The map $\mathbb{E} : c \mapsto \langle \xi, c\xi \rangle$ is a normal conditional expectation from C onto B and we set $\psi = \varphi \circ \mathbb{E}$. We identify the submodule $\xi B \oplus \overset{o}{H}_0$ of H with H_0 and we denote by q_0 the orthogonal projection from H onto H_0 . Then $\tilde{\mathbb{E}}_0 : c \mapsto q_0 c q_0$ defines a normal conditional expectation from C onto $j_0(A_0)$ (via the identification $x \equiv j_0 \circ \pi_0^{-1}(x)$ between $\pi_0(A_0)$ and $j_0(A_0)$) such that $\tilde{\mathbb{E}}_0 \circ j_1(a) = \mathbb{E}_1(a)$ for $a \in A_1$ (see [4, Lemma 3.5]). More precisely, if we avoid obvious identifications, we have to write $j_0^{-1} \circ \tilde{\mathbb{E}}_0 \circ j_1(a) = i_0 \circ i_1^{-1} \circ \mathbb{E}_1(a)$. We have $\mathbb{E} = \mathbb{E}_0 \circ \tilde{\mathbb{E}}_0$ and therefore $\varphi_0 \circ \tilde{\mathbb{E}}_0 = \psi$. Similarly we define $\tilde{\mathbb{E}}_1 : C \rightarrow j_1(A_1)$.

Finally, we observe that the equality $j_0^{-1} \circ \tilde{\mathbb{E}}_0 \circ j_1 = i_0 \circ i_1^{-1} \circ \mathbb{E}_1$ is nothing else than $j_0^* \circ j_1 = i_0 \circ i_1^*$. □

Note that C is generated, as a von Neumann algebra, by $j_0(A_0)$ and $j_1(A_1)$.

Theorem 6.5. *A ϕ -preserving Markov operator P on (B, ϕ) is factorizable if and only if theorem 4.1 holds. As a consequence, the noncommutative Rota theorem 4.2 applies to factorizable Markov operators.*

In fact, we shall prove a more complete theorem.

Theorem 6.6 (Noncommutative Markov chain construction). *Let $P : (B, \phi) \rightarrow (B, \phi)$ be a factorizable ϕ -preserving Markov operator. There exist*

- a von Neumann algebra \mathcal{B} ,
- a normal faithful state Φ on \mathcal{B} ,
- a normal unital endomorphism $\beta : \mathcal{B} \rightarrow \mathcal{B}$ with $\Phi \circ \beta = \Phi$ and $\sigma_t^\Phi \circ \beta = \beta \circ \sigma_t^\Phi$ for $t \in \mathbb{R}$,
- a normal unital homomorphism $J_0 : \mathcal{B} \rightarrow \mathcal{B}$ with $\Phi \circ J_0 = \phi$ and $\sigma_t^\Phi \circ J_0 = J_0 \circ \sigma_t^\Phi$ for $t \in \mathbb{R}$,

such that, if we set $J_n = \beta^n \circ J_0$ for $n \geq 0$, and if $\mathcal{B}_{[n]}$ and $\mathcal{B}_{\geq n}$ denotes the von Neumann subalgebras of \mathcal{B} generated by $\bigcup_{k \leq n} J_k(B)$ and $\bigcup_{k \geq n} J_k(B)$ respectively, then

- (i) the algebras $\mathcal{B}_{[n]}$ and $\mathcal{B}_{\geq n}$ are Φ -invariant;
- (ii) if $\mathbb{E}_{[n]}$ and $\mathbb{E}_{\geq n}$ are the corresponding Φ -preserving conditional expectations, we have

$$\mathbb{E}_{[n]} \circ J_q = J_n \circ P^{q-n}, \quad \forall q \geq n \geq 0, \tag{6.1}$$

$$\mathbb{E}_{[n+q]} \circ \beta^q = \beta^q \circ \mathbb{E}_{[n]}, \quad \forall n, q \in \mathbb{N}, \tag{6.2}$$

$$\mathbb{E}_{[n]} \circ J_0 = J_n \circ (P^*)^n, \quad \forall n \in \mathbb{N}. \tag{6.3}$$

Proof. Assume the existence of (B_1, ϕ_1) and of homomorphisms α_1, β_1 from (B, ϕ) into (B_1, ϕ_1) with $P = \alpha_1^* \circ \beta_1$. Let us set $(B_0, \phi_0) = (B, \phi)$. Using lemma 6.4, we construct inductively a sequence (B_n, ϕ_n) of noncommutative probability spaces, and of pairs $(\alpha_{n+1}, \beta_{n+1})$ of homomorphisms from (B_n, ϕ_n) into (B_{n+1}, ϕ_{n+1}) such that

$$\beta_n \circ \alpha_n^* = \alpha_{n+1}^* \circ \beta_{n+1} \tag{6.4}$$

$$\alpha_{n+1} \circ \beta_n = \beta_{n+1} \circ \alpha_n. \tag{6.5}$$

Let (\mathcal{B}, Φ) be the inductive limit of (B_n, ϕ_n) with respect to the morphisms $\alpha_n : (B_{n-1}, \phi_{n-1}) \rightarrow (B_n, \phi_n)$ and let λ_n be the canonical embedding from (B_n, ϕ_n) into (\mathcal{B}, Φ) . We also define the endomorphism β of (\mathcal{B}, Φ) , characterized by the condition

$$\lambda_{n+1} \circ \beta_{n+1} = \beta \circ \lambda_n \quad \text{for all } n \geq 0.$$

Now we set $J_0 = \lambda_0 : B \rightarrow \mathcal{B}$ and $J_n = \beta^n \circ J_0$ for $n \geq 0$.

We shall identify B_n with the subalgebra $\lambda_n(B_n)$ of \mathcal{B} when convenient. Then, for $0 \leq q \leq r$, the homomorphism J_q , when viewed as a homomorphism from B into $B_r \subset \mathcal{B}$ is written

$$\begin{aligned} J_q &= \alpha_r \circ \cdots \circ \alpha_{q+1} \circ \beta_q \circ \cdots \circ \beta_1 \\ &= \beta_r \circ \cdots \circ \beta_{r-q+1} \circ \alpha_{r-q} \circ \cdots \circ \alpha_1, \end{aligned}$$

if $q > 0$. For $q = 0$,

$$J_0 = \alpha_r \circ \cdots \circ \alpha_1.$$

We shall also need the following formula:

$$P^n = \alpha_1^* \circ \cdots \circ \alpha_n^* \circ \beta_n \circ \cdots \circ \beta_1. \quad (6.6)$$

Starting with $P = \alpha_1^* \circ \beta_1$, this formula is proved by using (6.4) repeatedly.

Since for every $n \geq 1$, the von Neumann algebra B_n is generated by $\alpha_n(B_{n-1})$ and $\beta_n(B_{n-1})$, we see inductively that

$$B_n \equiv \lambda_n(B_n) = B_{[n]}.$$

Observe that for $q \geq n \geq 0$, the conditional expectation $\mathbb{E}_{[n]}$, restricted to B_q is $\alpha_{n+1}^* \circ \cdots \circ \alpha_q^*$. It follows that

$$\mathbb{E}_{[n]} \circ J_q = \alpha_{n+1}^* \circ \cdots \circ \alpha_q^* \circ \beta_q \circ \cdots \circ \beta_1,$$

and, using (6.4), we get

$$\begin{aligned} \mathbb{E}_{[n]} \circ J_q &= \beta_n \circ \cdots \circ \beta_1 \circ \alpha_1^* \circ \cdots \circ \alpha_{q-n}^* \circ \beta_{q-n} \circ \cdots \circ \beta_1 \\ &= J_n \circ P^{q-n}, \end{aligned}$$

by (6.6). This concludes the proof of (6.1).

Let us show now that $\mathbb{E}_{[n+q]} \circ \beta^q = \beta^q \circ \mathbb{E}_{[n]}$ for every $n, q \in \mathbb{N}$. It is enough to fix an integer $r \geq n$ and to prove this formula in B_r . Therefore we have to check that

$$\alpha_{r+q+1}^* \circ \cdots \circ \alpha_{r+q}^* \circ \beta_{r+q} \circ \cdots \circ \beta_{r+1} = \beta_{n+q} \circ \cdots \circ \beta_{n+1} \circ \alpha_{n+1}^* \circ \cdots \circ \alpha_r^*,$$

which is still a consequence of (6.4).

Finally, let us prove that $\mathbb{E}_{[n]} \circ J_0 = J_n \circ (P^*)^n$. We have to check that for $b \in B$, $r \geq n$ and $c \in \left(\bigcup_{n \leq k \leq r} J_k(B) \right)''$ then

$$\Phi(cJ_0(b)) = \Phi(cJ_n \circ (P^*)^n(b)).$$

In order to shorten notations, for any pair of integers $k \leq \ell$, we shall write $\alpha_{[\ell,k]}$ for $\alpha_\ell \circ \dots \circ \alpha_k$, and we shall adopt the same convention for $\beta_{[\ell,k]}$. We denote by $B_{[n,r]}$ the algebra $\left(\bigcup_{n \leq k \leq r} J_k(B)\right)''$ when viewed into B_r . One shows immediately that for every $1 \leq k \leq r$,

$$B_{[r-k,r]} = \beta_{[r,k+1]}(B_k).$$

We have to prove that for $n \leq r$,

$$\mathbb{E}_{\beta_{[r,r-n+1]}(B_{r-n})}^{B_r} \circ \alpha_{[r,1]} = \alpha_{[r,n+1]} \circ \beta_{[n,1]} \circ (P^*)^n, \tag{6.7}$$

where $\mathbb{E}_{\beta_{[r,r-n+1]}(B_{r-n})}^{B_r}$ is the natural conditional expectation from B_r onto $\beta_{[r,r-n+1]}(B_{r-n}) = B_{[n,r]}$. We fix n and proceed by induction on r .

First, for $r = n$, let us prove that $\mathbb{E}_{\beta_{[n,1]}(B)}^{B_n} \circ \alpha_{[n,1]} = \beta_{[n,1]} \circ (P^*)^n$: by (6.6) we have $(P^*)^n = \beta_{[n,1]}^* \circ \alpha_{[n,1]}$ and therefore

$$\beta_{[n,1]} \circ (P^*)^n = \beta_{[n,1]} \circ \beta_{[n,1]}^* \circ \alpha_{[n,1]} = \mathbb{E}_{\beta_{[n,1]}(B)}^{B_n} \circ \alpha_{[n,1]}.$$

Assume now that (6.7) is true and let us prove the formula for $r + 1$. Take $c \in B_{r+1-n}$ and $b \in B$ and compute

$$\phi_{r+1}(\beta_{[r+1,r-n+2]}(c)\alpha_{[r+1,1]}(b)) = \phi_r(\alpha_{r+1}^* \circ \beta_{[r+1,r-n+2]}(c)\alpha_{[r,1]}(b)).$$

Using again (6.4), we get

$$\alpha_{r+1}^* \circ \beta_{[r+1,r-n+2]}(c) = \beta_{[r,r-n+1]} \circ \alpha_{r-n+1}^*(c).$$

We have $\alpha_{r-n+1}^*(c) \in B_{r-n}$ and therefore the assumption (6.7) gives

$$\begin{aligned} &\phi_{r+1}(\beta_{[r+1,r-n+2]}(c)\alpha_{[r+1,1]}(b)) \\ &= \phi_r(\alpha_{r+1}^* \circ \beta_{[r+1,r-n+2]}(c)\alpha_{[r,1]}(b)) \\ &= \phi_r(\alpha_{r+1}^* \circ \beta_{[r+1,r-n+2]}(c)\mathbb{E}_{\beta_{[r,r-n+1]}(B_{r-n})}^{B_r} \circ \alpha_{[r,1]}(b)) \\ &= \phi_r(\alpha_{r+1}^* \circ \beta_{[r+1,r-n+2]}(c)\alpha_{[r,n+1]} \circ \beta_{[n,1]} \circ (P^*)^n(b)) \\ &= \phi_{r+1}(\beta_{[r+1,r-n+2]}(c)\alpha_{r+1} \circ \alpha_{[r,n+1]} \circ \beta_{[n,1]} \circ (P^*)^n(b)). \end{aligned}$$

It follows that

$$\mathbb{E}_{\beta_{[r+1,r-n+2]}(B_{r+1-n})}^{B_{r+1}} \circ \alpha_{[r+1,1]}(b) = \alpha_{[r+1,n+1]} \circ \beta_{[n,1]} \circ (P^*)^n(b).$$

This concludes the proof of theorem 6.5. □

Remark 6.7. (a) Let us consider the commutative case (see (6.3) (a)). Given a μ -preserving Markov operator P from $L^\infty(X, \mu)$ into itself, there is a much simpler construction of $(\mathcal{B}, \Phi, \beta, J_n)$ satisfying theorem 6.6, known as the Daniell-Kolmogorov construction. The von Neumann \mathcal{B} is $L^\infty(\Omega, \nu)$, where $\Omega = X^{\mathbb{N}}$ is the trajectory space equipped with the usual σ -field of measurable subsets, and ν is the Markov measure associated with P and the starting measure μ . Let $(X_n)_{n \in \mathbb{N}}$ be the corresponding homogeneous Markov chain, that is, for $n \in \mathbb{N}$ the random variable $X_n : \Omega \rightarrow X$ is the projection $(x_i)_{i \in \mathbb{N}} \mapsto x_n$. Then J_n is the homomorphism $f \mapsto f \circ X_n$ from $L^\infty(X, \mu)$ into $L^\infty(\Omega, \nu)$.

- (b) When A is the von Neumann algebra \mathbb{C} , the construction made in section 4 is a particular case of the above one. It corresponds to the case where $X = I$, μ is the distribution $(p(i))_{i \in I}$ and $P : \mathbb{C}^I \rightarrow \mathbb{C}^I$ is given by the transition matrix $[p(ij)]$.
- (c) Given any unital completely positive map $P : B \rightarrow B$, where B is a von Neumann algebra (or a unital C^* -algebra), without any other assumption on P , several Daniell-Kolmogorov type constructions can be found in the literature (see [37] and [3] for example). In these constructions, equalities (6.1) and (6.2) are fulfilled. The main point here is that we also need condition (6.3).

- Problems.** (1) Are there examples of ϕ -preserving Markov operators not being factorizable?
- (2) For a factorizable Markov operator P , find a canonical construction for $(\mathcal{B}, \Phi, \beta, J_n)$ in theorem 6.6.
- (3) Let P be a normal unital completely positive map from (B, ϕ) onto itself such that $\phi \circ P = \phi$. Let P^* be the completely positive operator such that $\phi(P^*(b)\sigma_{-i/2}^\phi(a)) = \phi(\sigma_{i/2}^\phi(b)P(a))$ for all analytic elements $a, b \in B$. Study the convergence of the sequence $(P^n(P^*)^n)$.

Remark 6.8. Junge and Xu have generalized in [20] both the individual Dunford-Schwartz ergodic theorem and the classical Stein ergodic theorem proved in [40]. Using our terminology, the noncommutative Stein ergodic theorem reads as follows:

Theorem 6.9 (Junge-Xu, Corollaries 7.11 and 7.13, [20]). *Let (B, ϕ) be a noncommutative probability space and let $P : B \rightarrow B$ be a ϕ -preserving Markov operator such that $P = P^*$. Then for $p \in]1, +\infty]$ and $x \in L^p(B, \phi)$, the sequence $(P^{2n}(x))$ converges bilaterally almost surely to the canonical projection of x onto the subspace of P -invariant elements. If $p > 2$, the convergence also holds almost surely.*

This gives of course the noncommutative Rota theorem for self-adjoint ϕ -preserving Markov operators.

References

1. Accardi, L., Cecchini, C.: Conditional expectations in von Neumann algebras and a theorem of Takesaki. *J. Funct. Anal.* **45**, 245-273 (1982)
2. Arnol'd, V.I., Krylov, A.L.: Uniform distribution of points on a sphere and certain ergodic properties of solutions of linear ordinary differential equations in a complex domain. *Dokl. Akad. Nauk. SSSR* **148**, 9-12 (1963)
3. Bhat, B.V.R., Parthasarathy, K.R.: Markov dilations of nonconservative dynamical semi-groups and a quantum boundary theory. *Ann. Inst. Henri Poincaré* **31**, 601-651 (1995)
4. Blanchard, E., Dykema, K.: Embeddings of reduced free products of operator algebras. *Pacific J. Math.* **199**, 1-19 (2001)
5. Bratteli, O., Robinson, D.W.: *Operator algebras and quantum statistical mechanics I. Texts and Monographs in Physics*, Springer-Verlag, 1979
6. Bufetov, A.I.: Operator ergodic theorems for actions of free semigroups and groups. *Funct. Anal. Appl.* **34**, 239-251 (2000)

7. Bufetov, A.I.: Convergence of spherical averages for actions of free groups. *Annals of Math.* **155**, 929–944 (2002)
8. Chilin, V.I., Litvinov, S., Skalski, A.: A few new results in noncommutative ergodic theory. To appear in *J. Operator Theory*
9. Connes, A., Narnhofer, H., Thirring, W.: Dynamical entropy of C^* -algebras and von Neumann algebras. *Commun. Math. Phys.* **112**, 691–817 (1987)
10. Conze, J.P., Dang Ngoc, N.: Ergodic theorems for non-commutative dynamical systems. *Invent. Math.* **46**, 1–15 (1978)
11. Dang-Ngoc, N.: Pointwise convergence of martingales in von Neumann algebras. *Israel J. Math.* **34**, 273–280 (1979)
12. Defant, A., Junge, M.: Classical summation methods in noncommutative probability. In preparation
13. Goldstein, S., Lindsay, J.M.: KMS -symmetric Markov semigroups. *Math. Zeit.* **219**, 591–608 (1995)
14. Grigorchuk, R.I.: An individual ergodic theorems for actions of free groups. Proc. XII Workshop in the theory of operators in functional spaces. Tambov, 1987
15. Guivarc’h, Y.: Généralisation d’un théorème de von Neumann. *C. R. Acad. Sci. Paris, Ser. A* **268**, 1020–1023 (1969)
16. Jajte, R.: Strong limit theorems in non-commutative probability. *Lecture Notes in Math.* **1110**, Springer-Verlag, 1985
17. Jajte, R.: Strong limit theorems in non-commutative L_2 -spaces. *Lecture Notes in Math.* **1477**, Springer-Verlag, 1991
18. Junge, M.: Doob’s inequality for non-commutative martingales. *J. Reine Angew. Math.* **549**, 149–190 (2002)
19. Junge, M., Xu, Q.: Théorèmes ergodiques maximaux dans les espaces L^p non commutatifs. *C. R. Acad. Sci. Paris, Ser. I* **334**, 773–778 (2002)
20. Junge, M., Xu, Q.: Noncommutative maximal ergodic theorems, Preprint 2004
21. Junge, M., Xu, Q.: Non-commutative Burkholder/Rosenthal inequalities. Preprint
22. Junge, M., Xu, Q.: Haagerup’s reduction on noncommutative L_p -spaces and applications. In preparation
23. Haagerup, U.: L^p -spaces associated with an arbitrary von Neumann algebra. In: “Algèbres d’opérateurs et leurs applications en physique mathématique”. 175–184, Colloque Internat. CNRS, Marseille 1977
24. Kadison, R.V., Ringrose, J.R.: Fundamentals of the theory of operator algebras, Volume II. Graduate Studies in Math. **16**, AMS, 1997
25. Kaimanovich, V.: Measure-theoretic boundaries of Markov chains, 0 – 2 laws and entropy. In: Picardello, M.A., (ed) Proceedings of the Conference on Harmonic Analysis and Discrete Potential Theory. (Frascati 1991), Plenum 1992, pp 145–180
26. Kosaki, H.: Applications of the complex interpolation method to a von Neumann algebra: non-commutative L^p -spaces. *J. Funct. Anal.* **56**, 29–78 (1984)
27. Kümmerer, B.: A non-commutative individual theorem. *Invent. Math.* **46**, 139–145 (1978)
28. Lance, E.C.: Ergodic theorems for convex sets and operator algebras. *Invent. Math.* **37**, 201–214 (1975)
29. Nelson, E.: Notes on non-commutative integration. *J. Funct. Anal.* **15**, 29–78 (1974)
30. Neveu, J.: Bases mathématiques du calcul des probabilités. Masson, 1970
31. Nevo, A.: Harmonic analysis and pointwise ergodic theorems for noncommuting transformations. *J. Amer. Math. Soc.* **7**, (1994) 875–902
32. Nevo, A., Stein, E.M.: A generalisation of Birkoff’s pointwise ergodic theorem. *Acta Math.* **173**, 135–154 (1994)

33. Paschke, W.L.: Inner product modules over B^* -algebras. *Trans. Amer. Math. Soc.* **182**, 443–468 (1973)
34. Pisier, G.: Non-commutative vector valued L_p -spaces and completely p -summing maps. *Astérisque*, **247**, 1998
35. Revuz, D.: Markov chains. North-Holland, 1984
36. Rota, G.-C.: An “alternierende verfahren” for general positive operators. *Bull. A. M. S.* **68**, 95–102 (1962)
37. Sauvageot, J.-L.: A note on almost uniform convergence in von Neumann algebras. In: “Quantum probability and related topics”. *Word Sc. Publ.*, 1991, pp 385–390
38. Sauvageot, J.-L.: Markov quantum semigroups admit covariant Markov C^* -dilations. *Commun. Math. Phys.* **106**, 91–103 (1986)
39. Segal, I.E.: A noncommutative extension of abstract integration. *Annals of Math.* **57**, 401–457, 595–596 (1953)
40. Stein, E.M.: On the maximal ergodic theorem. *Proc. Nat. Acad. Sci. U.S.A.* **47**, 1894–1897 (1961)
41. Takesaki, M.: Conditional expectations in von Neumann algebras. *J. Funct. Anal.* **9**, 306–321 (1972)
42. Takesaki, M.: Theory of operator algebras II. *Encyclopaedia of mathematical sciences*, **125**, Operator algebras and non-commutative geometry. **6**, Springer-Verlag, Berlin, 2003
43. Terp, M.: Interpolation between a von Neumann algebra and its predual. *J. Operator Theory* **8**, 327–360 (1982)
44. Voiculescu, D.: Symmetries of some reduced free product C^* -algebras. In: “Operator algebras and their connections with topology and ergodic theory”. 556–588, *Lecture Notes in Math.* **1132**, Springer-Verlag, 1985
45. Walker, T.E.: Ergodic theorems for free group actions on von Neumann algebras. *J. Funct. Anal.* **150**, 27–47 (1997)
46. Yeadon, F.J.: Ergodic theorems for semi-finite von Neumann algebras I. *J. London Math. Soc.* **16**, 326–332 (1977)