ISRAEL JOURNAL OF MATHEMATICS **138** (2003), 93–123

NONSINGULAR DYNAMICAL SYSTEMS, BRATTELI DIAGRAMS AND MARKOV ODOMETERS

BY

A. H. DOOLEY

School of Mathematics, University of New South Wales
Sydney 2052, Australia
e-mail: a.dooley@unsw.edu.au

AND

TOSHIHIRO HAMACHI

Graduate School of Mathematics, Kyushu University
Ropponmatsu, Chuo-Ku, Fukuoka 810, Japan
e-mail: hamachi@math.kyushu-u.ac.jp

ABSTRACT

We prove that any ergodic non-singular transformation is orbit equivalent to a Markov odometer which is uniquely ergodic.

1. Introduction

It was proved by Henry Dye [D1], [D2] that every ergodic measure-preserving dynamical system $(X, \mathcal{B}, T, \mu)$ on a standard measure space $(X, \mathcal{B}, \mu)$ is orbit equivalent[1] to the odometer acting on an infinite product measure space. More recently, Vershik, Livshitz, Lodkin and other collaborators ([VL], [LV]) have shown

1 $(X, \mathcal{B}, T, \mu)$ is **orbit equivalent** to $(Y, \mathcal{C}, S, \nu)$ if there is a bi-measurable mapping $\Phi: X \to Y$ with $\mu \circ \Phi^{-1}$ and ν mutually absolutely continuous, and which maps almost all T-orbits to S-orbits, i.e., for almost all $x \in X$

$$\{\Phi T^n x : n \in \mathbb{Z}\} = \{S^n \Phi x : n \in \mathbb{Z}\}.$$

This is the natural notion of equivalence between measurable dynamical systems.

Received August 1, 2002

that the orbit equivalence may be achieved by a transversal mapping and in such a way that the odometer action is an adic action on a Markov compactum, which is moreover uniquely ergodic. In fact, Vershik's Markov compactum is the same thing in different language as the Bratteli diagram used in operator algebras ([B], [HPS]). The Vershik transformation is a natural generalisation of the odometer action.

Krieger noticed that Dye's theorem extends to the non-singular case, where we no longer have $\mu = T\mu$, but rather $\mu \sim T\mu$. In this case, μ is defined to be ergodic if every T-invariant set is either null or co-null. In fact, Connes, Feldman and Weiss ([CFW], [OW1]) also showed that the theorem is true when T is replaced by the action of any discrete amenable group.

The original classification of ergodic dynamical systems up to orbit equivalence, into type I (atomic spaces), type II (nonatomic spaces where there is an equivalent measure which is invariant under T) and type III (nonatomic spaces where there is no equivalent invariant measure) was later refined by Krieger and Araki-Woods, who introduced the ratio set. This allows one to further classify the type III systems into type III_λ for $0 \leq \lambda \leq 1$.

Krieger's Theorem ([Kr1], [Kr2]) then asserts that, up to orbit equivalence, there is a unique system of type III_λ for $0 < \lambda \leq 1$. Since one may exhibit a product odometer of each of these types, and since product actions are a fortiori Markovian in the sense of Vershik, one then sees for all actions, except for type III_0 actions, that some analogue of Vershik's theorem holds, although the questions of unique ergodicity have not, to the best of our knowledge, so far been addressed.

In this paper, we will demonstrate an analogue of Vershik's theorem for all non-singular systems. Our main theorem states:

THEOREM 1.1: *Every ergodic non-singular dynamical system $(X, \mathcal{B}, T, \mu)$ on a standard measure space*[2] *is orbit equivalent to a Markov odometer. Furthermore, when considered as a G-measure, the Markov odometer may be taken to be uniquely ergodic.*

In order to make precise the meaning of this theorem, we will need to define a Markov odometer on a Bratteli–Vershik diagram (section 2), and the notion of unique ergodicity (section 3). This is a weighted version of the traditional unique ergodicity (see Lemma 3.1). In the proof of our main theorem, a key technical

2 A measure space is **standard** if $(X, \mathcal{B})$ is a standard Borel space and μ is a σ-finite measure on X. All measurable spaces will be assumed standard, and all measures σ-finite.

idea is the notion of refinement of multiple ordered towers (see [Ha]), and a study of how this affects measures under which the towers are of constant Jacobian. This material is explained in sections 4 and 5, and we believe that it might have some independent interest.

In particular, our theorem implies that every system of type III_0 is orbit equivalent to a Markov odometer. The study of type III transformation systems has a long history. For a beautiful recent survey and simplification of the classification theorem, see [KW]. It is known that there exist type III_0 systems which are not orbit equivalent to products ([Kr], [CW]), although there have been no explicit descriptions of Markov odometers which are not orbit equivalent to products. In a recent paper, the authors construct such an example ([DH1]).

The unique ergodicity is a rather strong property. It can be shown that this can fail rather dramatically for general Markov measures ([DR]).

These results generalise to the non-singular case the well-known Jewett–Krieger theorem, which states that every finite measure-preserving system may be realised up to measure-theoretic conjugacy by a Cantor minimal system with a unique invariant measure. Recently, this result was strengthened by Ormes ([O]), who showed that every finite measure-preserving Cantor minimal system may be realised in this way, so that we have both measure-theoretic conjugacy and strong topological orbit equivalence. It would be interesting to find analogues of Ormes' theorem in the non-singular setting.

The results of this paper were announced in [Do].

ACKNOWLEDGEMENT: We gratefully acknowledge the support of the Australian Research Council and the Japan Society for the Promotion of Science. In particular, the first author was supported by a JSPS-Australian Academy of Science Fellowship, and the second author is supported by JSPS Grant-in-Aid for Scientific Research (c) no. 12640214.

2. Bratteli–Vershik diagrams and Markov odometers

In this section, we shall set forth the basic notation and properties of Bratteli–Vershik diagrams. We take as our reference[HPS], in particular section 2 of that paper, which we regard as a nice exposition of the fundamental theory.

Definition 2.1: Suppose we have a vertex set V and an edge set E, such that V is a disjoint union of finite sets $V^{(n)}, n \geq 0$, and E is a disjoint union of finite sets $E^{(n)}, n \geq 1$. We suppose that E is equipped with source mappings $s = s_n\colon E^{(n)} \to V^{(n-1)}$, and range mappings $r = r_n\colon E^{(n)} \to V^{(n)}$. We assume that $V^{(0)}$ is a singleton, that $s^{-1}(v)$ is non-empty for each $v \in V$ and that $r^{-1}(v)$ is non-empty for all $v \in V \setminus V^{(0)}$. We assume that E is equipped with a partial order so that edges e and e' are comparable if and only if $r(e) = r(e')$.

The sequence of pairs $(V^{(n-1)}, E^{(n)})$ together with the order $>$ is called an ordered Bratteli–Vershik diagram. Given an ordered Bratteli–Vershik diagram, one defines P_l^k to be the set of paths from $V^{(k)}$ to $V^{(l)}$, with the obvious order, range and source maps (see [HPS]). We define the associated space $X = X(V,E)$ of infinite paths as follows:

$$X = \{(x_n)_{n\geq 1} : x_n \in E^{(n)},\ r(x_n) = s(x_{n+1}),\ \forall n \geq 1\}.$$

A partial order on the space X is defined to be

$$x < y, \text{ if } \exists n \geq 1, \text{ such that } x_n < y_n, \text{ and } x_i = y_i, \forall i > n.$$

Following [HPS] we will be considering *essentially simple* Bratteli–Vershik diagrams. To define these, let $E_{\max}$ and $E_{\min}$ denote the set of maximal edges and the set of minimal edges, respectively. We then say that a Bratteli–Vershik diagram is essentially simple if there is a unique infinitely long path in each of $E_{\max}$ and $E_{\min}$. In fact, a point of our construction is that the diagrams we obtain are essentially simple (for suitable choices of ordering).

Of course, here we are dealing with measurable dynamics, whereas [HPS] is dealing with topological dynamics. We can now define the Vershik transformation, a version of the odometer for Bratteli diagrams. Note that Vershik refers to this transformation as an **adic** transformation.

Definition 2.2: Let $(V, E, >)$ be an essentially simple Bratteli diagram, and let $X = X(V,E)$ be the path space defined above. We define a topology on X by taking as a neighbourhood base the sets

$$[f_1, \ldots, f_k]_1^k = \{(e_1, e_2, \ldots) \in X : e_1 = f_1, e_2 = f_2, \ldots, e_k = f_k\}.$$

These sets of course also define a σ-algebra on X. They will be referred to as cylinder sets. Let $\mathcal{C}_n$ denote the (finite) σ-algebra on X generated by the cylinder sets of length n, and let $\mathcal{C}$ denote the σ-algebra generated by all cylinder sets: it is the join of all the $\mathcal{C}_n$'s. We shall also denote by

$$[f_m, \ldots, f_k]_m^k = \{(e_1, e_2, \ldots) \in X : e_m = f_m, e_{m+1} = f_{m+2}, \ldots, e_k = f_k\}.$$

This is a finite union of cylinder sets. (We may also abbreviate this by $[f]_m^k$.)

We define $T\colon X \to X$ as follows. If $x_{\max}$ is the unique maximal path, then we take $Tx_{\max} = x_{\min}$, where $x_{\min}$ is the unique minimal path. If x is some path which is not the maximal path, then at least one of its edges is not in $E_{\max}$, and we may choose such an edge in $E^{(k)}$, with the smallest possible k. Let this edge be x_k. Let f_k be its successor, let $(f_1, \dots, f_{k-1})$ be the unique finite path from $V^{(0)}$ to $s(f_k)$ each of whose edges lies in $E_{\min}$, and let $Tx = y$, where $y = (f_1, \dots, f_k, x_{k+1}, x_{k+2}, \dots)$. T is known as the Vershik transformation of X.

It is easily seen that T is an invertible transformation of X, which is a homeomorphism of X. Hence it is bimeasurable. It is well-known that T reduces to the familiar odometer, for a suitable diagram. We will discuss this below, where we will show that under certain conditions, we may consider T to be an induced transformation on a closed set of a full odometer space. Notice that in the above definition, we have used the hypothesis of essential simplicity in order to guarantee that T is defined everywhere. In the absence of this condition, it may still be possible to define a Vershik transformation, but the details become more complicated (see [HPS] for some commentary on this). A key observation in our proof of Theorem 1.1 is that we can choose the order on the Bratteli diagram we construct so that it is essentially simple.

We now define the group of finite coordinate changes on X. Set

$$P_k^0(v) = (e_i) \in X : r(e_k) = v$$

and note that each $P_k^0(v)$ is a totally ordered set. Thus, we may define a cyclic transformation $S = S_k$ on P_k^0 by

$$S(x_1, \dots, x_k) = (y_1, \dots, y_k)$$

where, if r is the least integer such that x_r is not maximal, the elements $y_1, \dots, y_{r-1}$ are minimal, y_r is the successor of x_r, and $(y_{r+1}, \dots, y_k) = (x_{r+1}, \dots, x_k)$. If all the x_r are maximal, then we take all the y_r to be minimal.

If we extend S_k to a transformation on the subsets

$$\{(x_1, \dots, x_k, x_{k+1}) \in P_{k+1}^0 : (x_1, \dots, x_k) \text{ is not maximal}\}$$

by letting

$$S_k(x_1, \dots, x_k, x_{k+1}) = (S_k(x_1, \dots, x_k), x_{k+1}),$$

then it coincides with S_{k+1} on that subset. It is clear that the restriction of the Vershik transformation to P_k^0 is nothing but S_k.

Definition 2.3: We denote by Γ_k the cyclic group generated by S_k. Further, for $v \in V^{(k)}$, let $\Gamma_k(v)$ denote the orbit $\{\gamma x : x \in P_k^0 \text{ with } r(x) = v \text{ and } \gamma \in \Gamma_k\}$. We let Γ denote the union of all the groups Γ_k.

We shall extend the action of $\gamma \in \Gamma_k$ to the whole of X by assuming that it maps a path of the form $(e_1, e_2, \ldots)$ to the path $(f_1, f_2, \ldots, f_k, e_{k+1}, \ldots)$, where $\gamma(e_1, \ldots, e_k) = (f_1, \ldots, f_k)$.

In the case of the standard odometer, the group Γ_k is the group of finite changes of the first k coordinates. In general, it is easy to see that the orbits of Γ coincide outside the maximal path, with the orbits of the Vershik transformation T; both are the set of paths which differ from the given one in finitely many edges. Since the Vershik transformation in general changes infinitely many edges in the maximal path, the T and Γ orbits of this path differ.

The above remarks show that, as is the case with the full odometer, the dynamics of T are, apart from certain exceptional points, the dynamics of Γ. In fact, we shall not seriously use dynamics of the group Γ, but rather the orbits $\Gamma_k(v)$ in section 3 below.

A convenient combinatorial way of calculating $|\Gamma_k(v)|$ — or alternatively, the cardinality of $P_k^0(v)$ — is given by the adjacency matrices $M^{(n)}$ of the Bratteli diagram. These are defined as follows.

Definition 2.4: Let

$$M_{i,j}^{(n)} = \#\{e \in E(n) : s(e) = i,\ r(e) = j\}, \quad (i,j) \in V^{(n-1)} \times V^{(n)}.$$

The integer matrices $M^{(n)}$ are called **adjacency matrices**.

The following result is now easy.

LEMMA 2.1: *The number of paths in P_k^0 with target $v \in V^{(k)}$ is given by*

$$|\Gamma_k(v)| = (M^{(1)}M^{(2)} \cdots M^{(k)})_{v_0,v}.$$

We now introduce our standard example to show how the usual odometer fits into this picture. In the future, we shall refer to this example as the **full odometer**, to distinguish it from the Vershik transformation.

Example 2.1: The odometer. For each n, we let $V^{(n)}$ be a singleton, and for a sequence $\{\ell_n\}$, let $E^{(n)} = \{0, 1, 2, \ldots \ell_n - 1\}$. Then the Bratteli–Vershik space may be identified with the infinite product space $X = \Pi_{i=1}^{\infty} \mathbb{Z}_{\ell_i}$. The Vershik transformation is then the usual odometer, and the group Γ is the usual finite coordinate change group of X.

Example 2.2: The odometer (again). The rather pleasant treatment of the odometer in Example 2.1 is not always the most convenient. A different coordinatisation of the same space is as follows. Let $\{l_n\}$ be a sequence of natural numbers with $l_0 = 1$ and, for each $n \in \mathbb{N}$, suppose that $V^{(n)}$ consists of l_n points, $\{0, 1, \ldots, l_n - 1\}$. Define the incidence matrix $M^{(n)}$ to be the $l_{n-1} \times l_n$ matrix all of whose entries are 1. That is, we assume that every vertex at the $n-1$st level is connected to every vertex at the nth level by exactly one edge. Thus the edges are indexed by pairs (i, j) where $0 \leq i < l_{n-1}, 0 \leq j < l_n$. The order on the edges with common range j is simply the usual order on the set $\{0, \ldots, l_{n-1}\}$. It is easily seen that the Vershik transformation is indeed the usual odometer action on this space. The odometer as presented in Example 2.1 represents the odometer space as the **transitions** of the present space.

In general, there are a number of different possible presentations of the odometer. Given a general Bratteli–Vershik system, it is of some interest to understand when it is one of these representations of a full odometer. In order to analyse this, we introduce the notion of the symbolic adjacency matrix of a Bratteli–Vershik system. A related notion of symbolic adjacency matrix has been used in the theory of sofic shifts ([LM], chapter 3). These matrices encode the order of the edges inherent in the definition of an ordered Bratteli–Vershik system.

Definition 2.4: Let (V, E) be an ordered Bratteli–Vershik system, and for each w in $V^{(n)}$, let the edges in $E^{(n)}$ with range w be indexed by the totally ordered set $\Lambda_w^{(n)}$. Define the **symbolic adjacency matrix**

$$\mathcal{M}_{(v,w)} = \sum_{\{\lambda \in \Lambda^{(n)}{}_w : s(\lambda) = v\}} \lambda$$

for $v \in V^{(n-1)}, w \in V^{(n)}$.

This is a matrix whose entries are formal sums of elements of the totally ordered set $\Lambda^{(n)}{}_w$. Note that each column of $\mathcal{M}$ contains each element of $\Lambda^{(n)}{}_w$ once, partitioned according to their vertex of origin.

We say that two such column vectors are **isomorphic** if there is a bijective order-preserving map between the corresponding totally ordered sets, under which the columns correspond.

The following proposition is now clear:

PROPOSITION 2.1: *Suppose that for each n, every symbolic adjacency matrix of a Bratteli–Vershik system has all columns isomorphic. Then the system is a full odometer.*

In fact, a general Bratteli–Vershik transformation space may be considered to be a closed subset of a full odometer space, and we are interested in understanding when the Vershik transformation is a (topologically) induced transformation of the odometer on this closed subset. In order to understand this, we must inject the system into an odometer in a way that respects the orders. This is also conveniently expressed using symbolic adjacency matrices.

Definition 2.6: We say that we have **an order injection** from the Bratteli–Vershik system (V, E) to the Bratteli–Vershik system (V_1, E_1) if for each n, there is an injection $\varphi_n\colon E^{(n)} \to E_1^{(n)}$ such that

(i) if $e \in E^{(n-1)}, f \in E^{(n)}$ with $r(e) = s(f)$ then $r(\varphi_{n-1}(e)) = s(\varphi_n(f))$, and

(ii) the mapping $\varphi_n\colon \Lambda^{(n)_w} \to \Lambda_1^{(n_1)_{w_1}}$ is an order-preserving injection of totally ordered spaces.

(In the above definition, notice that condition (i) implies that φ_n maps edges with common range (resp. source) to edges with common range (resp. source), so condition (ii) makes sense.)

Given such an order injection, we naturally construct a mapping φ from the path space X to the path space X_1. This mapping is clearly continuous, and hence the image of X in X_1 is a clopen subset. The issue that requires further clarification is the relationship between the Vershik maps. We are looking for conditions under which the Vershik map T_1 induces a map conjugate to T on the image of φ; that is, for each $x \in X$, we would like to see whether $\varphi(Tx)$ coincides with $T_1^k(\varphi x)$, where k is the first return time $k = \min\{m > 0 : T_1^m(\varphi x) \in \varphi(X)\}$.

In fact, under certain conditions, we may construct an order injection of a Bratteli–Vershik system into a full odometer as follows:

Firstly, suppose Λ and Λ' are finite totally ordered sets, and $\varphi\colon \Lambda \to \Lambda'$ is an order preserving mapping. If $I \subset \Lambda$ and $J \subset \Lambda'$, we say that $\sum_{\lambda \in I} \lambda \leq \sum_{\lambda' \in J} \lambda'$ if $\varphi(I) \subset J$. Note that the order does not depend on the mapping φ.

Suppose now that we have two column vectors ξ and ξ' of formal sums of elements of Λ and Λ', respectively. We say that $\xi \leq \xi'$ if they have the same number of entries, n say, and for each $1 \leq j \leq n$ we have $\xi_j \leq \xi'_j$.

Lemma 2.2: *Let $n \in \mathbb{N}$. Suppose we have a finite collection of totally ordered sets, $\{\Lambda_i\}$, and, for each i, a column vector ξ_i with n entries, of formal sums of elements of Λ_i. Then there exist*

(i) *a totally ordered set Λ,*

(ii) *for each i, an order-preserving injection $\varphi_i\colon \Lambda_i \to \Lambda$,*

(iii) *an n-vector ξ of formal sums of entries of Λ,* *such that $\xi_i \leq \xi$ for all i.*

Proof: The proof, a simple induction, is left to the reader. ∎

Now suppose that (V, E) is a given Bratteli–Vershik system. Define the system $(\tilde{V}, \tilde{E})$ as follows:

For each n, starting from the totally ordered sets $\Lambda^{(w)}$ and from the columns of the symbolic adjacency matrix $\mathcal{M}$, use the Lemma to choose a totally ordered set $\Lambda^{(n)}$ and a single vector $\xi^{(n)}$ which is greater than or equal to all the columns of $\mathcal{M}$.

Set $\tilde{V}^{(n)} = V^{(n)}$ and let the symbolic adjacency matrix of $(\tilde{V}, \tilde{E})$ be given by

$$\mathcal{M} = (\xi^{(n)}, \xi^{(n)}, \ldots, \xi^{(n)}).$$

Proposition 2.2: *Suppose that there is an order injection of an essentially simple Bratteli–Vershik system X into another X^0. Suppose further that this injection maps the infinite minimal path of X to the infinite minimal path of X^0 and the infinite maximal path of X to the infinite maximal path of X^0. Then the image of the space X is a closed subset of X^0, and the Vershik transformation T on X is conjugate to the induced transformation of the Vershik transformation T^0 on X^0.*

Proof: The fact that we have taken an order injection of one system into another ensures that the induced transformation is conjugate to the transformation T on the image of X in X^0 on all paths except perhaps the maximal path, where an infinite number of edges might be changed. The assumption that the maximal path of X^0 is included in the image of X now guarantees that the induced action coincides everywhere. ∎

We shall use the following condition in section 6 to show that the transformations we produce are induced transformations of a full odometer.

Corollary 2.1: *Suppose that $X = X(V, E)$ is a Bratteli–Vershik space such that every vertex set $V^{(n)}$ contains two different vertices v_n, v'_n with the property that every minimal edge in $E^{(n+1)}$ has source v_n and every maximal edge in $E^{(n)}$ has source v'_n. Then X can be embedded as a topologically induced transformation in a full odometer space.*

Proof: It suffices to notice that, under the hypothesis given, the order injections $\varphi_i \colon \Lambda_i \to \Lambda$ can be chosen such that φ_i sends the maximal (resp. minimal) element to the maximal (resp. minimal) element of Λ, and hence the assumption of the Proposition is verified. ∎

The above Proposition and Corollary indicate that many Bratteli–Vershik systems (including the ones we shall construct in section 6) are familiar objects in ergodic theory: topologically induced transformations on closed subsets of odometers. Note that the proof is entirely topological; we have not so far discussed measures on X. We now proceed to do that. In particular, we would like to define a family of measures on X which generalise the traditional product measures on infinite product spaces.

In general, notice that T is a measurable transformation of the measurable space $(X, \mathcal{C})$. A measure μ is quasi-invariant for T if μ and $T\mu$ are equivalent as measures. The dynamical system $(X, \mathcal{C}, \mu, T)$ is a generalisation of the traditional odometer action on an infinite product space. The simplest kinds of measures one may define on these spaces in general are Markov measures. In some cases, they reduce to traditional infinite product measures.

Definition 2.7: Markov measures. We say that a matrix

$$P^{(n)} = \{P^{(n)}_{v,e}\}_{(v,e)\in V^{(n-1)}\times E^{(n)}}$$

is a **stochastic matrix** if it satisfies the following two conditions:

(i) $P^{(n)}_{v,e} > 0 \Leftrightarrow s(e) = v$,

(ii) $\sum_{\{e\in E^{(n)}:s(e)=v\}} P^{(n)}_{v,e} = 1 \ \forall v \in V^{(n-1)}$.

Given a sequence $P^{(n)}$ of stochastic matrices and a probability measure ν_0 on $V^{(0)}$ such that

$$\nu_0(v) > 0, \ \forall v \in V^{(0)},$$

we define a measure μ on cylinder sets by

$$\mu([e_1, e_2 \cdots e_n]_1^n) = \nu_0(s(e_1))P^{(1)}_{s(e_1),e_1}P^{(2)}_{s(e_2),e_2}\cdots P^{(n)}_{s(e_n),e_n}.$$

A measure defined in this way is called a **Markov measure** and the dynamical system $(X, \mathcal{B}, T, \mu)$ is said to be a **Markov odometer** if μ in addition satisfies the following two conditions:

(M1) For almost every x, there exists an integer $n \geq 1$ and a block $y_1y_2\cdots y_n$ such that $x_n < y_n$ and $\mu([y_1y_2\cdots y_nx_{n+1}]_1^{n+1}) > 0$.

(M2) For almost every x, there exists an integer $m \geq 1$ and a block $z_1z_2\cdots z_m$ such that $z_m < x_m$ and $\mu([z_1z_2\cdots z_mx_{m+1}]_1^{m+1}) > 0$.

We note that a Markov odometer T is invertible, and has the property that $n = \sup\{i \geq 1 : (Tx)_i \neq x_i\} < \infty$, a.e. x. Furthermore, T is non-singular; in fact

$$\frac{dT\mu}{d\mu} = \frac{\nu_0(s(y_1))P^{(1)}_{s(y_1),y_1} P^{(2)}_{s(y_2),y_2} \cdots P^{(n)}_{s(y_n),y_n}}{\nu_0(s(x_1))P^{(1)}_{s(x_1),x_1} P^{(2)}_{s(x_2),x_2} \cdots P^{(n)}_{s(x_n),x_n}}$$

where $y_i = (Tx)_i, 1 \leq i \leq n$.

Example 2.3: The odometer revisited. It is easy to describe the Markov measures on the odometer space presented in Example 2.1. In fact, all of the sets $V^{(n)}$ are singletons, so in particular, $\nu(v_0) = 1$, and the stochastic matrices $P^{(n)}$ are just vectors depending on the edges. The Markovian condition (ii) ensures that the entries of these vectors sum to 1. Hence, we can think of $\eta_k(\{e\}) := P^{(k)}_{v_k,e}$ as a probability measure on the set of edges $E^{(k)} = \{e_0, \ldots, e_{\ell_k - 1}\}$. A Markov measure on this Bratteli diagram is thus exactly the traditional infinite product measure $\bigotimes_{i=1}^{\infty} \eta_i$.

In fact, if we change our presentation of the odometer just a little, the notion of Markov measure changes too, and we no longer stay within the confines of product measures, as is shown by the coordinatisation of Example 2.2. In fact, in that coordinatisation, the entries of the stochastic matrices are indexed by pairs of vertices, and we may write $P^{(n)}_{s(e),e} = P^{(n)}_{s(e),r(e)}$. Thus $P^{(n)}$ is an $l_{n-1} \times l_n$ matrix, all of whose entries are positive, and whose rows sum to 1. The measure of a cylinder set $[v_1, \ldots, v_n]_1^n$ is then given by the product $\nu(\{v_0\})P^{(1)}_{v_0,v_1} P^{(2)}_{v_1,v_2} \cdots P^{(n)}_{v_{n-1},v_n}$. This is not a product measure unless the P matrices are independent of their first indices, that is, all rows are the same. If this is the case, then we may let $\mu^{(n)}(\{y\}) = P^{(n)}_{v_{n+1},y}$, and the Markov measure equals the infinite product $\bigotimes_{n=0}^{\infty} \mu^{(n)}$.

In general, this description of Markov odometers as measures whose values depend on successive pairs of vertices of elements of X, is valid whenever the diagram has the property that there is at most one edge between two successive vertices (i.e., the adjacency matrices have all entries 0 or 1). In this case, the entries of the P-matrices correponding to non-zero entries of the M-matrices must be non-zero. The notion of a product measure is defined only if every vertex in $V^{(n-1)}$ has the same number of successors, and there is exactly one edge between them. That is, we have $E^{(n)} = V^{(n-1)} \times V^{(n)}$. In this case, we have a product if and only if $P^{(n)}_{i,j}$ is independent of i.

3. The G-measure formalism

The G-measure formalism ([BD]) has proved useful in analysing the structure of measures on the classical odometer space. In this section, we shall adapt it to the setting of Bratteli–Vershik systems, and apply it to Markov measures.

Definition 3.1: Let $X = X(V, E)$ be the space associated to an ordered Bratteli diagram with Vershik transformation T, and let μ be a T-quasi-invariant measure on X. For a fixed k, define the tail measure $\mu^{(k)}$ by setting, for $n \geq k$,

$$\mu^{(k)}([a_1, \ldots, a_n]_1^n) = \frac{1}{|\Gamma_k(r(a_k))|} \sum_{\gamma \in \Gamma_k} \mu^{(k)}(\gamma[a_1, \ldots, a_n]_1^n).$$

(Notice that the action of Γ_k was defined so that the summation above is over all cylinder sets $[x_1, \ldots, x_k, a_{k+1}, \ldots, a_n]_1^n$ with $r(x_k) = r(a_k)$.)

It is then easily seen that $\mu^{(k)}$ is a Γ_k-invariant measure which is equivalent to μ. We denote by G_k the Radon–Nikodým derivative $d\mu/d\mu^{(k)}$.

Similarly, we denote by g_k the Radon–Nikodým derivative $d\mu^{(k-1)}/d\mu^{(k)}$.

It is then clear that the function G_k satisfies

$$\frac{1}{|\Gamma_k(r(a_k))|} \sum_{\gamma \in \Gamma_k} G_k(\gamma a) = 1$$

for all infinite paths $a \in X$.

Furthermore, $G_k(a) = g_1(a) g_2(a) \ldots g_k(a)$, where for each i, g_i satisfies the two conditions:

(i) (*invariance*) $g_i(a)$ is independent of $(a_0, \ldots, a_{i-1})$, and

(ii) (*normalisation*) $\frac{1}{|E^{(i)}(v,w)|} \sum_{\{e \in E^{(i)}(v,w)\}} g_i(a_1, \ldots, a_{i-1}, e, a_{i+1}, \ldots) = 1$.

Here, v denotes $r(a_{i-1})$, w denotes $s(a_{i+1})$, and $E^{(i)}(v, w)$ denotes $\{e \in E^{(i)} : s(e) = v, r(e) = w\}$.

Definition 3.2: A family of functions $\{g_k\}$ satisfying conditions (i) and (ii) above is called a **normalised invariant family**. The corresponding sequence of functions $G = \{G_k\}$ is called a **normalised compatible family**.

A measure μ on X satisfying

$$G_k = d\mu/d\mu^{(k)}$$

is called a G-**measure**.

It is clear that, for a fixed normalised compatible family G, the set of G-measures is a convex set inside the set of all T-quasi-invariant probability measures. Moreover, it follows easily (as in[BD]) that the extreme points in this convex set are T-ergodic.

Furthermore, the existence of G-measures is clear:

LEMMA 3.1: *Let $\{G_n\}$ be a normalised compatible family and let $\{\nu^n\}$ be any sequence of Γ_n quasi-invariant probability measures on X. Then any weak* limit of the sequence $\{\mu^n\}$ defined by*

$$\mu^n = G_n\left(\frac{1}{|\Gamma_n|}\sum_{\gamma\in\Gamma_n}\nu^n\circ\gamma\right)$$

is a G-measure

The easy proof is given in [BD].

We single out for special attention the case when there is just one element in this convex set. In this case, the unique G-measure, μ say, is T-ergodic, and we say that we have a **uniquely ergodic G-measure.**

This condition generalises the traditional notion of unique ergodicity for homeomorphisms to the non-singular setting. The following result generalises Proposition 1.6 of [BD] from the odometer to the cadre of Bratteli diagrams.

PROPOSITION 3.1: *Let $G = \{G_n\}$ be a normalised compatible family of continuous functions. The following are equivalent.*

(i) *There is a unique G-measure — which is therefore T-ergodic.*

(ii) *For every $f\in C(X)$ the sequence*

$$A_k(f)(x) = \frac{1}{|\Gamma_k(r(x_k))|}\sum_{\gamma\in\Gamma_k(r(x_k))} G_k(\gamma x)f(\gamma x)$$

converges uniformly to a constant.

(iii) *For every $f\in C(X)$ the sequence*

$$A_k(f)(x) = \frac{1}{|\Gamma_k(r(x_k))|}\sum_{\gamma\in\Gamma_k(r(x_k))} G_k(\gamma x)f(\gamma x)$$

converges pointwise (for every $x\in X$) to a constant.

Proof: Notice that we can write $A_n(f)(x)$ as

$$A_n(f)(x) = \left(G_n\cdot\left(\frac{1}{|\Gamma_n|}\sum_{\gamma\in\Gamma_n}\delta_x\circ\gamma\right)\right)(f),$$

where δ_x is the Dirac measure at x.

Thus for a fixed x or for x_n depending on n, we have the format of Lemma 3.1, and any weak* limit of the sequence of linear functionals $A_n(x)(f) = A_n(f)(x)$ will be a G-measure.

If (i) held, but (ii) did not, and if convergence did occur, the limit would necessarily be $\int f d\mu$, so by our assumption, there would exist $f \in C(X)$, $\varepsilon > 0$ and a subsequence (n_i) such that

$$|A_{n_i}(x_{n_i})(f) - \int f d\mu| \geq \varepsilon.$$

We could then obtain a G-measure different from μ by taking any limit point of the sequence of measures $A_{n_i}(x_{n_i})$.

Clearly (ii) implies (iii), so let us show that (iii) implies (i).

We know that for any fixed x, $A_n(x)$ converges to a G-measure μ_x, and for each fixed f, that the limit $\mu_x(f)$ is independent of x. In other words, there is a G-measure μ such that

$$\lim_n A_n(x)(f) = \int f d\mu$$

for all $x \in X$.

By the bounded convergence theorem, we have for any probability measure $\nu \in \mathcal{M}(X)$,

$$\int f d\mu = \int \left(\int f d\mu \right) d\nu = \lim_n \int A_n(x)(f) d\nu.$$

Suppose now that ν is a G-measure. Then we have

$$\begin{aligned}\int A_n(f)(x) d\nu &= \int \frac{1}{|\Gamma_n(r(x_n))|} \sum_{\gamma \in \Gamma_n(r(x_n))} G_n(\gamma x) f(\gamma x) d\nu(x) \\ &= \int G_n(x) f(x) d\nu^n(x) = \int f(x) d\nu(x).\end{aligned}$$

We have shown that $\nu = \mu$, and this completes the proof. ∎

We now show how to compute the G-functions for Markov measures, and particularise Proposition 3.1 to that setting.

Suppose that we are given the data of Definition 2.5, that is, a Bratteli–Vershik space $X(V, E)$, a probability measure ν_0 on $V^{(0)}$, and a sequence of Markov matrices $(P^{(n)})$, whose entries $P^{(n)}_{v,e}$ are indexed by $v \in V^{(n-1)}$ and $e \in E^{(n)}$. We are thinking of the P-matrices as transporting the measure ν from one coordinate space to the next. Thus it is natural to define

$$\nu_{k+1}(\{v\}) := \sum_{\{e \in E^{k+1}; r(e) = v\}} \nu_k(\{s(e)\}) P^{(k+1)}_{s(e),e}.$$

Actually, using a slight abuse of notation, we are going to write $\nu(v)$ instead of $\nu(\{v\})$.

Given this notation, we may now state:

PROPOSITION 3.2: *Let $e = (e_n)_{n\geq 1} \in X$ and $k \geq 1$. Then:*

(i) $\mu^{(k)}([e_{k+1},\dots,e_n]_{k+1}^n) = \frac{1}{|\Gamma_k(w)|}\nu_{k+1}(w)P^{(k+1)}_{w,e_{k+1}}\cdots P^{(n)}_{s(e_n),e_n}$, *where* $w = s(e_{k+1})$.

(ii) $G_k(e) = |\Gamma_k(r(e_k))|\frac{\nu_0(s(e_1))}{\nu_k(r(e_k))}P^{(1)}_{s(e_1),e_1}\cdots P^{(k)}_{s(e_k),e_k}$.

(iii) $g_k(e) = \frac{|\Gamma_k(w)|}{|\Gamma_k(v)|}\frac{\nu_k(v)}{\nu_{k+1}(w)}P^{(k)}_{v,e_k}$, *where* $s(e_k) = v$ *and* $r(e_k) = w (= s(e_{k+1}))$.

(iv) *Let* $f \in C(X)$. *Then*

$$A_k(f)(e) = \frac{1}{\nu_k(w)}\sum \nu_0(s(x_1))P^{(1)}_{s(x_1),x_1}\cdots P^{(k)}_{s(x_k),x_k}f(x_1,\dots,x_k,e_{k+1},e_{k+2},\dots),$$

where $w = s(e_{k+1}) = r(e_k)$ *and the sum is over all* $x \in P_k^0$ *with* $r(x) = w$.

Proof: Let $w = s(e_{k+1})$. Then

$$\sum_{\gamma\in\Gamma_k}\mu^{(k)}(\gamma[e_1,\dots,e_n]_1^n) = \nu_{k+1}(w)P^{(k+1)}_{w,e_{k+1}}\cdots P^{(n)}_{s(e_n),e_n}$$

where we note that the summation is over all cylinder sets

$$[x_1,\dots,x_ke_{k+1},\cdots,e_n]_1^n$$

with $r(x_k) = r(e_k)$.

Hence,

$$\mu^{(k)}([e_1,\dots,e_n]_1^n) = \frac{1}{|\Gamma_k(w)|}\nu_{k+1}(w)P^{(k+1)}_{w,e_{k+1}}\cdots P^{(n)}_{s(e_n),e_n}.$$

Now (i) follows.

Further, note that

$$G_k(e) = \frac{d\mu}{d\mu^{(k)}}(e) = \frac{\mu([e_1,\dots,e_{k+1}]_1^{k+1})}{\mu^{(k)}([e_1,\dots,e_{k+1}]_1^{k+1})}$$

and that

$$g_k(e) = \frac{d\mu^{(k-1)}}{d\mu^{(k)}}(e) = \frac{\mu^{(k-1)}([e_1,\dots,e_{k+1}]_1^{k+1})}{\mu^{(k)}([e_1,\dots,e_{k+1}]_1^{k+1})}.$$

(ii) and (iii) now follow, and (iv) is an easy consequence of (ii). ∎

In Proposition 3.2 (iv), the expression for $A_k(f)(e)$ may be interpreted as

$$\mathbb{E}_\mu[f|\mathcal{C}^k](e_{k+1},e_{k+2},\dots)$$

where $\mathcal{C}^k$ is the σ-algebra generated by the Γ_k-invariant sets.

We now give a sufficient condition for uniqueness, which we shall actually use in section 6. This condition should be compared with the Lipschitz-type conditions of [BD].

Definition 3.3: We shall say that a Markov measure is **uniquely ergodic** if it is uniquely ergodic when considered as a G-measure

PROPOSITION 3.3: *For each k and for $v \in V_{k-1}, w \in V_k$, let*

$$\Theta^{(k)}(v,w) = \frac{1}{\nu_k(w)} \sum_{\{x \in E_k : s(x)=v, r(x)=w\}} P^{(k)}_{v,x}.$$

Then if $\max_{v \in V_{k-1}, w \in V_k} |\Theta^{(k)}(v,w) - 1| \to 0$ as $k \to \infty$, we have a uniquely ergodic Markov measure.

Proof: Firstly, notice that for any continuous function f, and for any $\varepsilon > 0$, there exists $m > 0$ and a continuous function f_0, so that $f_0(e)$ depends only on $(e_0, \ldots, e_m)$, and $||f - f_0||_\infty < \varepsilon$. By the normalisation condition on the functions G_k, we then have for all k, $||A_k(f) - A_k(f_0)||_\infty < \varepsilon$. Hence, it is sufficient to prove that the sequence $A_k(f)$ converges uniformly under the assumption that f depends only on the coordinates $1, \ldots, m$.

Now if $k \geq m+1$, then

$$\begin{aligned}
\mathbb{E}_\mu(f) - A_k(f)(x) &= \sum_1 \sum_2 \nu_0(s(e_1)) P^{(1)}_{s(e_1),e_1} \cdots P^{(k-1)}_{s(e_{k-1}),e_{k-1}} f(e_1, \ldots, e_m) \\
&- \frac{1}{\nu_k(r(x_k))} \sum_1 \sum_2 \nu_0(s(e_1)) P^{(1)}_{s(e_1),e_1} \cdots P^{(k-1)}_{s(e_{k-1}),e_{k-1}} f(e_1, \ldots, e_m) \sum_3 P^{(k)}_{v,e} \\
&= \sum_1 \sum_2 \nu_0(s(e_1)) P^{(1)}_{s(e_1),e_1} \cdots P^{(k-1)}_{s(e_{k-1}),e_{k-1}} f(e_1, \ldots, e_m) \{1 - \Theta^{(k)}(v, r(x_k))\},
\end{aligned}$$

where the subscript 1 indicates that the sum is over $v \in V^{(k-1)}$, the subscript 2 that the sum is over $e \in P^0_k(v)$, and the subscript 3 that the sum is over $e \in E^{(k)}(v, r(x_k))$. Hence,

$$|\mathbb{E}_\mu(f) - A_k(f)(x)| \leq |\mathbb{E}_\mu[f]| \cdot \max_{v,w} |\Theta^{(k)}(v,w) - 1|.$$

This completes the proof. ∎

More generally, for $v \in V^{(n)}$ and $w \in V^{(m)}$, $n < m$, let

$$\Theta(v,w) = \Theta^{(n,m)}(v,w) := \frac{1}{\nu_m(w)} \sum_{\{x \in P^n_m : s(x)=v, r(x)=w\}} P^{(n)}_{v,x_n} \cdots P^{(m)}_{s(x_m),x_m}.$$

By considering a contraction of the Bratteli–Vershik diagram, we have as an immediate consequence:

COROLLARY 3.1: *Suppose that there exists a stictly increasing sequence (N_k) of integers, such that*

$$\max_{v\in V_{N_k}, w\in V_{N_{k+1}}} |\Theta^{(N_k,N_{k+1})}(v,w)-1| \to 0 \quad \text{as } k\to\infty.$$

Then there is a uniquely ergodic Markov measure.

4. Ordered multiple towers

We are going to construct a Bratteli diagram inside any ergodic non-singular dynamical system. This will be done by using tower constructions. In this section, we will give a careful definition of a tower, and of a multiple tower. The notion of ordered tower defined below was called an **array** by Krieger.

Throughout this section, we assume that $(X,\mathcal{B},\mu,T)$ is an ergodic non-singular transformation with $\mu(X)<\infty$. In fact, without loss of generality, we may assume that μ is a probability measure on X.

The towers will be constructed from partially defined transformations of X, that is, invertible measurable mappings ϕ both of whose domain $\mathrm{Dom}(\phi)$ and image $\Im(\phi)$ are sets of positive measure.

$[T]_*$ is the set of all partially defined transformations ϕ satisfying

$$\phi x\in \mathrm{Orb}_T(x), \quad \text{a.e. } x\in \mathrm{Dom}(\phi),$$

where $\mathrm{Orb}_T(x)=\{T^n x : n\in\mathbb{Z}\}$.

Recall that the **full group** $[T]$ of T is the set of elements of $[T]_*$ whose domain and range are both all of X up to a null set.

Definition 4.1: Let A be a set of positive measure in $\mathcal{B}$ and $\mathcal{P}=\{A_\alpha:\alpha\in\Lambda\}$ be a finite partition of A by measurable subsets $A_\alpha\in\mathcal{B}$, where the index set Λ is a finite totally ordered set. Suppose that for each pair $\alpha,\beta\in\Lambda$, we are given partially defined transformations $\xi_{\alpha,\beta}$: $A_\beta\to A_\alpha$, belonging to $[T]_*$, which satisfy:

(i) $\xi_{\alpha,\alpha}=\mathrm{Id}|_{A_\alpha}$, and

(ii) $\xi_{\alpha,\beta}\xi_{\beta,\gamma}=\xi_{\alpha,\gamma}$.

Then we shall say that the collection $\xi=\{\xi_{\alpha,\beta}:\alpha,\beta\in\Lambda\}$ is an **ordered tower of A**.

For ease of notation, we will consider ξ as the primary object, and will denote the partition $\mathcal{P}$ by $\mathcal{P}(\xi)$ and the sets A_α as $\mathcal{P}(\xi)_\alpha$. Similarly, we write $\Lambda(\xi)=\Lambda, \mathrm{supp}(\xi)=A$

An ordered tower always has an **associated transformation** $S = S(\xi) \in [T]_*$ defined by

$$Sx = \xi_{\beta,\alpha}x, \quad x \in A_\alpha$$

where

$$\beta = \min_{\theta \in \Lambda}\{\theta : \theta > \alpha\}.$$

Notice that if α is maximal, then S is not defined on $\mathcal{P}(\xi)_\alpha$, and that if α is minimal, then S^{-1} is not defined on $\mathcal{P}(\xi)_\alpha$.

The orbit of $x \in \mathcal{P}(\xi)_\alpha$ under S, $\mathrm{Orb}_S(x) = \mathrm{Orb}_\xi(x)$, is given by

$$\{\xi_{\beta,\alpha}x : \beta \in \Lambda\}.$$

A measurable subset $E \subset \mathrm{supp}(\xi)$ is said to be ξ-invariant if $\mathrm{Orb}_\xi(x) \subset E$ for a.e. $x \in E$.

It should be apparent to the reader that the structure of an ordered tower is not unlike that of a set of edges in a Bratteli diagram with a common range.

An important tool in defining our Markov measures will be the concept of a measure of constant Jacobian.

Definition 4.2: We say that a tower ξ is of **constant Jacobian** with respect to μ if $\frac{d\xi_{\alpha,\beta}\mu}{d\mu}(x)$ is a constant independent of x for a.e. $x \in A_\beta$.

Actually, we can identify the constant in Definition 4.2. It is given by

$$\frac{d\xi_{\alpha,\beta}\mu}{d\mu}(x) = \frac{\mu(A_\alpha)}{\mu(A_\beta)}.$$

We now wish to define the concept of a multiple tower: this will eventually become a set of vertices $V^{(n)}$ in a layer of our Bratteli diagram.

Definition 4.3: Let $\xi^{(v)}, v = 1, \ldots, m$, be a finite collection of ordered towers, $\xi^{(v)} = \{\xi^{(v)}_{\alpha,\beta} : \alpha, \beta \in \Lambda^{(v)}\}$, such that the sets $\mathrm{supp}(\xi^{(v)})$, $1 \le v \le m$, are disjoint.

Define

$$\Lambda = \{\alpha v : 1 \le v \le m, \alpha \in \Lambda^{(v)}\}$$

and define a partial order on Λ by

$$\alpha v < \alpha' v' \quad \text{if } v = v' \text{ and } \alpha < \alpha'.$$

We call the collection

$$\xi = \{\xi^{(v)} :\ 1 \le v \le m\}$$

an **ordered multiple tower**.

As we did in the case of a single tower, we wish to regard ξ as the principal object, and so we establish the following notation.

$$\begin{aligned}\mathcal{P}(\xi) =& \{\mathcal{P}(\xi)_\alpha : \alpha \in \Lambda(\xi^{(v)}),\ 1 \le v \le m\},\\ \operatorname{supp}(\xi) =& \bigcup_{v=1}^{m} \operatorname{supp}(\xi^{(v)}),\\ \operatorname{Orb}_\xi(x) =& \operatorname{Orb}_{\xi^{(v)}}(x), \quad \text{if } x \in \operatorname{supp}(\xi^{(v)}),\\ \xi_{\alpha v,\beta v}x =& \xi^{(v)}_{\alpha,\beta}x \quad \text{if } \alpha,\beta \in \Lambda^{(v)} \text{ and } x \in \mathcal{P}(\xi^{(v)})_\beta,\\ \Lambda(\xi) =& \Lambda.\end{aligned}$$

A union of $\xi^{(v)}$-invariant subsets is a ξ-invariant set.

We likewise define the **associated transformation of** ξ to be the partially defined transformation obtained by taking the associated transformation of the $\xi^{(v)}$ on each $\operatorname{supp}(\xi^{(v)})$.

If μ is a measure on $\operatorname{supp}(\xi)$, we say that ξ is of constant Jacobian (for μ) when every $\xi^{(v)}$ is of constant Jacobian (for μ).

Example 4.1: Let $X(V, E)$ be a Bratteli–Vershik system, and fix an integer n. Then $V^{(n)}$ provides the index set for a multiple tower ξ, made up of the towers $\xi^{(v)} = \{e \in P_n^0 : r(e_n) = v\}$. The associated transformation of ξ is exactly the Vershik transformation (restricted to P_n^0), except on the top level. Furthermore, a sequence of stochastic matrices $P^{(n)}$ provides a measure μ of constant Jacobian for ξ such that

$$\frac{d\xi^{(v)}_{\alpha,\beta}\mu}{d\mu}(x) = \frac{\nu_0(s(\alpha_1))P^{(1)}_{s(\alpha_1),\alpha_1}\cdots P^{(n)}_{s(\alpha_n),\alpha_n}}{\nu_0(s(\beta_1))P^{(1)}_{s(\beta_1),\beta_1}\cdots P^{(n)}_{s(\beta_n),\beta_n}},$$

where $x \in [\beta]_1^n$.

5. Refinements of multiple towers

The proof of our main theorem will use a notion of refinement of multiple towers. This will be done in section 6, using successive applications of Rokhlin's Lemma. The key idea is to take a union of layers, one from each of the towers in our collection ξ, and refine the partition arising from the fact that the supports of the individual towers in ξ are disjoint. We will see how this gives rise to a new multiple tower, which we shall call a refinement of ξ.

In terms of Bratteli diagrams, this procedure can be thought of as creating a new set of vertices "underneath" the layer provided by ξ.

Throughout this section, we let $\xi = \{\xi^{(1)}, \dots, \xi^{(n)}\}$, $\xi^{(v)} = \{\xi^{(v)}_{\alpha,\beta} : \alpha, \beta \in \Lambda^{(v)}\}$, be an ordered multiple tower of X. For each v, we let $\alpha^{(v)} \in \Lambda^{(v)}$ be the minimal element.

Given these choices, we let $E(v) = \mathcal{P}(\xi^{(v)})_{\alpha^{(v)}}$, and $E = \bigcup_{v=1}^{n} E(v)$.

The reader might find it helpful to think of the sets $E(v)$ as the base sets in the towers $\xi^{(v)}$, and of E as their union.

Clearly the sets $E(v)$ form a partition of E. The other piece of data which we will need for the construction is an ordered multiple tower η of E such that the partition $\mathcal{P}(\eta)$ of E is a refinement of the partition induced by the $E(v)$'s. Actually, we need something slightly less; it is sufficient that the support of η be a subset of E of positive measure, and that the partition $\mathcal{P}(\eta)$ be a refinement of the partition of $\mathrm{supp}(\eta)$ given by $\{E(1) \cap \mathrm{supp}(\eta), \dots, E(n) \cap \mathrm{supp}(\eta)\}$.

The reader may think of η as a tower which "cuts across" the base level of the towers $\xi^{(v)}$.

We will denote η as $\{\eta^{(1)}, \dots, \eta^{(m)}\}$, $\eta^{(w)} = \{\eta^{(w)}_{\epsilon,\delta} : \epsilon, \delta \in \Lambda^{(w)}\}$.

With these data established, we may now define the refinement ζ of ξ induced by η.

Definition 5.1: Let $\xi, \alpha^{(v)}$ and η be as above.

Define the ordered multiple tower $\zeta = \{\zeta^{(w)} : 1 \le w \le m\}$ as follows.

(i) The support of ζ is the ξ-invariant set $\bigcup_{v=1}^{n} \bigcup_{\beta \in \Lambda^{(v)}} \xi_{\beta,\alpha^{(v)}}(\mathrm{supp}(\eta))$.

(ii) The index set is given by

$$\Lambda(\zeta^{(w)}) = \{\beta v \epsilon : 1 \le v \le n,\ \beta \in \Lambda(\xi^{(v)}), \epsilon \in \Lambda(\eta^{(w)}), \mathcal{P}(\eta^{(w)})_\epsilon \subset E(v)\}.$$

(iii) On each $\Lambda(\zeta^{(w)})$, we take a total order, defined as follows:
For $\beta v \epsilon, \beta' v' \epsilon' \in \Lambda(\zeta^{(w)})$,

$$\beta v \epsilon < \beta' v' \epsilon' \text{ if either (1) } \epsilon < \epsilon',\text{ or, (2) } \epsilon = \epsilon',\ v = v' \text{ and } \beta < \beta'.$$

(iv) Putting

$$A^{(w)}_{\beta v \epsilon} = \xi^{(v)}_{\beta,\alpha^{(v)}}(\mathcal{P}(\eta^{w})_\epsilon) \subset \mathcal{P}(\xi^{(v)})_\beta,$$

the partially defined transformations of $[T]_*$ are given by

$$\begin{aligned}
\zeta^{(w)}_{\alpha^{(v)} v\delta, \alpha^{(v')} v' \epsilon} &= \eta^{(w)}_{\delta,\epsilon} \quad \text{for } \alpha^{(v)} v \delta, \alpha^{(v')} v' \epsilon \in \Lambda(\zeta^{(w)}),\\
\zeta^{(w)}_{\beta v' \epsilon, \alpha^{(v')} v' \epsilon} &= \xi^{(v')}_{\beta,\alpha^{(v')}} \quad \text{for } \beta v' \epsilon, \alpha^{(v')} v' \epsilon \in \Lambda(\zeta^{(w)}),\\
\zeta^{(w)}_{\beta v \delta, \gamma v' \epsilon} &= \zeta^{(w)}_{\beta v \delta, \alpha^{(v)} v \delta} \cdot \zeta^{(w)}_{\alpha^{(v)} v \delta, \alpha^{(v')} v' \epsilon} \cdot (\zeta^{(w)}_{\gamma w' \epsilon, alpha^{(v')} v' \epsilon})^{-1}\\
&\qquad \text{for } \beta v \delta, \gamma v' \epsilon, \alpha^{(v)} v \delta, \alpha^{(v')} v' \epsilon \in \Lambda(\zeta^{(w)}).
\end{aligned}$$

These data uniquely determine an ordered multiple tower $\zeta = \{\zeta^{(1)}, \dots, \zeta^{(m)}\}$, where $\zeta^{(w)} = \{\zeta^{(w)}_{\beta v\delta,\gamma v'\epsilon} : \beta v\delta, \gamma v'\epsilon \in \Lambda(\zeta^{(w)})\}$.

We call the multiple ordered tower ζ the **refinement of ξ induced by η**.

This definition is somewhat standard, and is indeed similar to Definition 1.1 of [VL], where it is applied to measure-preserving systems.

Our next task, and in some senses the main novelty of this section, is to see what effect refinements of multiple towers have on measures. In fact, we will see that the prescription of measures of constant Jacobian for the two towers gives rise naturally to a stochastic matrix P.

PROPOSITION 5.1: *Suppose that the ordered multiple tower ξ is of constant Jacobian with respect to μ and that there is a measure ν on E, equivalent to μ, and under which η is of constant Jacobian.*

Then there is a unique measure ρ on $\mathrm{supp}(\zeta)$, *equivalent to μ, satisfying the following conditions:*

(i) *$\rho = \nu$ on the set* $\mathrm{supp}(\zeta)$.

(ii) *ρ is of constant Jacobian for ζ. Indeed, for $\beta v\delta, \gamma v'\epsilon \in \Lambda(\zeta^{(w)})$ and $x \in A^{(w)}_{\gamma v'\epsilon}$ we have*

$$\frac{d\zeta^{(v)}_{\beta v\delta,\gamma v'\epsilon}\rho}{d\rho}(x) = \frac{\mu(\mathcal{P}(\xi^{(v)})_\beta)}{\mu(E(v))} \cdot \frac{\nu(\mathcal{P}(\eta^{(w)})_\delta)}{\nu(\mathcal{P}(\eta^{(w)})_\epsilon)} \cdot \frac{\mu(E(v'))}{\mu(\mathcal{P}(\xi^{(v')})_\gamma)}.$$

Proof: In fact, it is easily seen that $\rho(W)$ must be defined for a measurable subset W of $\mathcal{P}(\xi^{(v)})_\gamma$ by

$$\rho(W) = \frac{\mu(\mathcal{P}(\xi^{(v)})_\gamma)}{\mu(\mathcal{P}(\xi^{(v)})_{\alpha^{(v)}})} \cdot \nu(\xi^{(v)}_{\alpha^{(v)},\gamma}(W)). \qquad \blacksquare$$

Proposition 5.1 implies that

$$\frac{d\rho}{d\mu}(x) = \frac{d\nu}{d\mu}(\xi^{(v)}_{\alpha^{(v)},\gamma}x), \ x \in \mathcal{P}(\xi^{(v)})_\gamma.$$

We shall write $\rho = \mu_\nu$ and call it the **Markov extension of the measure μ by the measure ν**.

We now continue to make the link with Bratteli diagrams. At the end of section 4, we saw how a multiple tower could be thought of as a set of vertices of a Bratteli diagram. We now see how a multiple tower together with a refinement of that tower can be thought of as two successive layers of a diagram, and we see how the Markov extension of a measure as introduced in Proposition 5.1 naturally induces a stochastic matrix.

In fact, maintaining the notation of Definition 5.1 and Proposition 5.1, one defines an adjacency matrix M and a stochastic matrix P as follows. Firstly let $V = \{1, \ldots, m\}$, $V' = \{1, \ldots, n\}$ and $E = \bigcup_{w=1}^{m} \Lambda(\eta^{(w)})$. Recall that each $\Lambda(\eta^{(w)})$ is a totally ordered set. The source mapping s: $E \to V$ and range mapping r: $E \to V'$ are defined for $\epsilon \in \Lambda(\zeta^{(w)})$ by

$$s(\epsilon) = v \quad \text{and} \quad r(\epsilon) = w,$$

where $\mathcal{P}(\eta^{(w)})_\epsilon \subset E(v)$. For $1 \le v \le m$ and $1 \le w \le n$ set

$$M_{v,w} = \#\{\epsilon : \mathcal{P}(\eta^{(w)})_\epsilon \subset E(v)\}$$

and for $\epsilon \in \Lambda(\eta^{(w)})$ with $\mathcal{P}(\eta^{(w)})_\epsilon \subset E(v)$,

$$P_{v,\epsilon} = \frac{\nu(\mathcal{P}(\eta^{(w)})_\epsilon)}{\nu(E(v))}.$$

Definition 5.2: We call M the **adjacency matrix** and P the **stochastic matrix associated with** ζ.

LEMMA 5.1: *The above definitions are consistent with the statement that V and V' are two successive sets of vertices of a Bratteli–Vershik diagram, that E is the edge set between these sets of vertices, that M is the adjacency matrix, and that P is a stochastic matrix in the sense of Definition 2.5.*

We can now calculate the expression $\Theta(v,w)$ of Proposition 3.3 for the Markov matrices in the extension of a multiple tower. In fact, we have:

PROPOSITION 5.2: *In the notation of this section, let ζ be the refinement of ξ by η and the measures μ, ν and μ_ν and the stochastic matrix P be as above. Then for $v \in V$ and $w \in V'$, we have*

$$\Theta(v,w) = \frac{\mu_\nu(\mathrm{supp}(\zeta^w) \cap \mathrm{supp}(\xi^v))}{\mu_\nu(\mathrm{supp}(\zeta^w))\mu_\nu(\mathrm{supp}(\xi^v))}.$$

Proof: By definition,

$$\begin{aligned}\Theta^{(k)}(v,w) =& \frac{1}{\nu_k(w)} \sum_{\{x \in E_k : s(x)=v, r(x)=w\}} \frac{\nu(\mathcal{P}(\eta^{(w)})_x)}{\nu(E(v))} \\ =& \frac{1}{\mu_\nu(\mathrm{supp}(\zeta^w))} \cdot \frac{\nu(\mathrm{supp}(\eta^w) \cap E(v))}{\nu(E(v))}.\end{aligned}$$

Since $\xi^{(v)}$ is of constant Jacobian,

$$\begin{aligned}\mu(\mathrm{supp}(\zeta^{(w)}) \cap \mathrm{supp}(\xi^{(v)})) =& \mu_\nu(\cup_\beta \xi^{(v)}_{\beta,\alpha^{(v)}}(\mathrm{supp}(\zeta^{(w)}) \cap E(v))) \\ =& \sum_\beta C_\beta \mu_\nu(\mathrm{supp}(\zeta^{(w)}) \cap E(v)) \\ =& \sum_\beta C_\beta \mu_\nu(\mathrm{supp}(\eta^{(w)}) \cap E(v)),\end{aligned}$$

for some constants $C_\beta > 0$.

Similarly, $\mu_\nu(\mathrm{supp}(\xi^{(v)}) = \sum_\beta C_\beta \mu_\nu(E(v))$. Hence we have

$$\frac{\nu(\mathrm{supp}(\eta^w) \cap E(v))}{\nu(E(v))} = \frac{\mu(\mathrm{supp}(\zeta^{(w)}) \cap \mathrm{supp}(\xi^{(v)}))}{\mu_\nu(\mathrm{supp}(\xi^{(v)}))}.$$

This completes the proof. ∎

The statement that $\Theta(v, w)$ is close to 1 can thus be seen as a kind of weak independence condition for the measure μ_ν with respect to the supports of ζ and ξ. Note that this condition in no way implies that the individual edges are independent. The condition is a natural analogue of the statement in classical Perron–Frobenius theory that one long block of transitions is almost independent from another subsequent long block. (Indeed, in that setting, a refinement of a multiple tower is exactly the process of passing from one block of edges to the next.)

6. Approximation by ordered multiple towers

In this section, we give the proof of Theorem 1.1. The basic idea is to approximate our non-singular transformation by a sequence of multiple towers, which converge, as in the previous section, to a Bratteli–Vershik system. Of course, we have to keep track of the measures as well, and to show that the resulting system is uniquely ergodic in the sense of Section 3.

Throughout this section, $(X, \mathcal{B}, \mu, T)$ will denote an ergodic non-singular dynamical system of type III. It follows that μ is nonatomic and conservative for the action of T.

Our basic tool is the following well-known version of Rokhlin's Lemma.

Lemma 6.1: *Let T be an ergodic non-singular transformation. Then for any $A_1, \ldots, A_k \in \mathcal{B}$ of positive measure, $\varepsilon > 0$ and $N \geq 1$, there exists a measurable subset E of positive measure satisfying the following conditions:*

(i) *$E, TE, T^2E, \ldots, T^{N-1}E$ are disjoint.*

(ii) $\mu(\bigcup_{i=0}^{N-1} T^i E) > 1 - \varepsilon$.

(iii) *Each set A_i is approximated by a union of the sets $\{T^k E\}$ in the sense that the measure of the symmetric difference is less than ε.*

The standard proof of the Dye–Krieger representation theorem (see [HO]) uses this lemma to approximate a dynamical system by an odometer as follows. First, we use the lemma to choose a single tower $\xi = \{\xi_{\alpha,\beta} : 0 \leq \alpha, \beta < N\}$. Then, choosing any level $\mathcal{P}(\xi)_\alpha$ of ξ, we consider the induced transformation of T on that level. One then applies the Lemma again, to obtain a tower $\eta = \{\eta_{i,j} : i, j \in \Theta\}$ whose support is a subset of the level which is close to the level (in the sense that the measure of the symmetric difference is small). The refinement of this tower by the tower η then has support close to $\mathrm{supp}(\xi)$, and is a much better approximation than ξ.

Repeating this process, in the limit one finally obtains a measurable subset F of positive measure so that if we restrict the refinements of towers to F, the orbit of the induced transformation $T|_F$ is a union of orbits of these refinements of towers and the restriction of the whole σ-algebra to F is approximated by the levels of these towers.

Using the above sequence of refinements of towers, one then constructs a measure space isomorphism from F to a full odometer space so that the conjugacy image of $T|_F$ has the same full group as that of the full odometer. However, in this proof, we lose control of the push-forward of μ by the measure space isomorphism.

Our proof of Theorem 1.1 follows the same path as that outlined in the previous paragraph, but instead of single towers, we use multiple towers, and show how the sequence of the refinements of these multiple ordered towers of F can be taken to have constant Jacobian. Hence we will obtain a measure space isomorphism under which a generator S of $[T|_F]$ (i.e., $[S] = [T|_F]$) is conjugate to a Markov odometer and for which the push-forward measure of the normalized measure of the restriction of μ on F is the corresponding Markov measure.

The proof of the Theorem will be preceded by a series of lemmas concerning refinements of towers. These use the Hurewicz ergodic theorem and Lusin's theorem of measure theory, to guarantee that we have the quasi-independence condition of Proposition 5.2, together with Proposition 3.3, and hence unique ergodicity. Rokhlin's lemma allows successive refinements of the towers, while keeping control of the measures.

We shall also show that the diagrams we construct have unique infinite maximal and minimal paths, so that the Vershik transformation is uniquely defined. In order to keep track of this, we shall use the following notation. Choose an ordered multiple tower of constant Jacobian,

$$\xi = \{\xi^{(v)}\}_{v\in V}, \quad \text{where } \xi^v = \{\xi^{(v)}_{e,e'} : e, e' \in \Lambda^{(v)}\}.$$

Recall that $\Lambda^{(v)}$ is a totally ordered set; denote its maximum and minimum elements by $e_{\max}(v)$ and $e_{\min}(v)$. We let $\alpha^{(v)} = e_{\min}(v)$, set $E(v) = \mathcal{P}(\xi(v))_{\alpha^{(v)}}$ and $E = \bigcup_{v\in V} E(v)$.

Let S denote the induced transformation $T|_E$ of T on E.

LEMMA 6.2: *Let $\varepsilon > 0$. There exists a subset E_0 of E of measure less than ε and there exists $N > 0$ such that for $x \notin E_0$, and for all $n > N$,*

$$e^{-\varepsilon}\frac{\mu(E(v))}{\mu(E)} \le \frac{\sum_{i=0}^{n-1}\frac{dS^i\mu}{d\mu}(x)\mathbf{1}_{E(v)}(S^i x)}{\sum_{i=0}^{n-1}\frac{dS^i\mu}{d\mu}(x)} \le e^{\varepsilon}\frac{\mu(E(v))}{\mu(E)}.$$

Proof: Since μ is nonatomic and conservative for the T-action, we may apply the Hurewicz ergodic theorem to see that the expression in the central inequality tends to $\mu(E(v))/\mu(E)$ almost everywhere as $n \to \infty$. By Lusin's theorem, almost everywhere convergence implies uniform convergence off a set of arbitrarily small measure. That is, we may delete from E a set E_0 of measure less than ε such that the convergence is uniform outside E_0. The Lemma now follows. ∎

Now we may apply Rokhlin's Lemma to the transformation S, the integer N and the set $E - E_0$. We obtain

LEMMA 6.3: *There exists a subset F of $E - E_0$, of positive measure, a finite set of integers $\{N_w : w \in W\}$, with each $N_w \ge N$, and a finite partition $\{F_w : w \in W\}$ of F such that:*

(i) *The sets $\{S^j F_w : 0 \le j < N_w, w \in W\}$ are disjoint subsets of E.*

(ii) *Each $S^j F_w$ is included in some $E(v)$.*

(iii) *For each $v \in V$ and $w \in W$, $|\{j : S^j F_w \subset E(v)\}| \ge |V|$.*

(iv) *For all $x \in F_w, w \in W$,*

$$e^{-\varepsilon}\frac{\mu(S^j F_w)}{\mu(F_w)} \le \frac{dS^j\mu}{d\mu}(x) \le e^{\varepsilon}\frac{\mu(S^j F_w)}{\mu(F_w)}.$$

(v)

$$e^{-\varepsilon} \le \frac{\mu(\cup_{w\in W}\cup_{j=0}^{N_w-1} S^j F_w)}{\mu(E)} \le e^{\varepsilon}.$$

The proof of this lemma is a standard application of Rokhlin's Lemma. ∎

Keeping the notation from the previous two lemmas, we now define the ordered multiple tower $\eta = \{\eta^w : w \in W\}$, supported in $E - E_0$, as follows:

(i) $\mathcal{P}(\eta^{(w)})_i := S^i F_w$, $0 \le i < N_w$.

(ii) $\eta^w_{i+1,i}(x) = Sx$, $x \in S^i F_w$.

(iii) $\eta^{(w)} = \{\eta^{(w)}_{i,j} : 0 \le i,j < N_w\}, w \in W$.

Before proceeding with the construction of η — we have yet to equip it with an order and a measure under which it is of constant Jacobian — let us record:

LEMMA 6.4:

$$e^{-\varepsilon}\frac{\mu(E(v))}{\mu(E)} \le \frac{\mu(\operatorname{supp}(\eta^{(w)}) \cap E(v))}{\mu(\operatorname{supp}(\eta^{(w)}))} \le e^{\varepsilon}\frac{\mu(E(v))}{\mu(E)}.$$

Proof: To see this, we start from the inequality in Lemma 6.2 for N_w, multiply throughout by the denominator of the centre term, $\sum_{i=0}^{n-1} \frac{dS^i\mu}{d\mu}(x)$, and integrate over F_w. We obtain an inequality of the form $e^{-\varepsilon}A \le B \le e^{\varepsilon}A$, where

$$A = \int_{F_w} \frac{\mu(E(v))}{\mu(E)} \sum_{i=0}^{n-1} \frac{dS^i\mu}{d\mu}(x)d\mu(x) = \frac{\mu(E(v))}{\mu(E)}\mu(\operatorname{supp}(\eta^w))$$

and

$$B = \int_{F_w} \sum_{i=0}^{n-1} \frac{dS^i\mu}{d\mu}(x)\mathbf{1}_{E(v)}(S^i x)d\mu(x) = \mu(\operatorname{supp}(\eta^w) \cap E(v)).$$

The Lemma now follows. ∎

Now, continuing with our construction of η, we equip it with an order:

(iv) Choose and fix any two vertices v, $v' \in V$.

For each $w \in W$, assign a total order on the set $\Lambda^{(w)} = \{i : 0 \le i < N_w\}$ such that

$$\mathcal{P}(\eta^{(w)})_{e_{\min}(w)} = S^i F_w \subset E(v) \quad \text{and}$$
$$\mu(S^i F_w) \le \frac{1}{|V|}\mu(\operatorname{supp}(\eta^{(w)}) \cap E(v))$$

and

$$\mathcal{P}(\eta^{(w)})_{e_{\max}(w)} = S^j F_w \subset E(v') \quad \text{and}$$
$$\mu(S^j F_w) \le \frac{1}{|V|}\mu(\operatorname{supp}(\eta^{(w)}) \cap E(v')).$$

Finally, we need to define a measure ν for which η is of constant Jacobian:

(v) Take measure equivalent to μ on $\mathrm{supp}(\eta)$, defined, for $A \subset F_w$ and $0 \le i < N_w$, by

$$\nu(S^i A) = \frac{\mu(A)}{\mu(\mathrm{supp}(\eta))} \frac{\mu(S^i F_w)}{\mu(F_w)}.$$

It is easy to see, with this definition, that for $x \in \mathrm{supp}(\eta)$

$$e^{-\varepsilon} \le \frac{d\nu}{d\mu}(x) \le e^{\varepsilon}.$$

This completes the construction of η. Now let us consider the refinement ζ of ξ by η, as constructed in Section 5. We claim that

LEMMA 6.5: *For all $v \in V, w \in W$*

$$e^{-7\varepsilon}\mu_\nu(\mathrm{supp}(\xi^v)) \le \frac{\mu(\mathrm{supp}(\zeta^w) \cap \mathrm{supp}(\xi^v))}{\mu_\nu(\mathrm{supp}(\zeta^w))} \le e^{7\varepsilon}\mu_\nu(\mathrm{supp}(\xi^v)).$$

Proof: Combining the inequality of Lemma 6.4 with the above estimate for the Radon–Nikodým derivative $\frac{d\nu}{d\mu}$, we obtain an inequality of the form $e^{-3\varepsilon}A_1 \le B_1 \le e^{3\varepsilon}A_1$, where

$$A_1 = \frac{\nu(E(v))}{\nu(\mathrm{supp}(\eta))} \quad \text{and} \quad B_1 = \frac{\nu(\mathrm{supp}(\eta^w) \cap E(v))}{\nu(\mathrm{supp}(\eta^w)}.$$

Since ζ is of constant Jacobian with respect to μ_ν, we have for all $e \in \Lambda^{(v)}$

$$\mu_\nu(\mathrm{supp}(\zeta^w) \cap \mathcal{P}(\xi^v)_e)) = \mu_\nu(\mathrm{supp}(\eta^w) \cap E(v))\frac{\mu(\mathcal{P}(\xi^v)_e)}{\mu(E(v))}.$$

Hence, summing over all $e \in \Lambda^{(v)}$, we obtain

$$\mu_\nu(\mathrm{supp}(\zeta^w) \cap \mathrm{supp}(\xi^v)) = \mu_\nu(\mathrm{supp}(\eta^w) \cap E(v))\frac{\mu(\mathcal{P}(\xi^v))}{\mu(E(v))}.$$

By the previous lemma, the right hand side differs by a factor of at most $e^{3\varepsilon}$ from

$$\frac{\nu(\mathrm{supp}(\eta^w))\nu(E(v))}{\nu(\mathrm{supp}(\eta))} \cdot \frac{\mu(\mathcal{P}(\xi^v))}{\mu(E(v))}.$$

In view of the fact that μ and ν coincide on the minimal edges, the above simplifies to

$$\frac{\nu(\mathrm{supp}(\eta^w))}{\nu(\mathrm{supp}(\eta))}\mu(\mathrm{supp}(\xi^v)).$$

Now summing over all $v \in V$, similar reasoning shows that

$$\mu_\nu(\mathrm{supp}(\zeta^w)) = \sum_{v\in V} \mu_\nu(\mathrm{supp}(\zeta^w) \cap \mathrm{supp}(\xi^v))$$

differs by a factor of at most $e^{3\varepsilon}$ from

$$\frac{\nu(\mathrm{supp}(\eta^w))}{\nu(\mathrm{supp}(\eta))}\sum_{v\in V}\mu_\nu(\mathrm{supp}(\xi^v)).$$

Now, by Lemma 6.3 (iv), the sum differs by a factor of at most e^ε from $\mu(\mathrm{supp}(\xi)) = 1$. Hence, $\mu_\nu(\mathrm{supp}(\zeta^w))$ differs from $\nu(\mathrm{supp}(\eta^w))/\nu(\mathrm{supp}(\eta))$ by a multiple of at most $e^{4\varepsilon}$.

This gives the desired estimate. ∎

Proof of Theorem 1.1: We may assume without loss of generality that the transformation is of type *III*.

Let $\{A_n\}_{n\geq 1}$ be a sequence of measurable sets generating the whole σ-algebra, and such that every set A_n appears infinitely often in the sequence.

Choose a sequence (ε_n) such that $\sum_{n=1}^\infty \varepsilon_n < \infty$. It follows from Lemmas 6.1 and 6.2 that there is a sequence ξ_n of multiple towers with associated vertex sets $\Lambda_n = V_n$ and transformations S_n; and a sequence of measures (μ_n) satisfying the following conditions:

(i) ξ_{n+1} is a refinement of ξ_n, for all $(n \geq 1)$.

(ii) For each n, μ_n is of constant Jacobian with respect to ξ_n.

(iii) μ_{n+1} is a Markov extension of μ_n, $(n \geq 1)$.

(iv) $\mu(\mathrm{supp}(\xi_{n+1})) > (1-\varepsilon_n)\mu(\mathrm{supp}(\xi_n))$.

(v) $e^{-\epsilon_n} < \frac{d\mu_{n+1}}{d\mu_n}(x) < e^{\varepsilon_n}$, $x \in \mathrm{supp}(\xi_{n+1})$.

(vi) For each n, for all $v \in V_n$ and for all $w \in V_{n+1}$,

$$e^{-7\varepsilon_n} \leq \frac{\mu_{n+1}(\mathrm{supp}(\xi_{n+1}^w)\cap \mathrm{supp}(\xi_n^v))}{\mu_{n+1}(\mathrm{supp}(\xi_{n+1}^w))\mu_{n+1}(\mathrm{supp}(\xi_n^v))} \leq e^{7\varepsilon_n}.$$

(vii) For each n, there are two vertices $v_n, v_n' \in V_n$ such that every minimal edge to level V_{n+1} starts from v_n, and every maximal edge to level V_{n+1} starts from v_n'.

(viii) $\mu\{x \in \mathrm{supp}(\xi_n) : S_n x \text{ or } S_n^{-1}x \text{ is not defined}\} \leq \varepsilon_n \mu(\mathrm{supp}(\xi_n))$.

(ix)

$$\mu\{x \in \mathrm{supp}(\xi_n) : T^i x \in \mathrm{Orb}_{S_n}(x), \text{ for all } i,\ -n \leq i \leq n\} > (1-\varepsilon_n)\mu(\mathrm{supp}(\xi_n)).$$

(x) A_n is approximated by a union of levels of ξ_n within ε_n in symmetric difference of measure.

We set $F = \bigcap_{n=1}^\infty \mathrm{supp}(\xi_n)$. Then by condition (iv), $\mu(F) > 0$ and hence, by (viii), together with the Borel–Cantelli lemma, we have that for a.e. $x \in F$

$$\{T|_F^i x : -\infty < i < \infty\} = \{S^i x : -\infty < i < \infty\},$$

where S is the associated transformation of ξ_n which is consistently defined for n on the set F.

Define

$$d\nu(x) = \prod_{n=1}^{\infty} \frac{d\mu_{n+1}}{d\mu_n}(x) d\mu(x), \quad x \in F;$$

then the measure ν is equivalent to μ on F and the restriction of each ξ_n on F is of constant Jacobian with respect to ν.

It is clear by (x) that one has a measure space isomorphism from the set F to the Bratteli diagram X constructed from the ordered multiple towers ξ_n, each restricted to F. Furthermore, this diagram is essentially simple by condition (vii). Moreover, by our construction, S is conjugate with the Vershik transformation, and the push-forward measure of μ by the conjugacy is a Markov measure on X. By condition (vi) above, combined with Proposition 3.3 and Proposition 5.2, the action of the associated Vershik transformation on X is uniquely ergodic for this measure in the sense of G-measures.

Since any ergodic non-singular transformation of type III is orbit equivalent with any of its induced transformations, T and S are orbit equivalent with each other. Thus the proof of the theorem is complete. ∎

Actually, we have shown a little more. The Bratteli–Vershik system we have constructed as a model for $(X, \mathcal{B}, \mu, \mathrm{T})$ has the property that every vertex at level n is connected to every vertex at level $n+1$. It is not hard to see that this implies that the only closed invariant sets in X are the empty set and X itself. Furthermore, the condition that every maximal (respectively minimal) edge in $E(n)$ has a common source, v_n' (resp. v_n) guarantees by Corollary 2.1 that the system is topologically induced as a closed subset of a full odometer.

Thus we have shown:

COROLLARY 6.1: *The Markov odometer constructed in the proof of Theorem 1.1 is a minimal homeomorphism for the topology of X. Furthermore, it is a topologically induced transformation of a full odometer.*

References

[B] O. Bratteli, *Inductive limits of finite dimensional C^*-algebras*, Transactions of the American Mathematical Society **171** (1972), 195–234.

[BD] G. Brown and A. H. Dooley, *Odometer actions on G-measures*, Ergodic Theory and Dynamical Systems **11** (1991), 279–307.

[BD1] G. Brown and A. H. Dooley, *On G-measures and product measures*, Ergodic Theory and Dynamical Systems **18** (1998), 95–107.

[Co] A. Connes, *Une classification des facteurs de Type III*, Annales Scientifiques de l'École Normale Supérieure **6** (1973), 133–252.

[CFW] A. Connes, J. Feldman and B. Weiss, *An amenable equivalence relation is generated by a single transformation*, Ergodic Theory and Dynamical Systems **1** (1981), 431–450.

[CW] A. Connes and J. Woods, *Approximately transitive flows and ITPFI factors*, Ergodic Theory and Dynamical Systems **5** (1985), 203–236.

[DH1] A. H. Dooley and T. Hamachi, *Markov odometer actions not of product type*, Ergodic Theory and Dynamical Systems **23** (2003), 1–17.

[DR] A. H. Dooley and D. J. Rudolph, *Non-uniqueness in G-measures*, submitted.

[Do] A. H. Dooley, *Markov odometers*, in *Topics in Dynamical Systems and Ergodic Theory* (S. Bezuglyi and S. Kolyada), Cambridge University Press, 2003.

[D1] H. Dye, *On groups of measure-preserving transformations I*, Transactions of the American Mathematical Society **81** (1959), 119–159.

[D2] H. Dye, *On groups of measure-preserving transformations II*, American Journal of Mathematics **85** (1963), 551–576.

[Ha] T. Hamachi, *A measure theoretical proof of the Connes–Woods theorem on AT-flows*, Pacific Journal of Mathematics **154** (1992), 67–85.

[HO] T. Hamachi and M. Osikawa, *Ergodic groups of automorphisms and Krieger's theorem*, Seminar report, **3**, Mathematical Sciences, Keio University, 1981.

[HPS] R. Herman, I. Putnam and C. Skau, *Ordered Bratteli diagrams, dimension groups and topological dynamics*, International Journal of Mathematics **6** (1992), 827–864.

[Ke] M. S. Keane, *Strongly mixing g-measures*, Inventiones Mathematicae **16** (1972), 827–864.

[Kr1] W. Krieger, *On non-singular transformations of measure spaces I*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **11** (1969), 83–97.

[Kr2] W. Krieger, *On non-singular transformations of measure spaces II*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **11** (1969), 98–119.

[Kr] W. Krieger, *On the infinite product construction of non-singular transformations of a measure space*, Inventiones Mathematicae **15** (1972), 144–163; *erratum*, Inventiones Mathematicae **26** (1974), 323–328.

[KW] Y. Katznelson and B. Weiss, *The classification of nonsingular actions revisited*, Ergodic Theory and Dynamical Systems **11** (1991), 531–543.

[LM] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, 1995.

[LV] A. A. Lodkin and A. M. Vershik, *Approximation of amenable groups and transversal automorphisms*, Lecture Notes in Mathematics **1132**, Springer-Verlag, Berlin, 1985, pp. 331–346.

[O] N. Ormes, *Strong orbit realisations for minimal homeomorphisms*, Journal d'Analyse Mathématique **71** (1997), 103–133.

[OW1] D. Ornstein and B. Weiss, *Entropy and isomorphism theorems for actions of amenable groups*, Journal d'Analyse Mathématique **48** (1987), 1–141.

[Q] A. Quas, *Non-ergodicity for C^1-expanding maps and g-measures*, Ergodic Theory and Dynamical Systems **16** (1996), 531–543.

[VL] A. M. Vershik and A. N. Livshits, *Adic models of ergodic transformations, spectral theory, substitutions and related topics*, in *Representation Theory and Dynamical Systems*, Advances in Soviet Mathematics **9** (1992), 185–204.