

HILBERT SPACES OF MARTINGALES SUPPORTING CERTAIN SUBSTITUTION-DYNAMICAL SYSTEMS

DORIN ERVIN DUTKAY AND PALLE E. T. JORGENSEN

Dedicated to the memory of Shizuo Kakutani

ABSTRACT. Let X be a compact Hausdorff space. We study finite-to-one mappings $r: X \rightarrow X$, onto X , and measures on the corresponding projective limit space $X_\infty(r)$. We show that certain quasi-invariant measures on $X_\infty(r)$ correspond in a one-to-one fashion to measures on X which satisfy two identities. Moreover, we identify those special measures on $X_\infty(r)$ which are associated via our correspondence with a function V on X , a Ruelle transfer operator R_V , and an equilibrium measure μ_V on X .

CONTENTS

1. Introduction
 2. Projective limits
 - 2.1. Definitions
 - 2.2. A scaling equation
 3. Induction of measures on $X_\infty(r)$ from measures on X
 - 3.1. Martingales
 - 3.2. From μ on X to $\hat{\mu}$ on $X_\infty(r)$
 - 3.3. Quasi-invariance
 4. A fixed-point problem
 5. Transformations of measures
 6. Extreme points
- Acknowledgments
References

1. INTRODUCTION

This paper is motivated by our desire to apply wavelet methods to some nonlinear problems in symbolic and complex dynamics. Recent research by many authors (see, e.g., [4, 2, 6]) on iterated function systems with affine scaling has suggested that the scope of the multiresolution method is wider than the more traditional wavelet context.

Received by the editors January 23, 2005 and, in revised form, February 2, 2005.

2000 *Mathematics Subject Classification*. Primary 42C40, 42A16, 42A65, 43A65, 46G15, 47D07, 60G18.

Key words and phrases. Measures, projective limits, transfer operator, martingale, fixed point.

©2005 American Mathematical Society
Reverts to public domain 28 years from publication

In this paper we exploit the connections between multiscale-resolution analysis (wavelet theory) and dynamics, in particular iteration theory. For this purpose, we study certain quasi-invariant measures, extending the familiar scaling equation from wavelet theory. Our proofs use martingale theory, a transfer operator, called the Ruelle transfer operator, and a fixed point theorem. A by-product of our analysis is Theorem 5.9, generalizing a theorem of Perron and Frobenius.

As suggested first by S. Mallat [29], a particularly productive approach to the construction of wavelet bases in $L^2(\mathbb{R})$ may be based on the notion of resolution from optics. In this model, a fixed image-resolution corresponds to a certain closed subspace V_0 in $L^2(\mathbb{R})$. A coarser resolution will be a proper subspace of V_0 , and a finer resolution is a subspace containing V_0 . What results is a \mathbb{Z} -indexed scale of nested Hilbert subspaces V_n , $n \in \mathbb{Z}$, in $L^2(\mathbb{R})$ such that the intersection is $\{0\}$ and the union is dense. Moreover the operation of dyadic scaling transforms each V_n to the next V_{n+1} . This is called a multiresolution approach to wavelets, see [12], and it is based on the interplay between the two abelian groups \mathbb{T} (the circle group = one-torus), and \mathbb{R} (the real line), with \mathbb{T} representing a period interval placed on the line \mathbb{R} .

Our present paper is based on the observation that multiresolutions really are martingales; and on our exploiting this fact. Thus we are able to adapt the geometric notion of scaling subspaces (V_n), or resolutions, to nonlinear dynamics in a variety of applications where such a pair of groups is not available. In the context of discrete dynamical systems, we have instead an endomorphism $r: X \rightarrow X$ of a compact space. And then successive iterations of r may serve to define a certain selfsimilarity (for example statistical similarity up to scale in time or in space), mirroring the more familiar notion of scale-similarity which is such a powerful tool in wavelet theory. The usefulness of this geometric viewpoint is illustrated for example in multivariable operator theory (e.g., [10], [23], [25], [26], [33], and [34]), as well as in recent work on encoding of quantum stochastic processes; see, e.g., [1], [3], [13], [18], [28], and [30].

The notion of multiresolution (from wavelets) is in fact closely related to more familiar concepts from dynamics, e.g., to sequences of partitions or tower constructions in dynamics, such as laminations [39], or successive subdivisions [40]; to Pesin boxes [9], to attractors [38].

Let r be an endomorphism in a compact metric space X (for example the Julia set of a given rational map), and suppose r is onto X and finite-to-one. Form a projective space X_∞ such that r induces an automorphism \hat{r} of X_∞ . Let V be a Borel function on X (naturally extended to a function on X_∞). Generalizing the more traditional approach to scaling functions, we give in Theorem 4.1 a complete classification of measures on X_∞ which are quasi-invariant under \hat{r} and have Radon-Nikodym derivative equal to V ; see also (1.1) and Section 2.2 for further details.

The interest in quasi-invariance derives in part from the observation that in the more familiar wavelet case, this invariance notion plays a fundamental role, that of a scaling equation (see [12, Chapter 5]). But in the wavelet case, the measure is prescribed *a priori*; it is the Lebesgue measure. Not so for other scaling problems in dynamics. Starting with a possibly nonlinear system (X, r) , we must examine measures μ on X which are suitably invariant for the inverse branches of the given endomorphism $r: X \rightarrow X$. This invariance amounts to equation (2.4) below, and we typically assume that μ (if it exists) is normalized. If a probability measure μ

exists, solving equation (2.4), we say that μ is strongly invariant. But existence of strongly invariant measures is now a question. (If for example $r: X \rightarrow X$ is known to have strictly contractive inverse branches [20], or if X is the Julia set of some rational map, $w = r(z)$, then we know that strongly invariant measures μ exist [11], but not in general.) However, assuming that a strongly invariant measure μ does exist, then the statement of the scaling equation involves $\hat{\mu}$, and it takes the form of

$$(1.1) \quad \frac{d(\hat{\mu} \circ \hat{r})}{d\hat{\mu}} = V \circ \theta_0$$

where θ_0 is the X -coordinate function on X_∞ . (Both sides in equation (1.1) are functions on X_∞ . The function V on the right-hand side is defined on X , so the composition is indeed a function on X_∞ .)

Hence there are three separate problems: (1) Given such a measure μ , find the functions V such that (1.1) holds. (2) For what functions V are there measures μ on X such that the corresponding $\hat{\mu}$ on X_∞ satisfies (1.1). (3) Starting with a measure ν on X_∞ for which the scaling equation (1.1) holds, then when is there a measure μ on X such that $\nu = \hat{\mu}$.

Our earlier paper [17] addressed (1); see also Theorem 2.1 below. Here we turn instead to questions (2) and (3).

We show in Section 2 that all three problems relate to a certain eigenvalue problem for a transfer operator R_V defined by equation (2.5) in terms of the function V . The operator R_V is of general interest in operator theory; see e.g., [8], [19], [22], [32], and [37].

Our analysis of the quasi-invariant measures is based on certain Hilbert spaces of martingales (Theorem 3.3) and on a transfer operator (Section 2.2) studied first by David Ruelle. In Theorem 6.3 we give a characterization of the extreme points in the set of V -quasi-invariant probability measures.

Example 1.1. We now give an example of an endomorphism r_A in a compact space $X(A)$ which is onto, and has finite pre-images. The construction is as follows:

Let A be a $k \times k$ matrix with entries in $\{0, 1\}$. Suppose every column in A contains an entry 1. (This guarantees that the restricted shift r_A below maps onto $X(A)$).

Set

$$X(A) := \left\{ (\xi_i)_{i \in \mathbb{N}} \in \prod_{\mathbb{N}} \{1, \dots, k\} \mid A(\xi_i, \xi_{i+1}) = 1 \right\}$$

and

$$r_A(\xi_1, \xi_2, \dots) = (\xi_2, \xi_3, \dots) \text{ for } \xi \in X(A).$$

Then r_A is a subshift, and the pair $(X(A), r_A)$ satisfies our conditions.

It is known [37] that, for each A , as described above, the corresponding system $r_A: X(A) \rightarrow X(A)$ has a unique strongly r_A -invariant probability measure, ρ_A , i.e., a probability measure on $X(A)$ such that

$$\int_{X(A)} f d\rho_A = \int_{X(A)} \frac{1}{\#r_A^{-1}(x)} \sum_{r_A(y)=x} f(y) d\rho_A(x),$$

for all bounded measurable functions f on $X(A)$.

In this paper we analyze the connection between measures on X and the induced measures on X_∞ , and we characterize those measures X_∞ which are quasi-invariant with respect to the invertible mapping \hat{r} ,

$$\begin{array}{ccc} X_\infty & \xrightarrow{\hat{r}} & X_\infty \\ \downarrow & & \downarrow \\ X & \xrightarrow{r} & X \end{array}$$

where

$$(1.2) \quad X_\infty := \left\{ \hat{x} = (x_0, x_1, \dots) \in \prod_{\mathbb{N}_0} X \mid r(x_{n+1}) = x_n, n \in \mathbb{N}_0 \right\},$$

$$\hat{r}(\hat{x}) = (r(x_0), x_0, x_1, \dots)$$

and

$$\hat{r}^{-1}(\hat{x}) = (x_1, x_2, \dots).$$

We recall the simplest instance in the wavelet context: As described in [12, Chapter 5], wavelets may be obtained from an X_∞ construction in the special case when $X = \mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$, and $r(z) = z^2$, as follows:

While the real line \mathbb{R} is not \mathbb{T}_∞ , we may still build the Hilbert space $L^2(\mathbb{R})$ with its multiresolution wavelets from an inductive and isometric procedure based on \mathbb{T}_∞ . The intuitive idea is to get \mathbb{R} in the limit by successive doubling of periods. Following [19], we are in fact using a recursive system of Hilbert spaces for this, and we note that this is really a martingale construction. To highlight the distinction between the compact group $G = \mathbb{T}_\infty$ (called a solenoid) and the reals \mathbb{R} , contrast the following two familiar short exact sequences of abelian groups, $0 \rightarrow C \rightarrow G \rightarrow \mathbb{T} \rightarrow 0$, where C is the Cantor group, on the one hand; and $0 \rightarrow \mathbb{Z} \rightarrow \mathbb{R} \rightarrow \mathbb{T} \rightarrow 0$, on the other. While the groups G and \mathbb{R} are very different, the use of G helps us even in this case to build the Hilbert space $L^2(\mathbb{R})$; but within the category of isometries in Hilbert space. Moreover, the X_∞ viewpoint is useful in important applications outside the group context, and this is the focus of the present paper.

We need a specific interplay between spaces (X, r) , $r: X \rightarrow X$ a non-invertible endomorphism, and induced spaces (X_∞, \hat{r}) where \hat{r} is an automorphism. This issue, and variants of it, arise in a number of areas of mathematics; first in probability theory, going back to [27]; and also more recently in a number of wavelet problems; see for example [5], [7], [8], [15], [16], [17], [19], [21], and [22]. In these applications, the problem is to carry along some isometric operator defined on a Hilbert space of functions on X , to the space X_∞ (see Proposition 2.2 below). In the language of operator theory, we wish to make a covariant unitary dilation from X to X_∞ . (See e.g., [24] and [30].) This means that we need to induce measures μ on X to measures $\hat{\mu}$ on X_∞ in such a way that a prescribed covariance is preserved.

To make our paper self-contained, we have recalled Doob's martingale convergence Theorem in Section 3, in the form in which we need it, but we refer the reader to the books [14] and [31] for background on martingale theory.

2. PROJECTIVE LIMITS

2.1. Definitions. Let X be a compact Hausdorff space, and let $r: X \rightarrow X$ be a finite-to-one mapping:

$$1 \leq \#r^{-1}(x) < \infty$$

where $r^{-1}(x) = \{y \in X \mid r(y) = x\}$. More generally, set

$$r^{-1}(E) = \{y \in X \mid r(y) \in E\}, \quad \text{if } E \subset X.$$

By the projective limit $X_\infty(r)$, we mean

$$X \xleftarrow[r]{} X \xleftarrow[r]{} X \xleftarrow[r]{} \cdots \xleftarrow[r]{} X_\infty(r);$$

see also (1.2).

We will further assume that r is not invertible, i.e., that $\#r^{-1}(x)$ is not the constant function one.

It is well known that, if r is continuous, then pull-backs of open sets in X define a topology on $X_\infty(r)$ making $X_\infty(r)$ compact.

The restriction to $X_\infty(r)$ of the coordinate projections $(x_0, x_1, \dots) \mapsto x_n$ will be denoted by θ_n , and we have

$$(2.1) \quad r \circ \theta_{n+1} = \theta_n \quad (n \in \mathbb{N}_0).$$

One advantage of passing from X to $X_\infty(r)$ is that r induces an invertible mapping $\hat{r}: X_\infty(r) \rightarrow X_\infty(r)$, such that

$$r \circ \theta_n = \theta_n \circ \hat{r} = \theta_{n-1}.$$

Moreover, both \hat{r} and \hat{r}^{-1} are continuous if r is.

If \mathfrak{B} is a sigma-algebra of subsets in X (typically we will take \mathfrak{B} to be the Borel subsets in X), then there are sigma-algebras

$$\mathfrak{B}_n := \theta_n^{-1}(\mathfrak{B}) = \{\theta_n^{-1}(E) \mid E \in \mathfrak{B}\}$$

and

$$(2.2) \quad \mathfrak{B}_\infty = \bigcup_{n \in \mathbb{N}_0} \mathfrak{B}_n.$$

These \mathfrak{B}_n 's are sigma-algebras of subsets of $X_\infty(r)$.

Using (2.1), we get

$$\theta_n^{-1} = \theta_{n+1}^{-1} \circ r^{-1}.$$

If r is measurable, then $r^{-1}(\mathfrak{B}) \subset \mathfrak{B}$, and we conclude that

$$(2.3) \quad \mathfrak{B}_n \subset \mathfrak{B}_{n+1}.$$

Then both of the mappings \hat{r} and \hat{r}^{-1} are measurable on $X_\infty(r)$ with respect to the sigma-algebra \mathfrak{B}_∞ .

2.2. A scaling equation. In [17], we studied the following restrictive setup: we assumed that X carries a probability measure μ which is *strongly r -invariant*. By this we mean that

$$(2.4) \quad \int_X f d\mu = \int_X \frac{1}{\#r^{-1}(x)} \sum_{y \in X, r(y)=x} f(y) d\mu(x) \quad (f \in L^\infty(X)).$$

If, for example $X = \mathbb{R}/\mathbb{Z}$, and $r(x) = 2x \bmod 1$, then the Haar measure on $\mathbb{R}/\mathbb{Z} =$ Lebesgue measure on $[0, 1)$, is the unique strongly r -invariant measure on X . We prescribe probabilities via a fixed function $V: X \rightarrow [0, \infty)$.

Suppose V is bounded and measurable. Then define $R = R_V$, the *Ruelle operator*, by

$$(2.5) \quad R_V f(x) = \frac{1}{r^{-1}(x)} \sum_{r(y)=x} V(y)f(y) \quad (f \in L^1(X, \mu)).$$

Theorem 2.1. ([17]) *Let $r: X \rightarrow X$ and $X_\infty(r)$ be as described in Section 2.1, and suppose that X has a strongly r -invariant measure μ . Let V be a non-negative, measurable function on X , and let R_V be the corresponding Ruelle operator.*

(i) *There is a unique measure $\hat{\mu}$ on $X_\infty(r)$ such that*

$$(a) \quad \hat{\mu} \circ \theta_0^{-1} \ll \mu \quad (\text{set } h := \frac{d(\hat{\mu} \circ \theta_0^{-1})}{d\mu}),$$

$$(b) \quad \int_X f d\hat{\mu} \circ \theta_n^{-1} = \int_X R_V^n(fh) d\mu \quad (n \in \mathbb{N}_0).$$

(ii) *The measure $\hat{\mu}$ on $X_\infty(r)$ satisfies*

$$(2.6) \quad \frac{d(\hat{\mu} \circ \hat{r})}{d\hat{\mu}} = V \circ \theta_0$$

and

$$R_V h = h.$$

In this paper we turn around the problem and take (2.6) as the fundamental axiom. We will not assume the existence of a strongly r -invariant measure μ .

Proposition 2.2. *Let $r: X \rightarrow X$ be as described above, and let $\hat{r}: X_\infty(r) \rightarrow X_\infty(r)$ be the corresponding automorphism. Let $m: X \rightarrow \mathbb{C}$ be a bounded measurable function on X . Then the operator*

$$(2.7) \quad \hat{S}_m f := (m \circ \theta_0) f \circ \hat{r}$$

defines an isometry in the Hilbert space $L^2(X_\infty(r), \hat{\mu})$ if and only if

$$(2.8) \quad \hat{\mu} \circ \hat{r} \ll \hat{\mu}$$

and

$$(2.9) \quad \frac{d(\hat{\mu} \circ \hat{r})}{d\hat{\mu}} = |m \circ \theta_0|^2.$$

Proof. The isometric property for (2.7) may be stated in the form.

$$(2.10) \quad \int_{X_\infty(r)} |m \circ \theta_0|^2 |f \circ \hat{r}|^2 d\hat{\mu} = \int_{X_\infty(r)} |f|^2 d\hat{\mu}, \text{ for all } f \in L^2(X_\infty(r), \hat{\mu}).$$

Setting $V := |m|^2$, and $g := |f \circ \hat{r}|^2$, (2.10) reads

$$\int_{X_\infty(r)} V \circ \theta_0 g d\hat{\mu} = \int_{X_\infty(r)} g \circ \hat{r} d\hat{\mu}.$$

But this amounts precisely to the two assertions (2.8) and (2.9). In other words, the V -quasi-invariance property for $V = |m|^2$ is equivalent to S_m defining an L^2 -isometry. \square

3. INDUCTION OF MEASURES ON $X_\infty(r)$ FROM MEASURES ON X

3.1. Martingales. In this subsection we introduce the sequence conditional expectations which support our martingale construction of the global Hilbert space.

If $\hat{\mu}$ is a measure defined on the sigma-algebra \mathfrak{B}_∞ on $X_\infty(r)$, then there is an associated sequence of measures (μ_n) on X defined as follows:

$$(3.1) \quad \mu_n := \hat{\mu} \circ \theta_n^{-1} \quad (n \in \mathbb{N}_0),$$

or more precisely

$$\mu_n(E) = \hat{\mu}(\theta_n^{-1}(E)) \quad (E \in \mathfrak{B}),$$

where

$$\theta_n^{-1}(E) = \{\hat{x} \in X_\infty(r) \mid \theta_n(\hat{x}) \in E\}.$$

Our measures will be assumed positive and finite, unless specified otherwise.

We now introduce the Hilbert spaces:

$$(3.2) \quad \hat{H}(\hat{\mu}) := L^2(X_\infty(r), \hat{\mu})$$

and

$$H_n(\hat{\mu}) := \{\xi \circ \theta_n \mid \xi \in L^2(X, \mu_n)\},$$

the orthogonal projections

$$P_n: \hat{H}(\hat{\mu}) \rightarrow H_n(\hat{\mu})$$

and

$$E_n: \hat{H}(\hat{\mu}) \rightarrow L^2(X, \mu_n),$$

setting

$$P_n(f) := E_n(f) \circ \theta_n \quad (n \in \mathbb{N}_0);$$

then

$$\int_{X_\infty(r)} \xi \circ \theta_n f d\hat{\mu} = \int_X \xi E_n(f) d\mu_n, \text{ for all } n \in \mathbb{N}_0, \xi \in L^2(X, \mu_n).$$

Lemma 3.1. *Let $\hat{\mu}$ be a measure on $(X_\infty(r), \mathfrak{B}_\infty)$ and let $\hat{H}(\hat{\mu})$ be the corresponding Hilbert space from (3.2).*

(i) *Then there is an isometric isomorphism*

$$(3.3) \quad J: \hat{H}(\hat{\mu}) \rightarrow \lim_n L^2(X, \mu_n)$$

where the inductive limit in (3.3) is defined by the isometric embeddings

$$\xi \mapsto \xi \circ r: L^2(X, \mu_n) \rightarrow L^2(X, \mu_{n+1}).$$

The isomorphism J in (3.3) is

$$Jf = (E_n f)_{n \in \mathbb{N}_0} \quad (f \in \hat{H}(\hat{\mu})).$$

(ii) *The Hilbert norm is given by*

$$\|f\|_{\hat{H}(\hat{\mu})}^2 = \lim_{n \rightarrow \infty} \|E_n(f)\|_{L^2(X, \mu_n)}^2;$$

and

(iii) *For every sequence $\xi_n \in L^2(X, \mu_n)$ such that $P_n(\xi_{n+1} \circ \theta_{n+1}) = \xi_n \circ \theta_n$, and*

$$\sup_n \|\xi_n\|_{L^2(X, \mu_n)} < \infty,$$

there is a unique $f \in L^2(X_\infty(r), \hat{\mu})$ such that

$$P_n f = \xi_n \circ \theta_n \text{ for all } n \in \mathbb{N}_0.$$

Moreover,

$$f = \lim_{n \rightarrow \infty} \xi_n \circ \theta_n, \text{ pointwise } \hat{\mu}\text{-a.e., and in } \hat{H}(\hat{\mu}).$$

Proof. The proof depends on Doob's martingale convergence theorem; see [14], [17], and [31]. \square

3.2. From μ on X to $\hat{\mu}$ on $X_\infty(r)$. In this subsection we introduce the sequence martingale measures (μ_n) needed in our construction of the global Hilbert space.

Lemma 3.2. *Let $(\mu_n)_{n \in \mathbb{N}_0}$ be a sequence of measures on (X, \mathfrak{B}) . Then there is a measure $\hat{\mu}$ on $(X_\infty(r), \mathfrak{B}_\infty)$ such that*

$$(3.4) \quad \hat{\mu} \circ \theta_n^{-1} = \mu_n \quad (n \in \mathbb{N}_0)$$

if and only if

$$(3.5) \quad \mu_{n+1} \circ r^{-1} = \mu_n \quad (n \in \mathbb{N}_0).$$

Proof. We first recall the definition, the measure $\mu_{n+1} \circ r^{-1}$ in (3.5). On a Borel subset E , it is

$$(3.6) \quad \mu_{n+1} \circ r^{-1}(E) := \mu_{n+1}(r^{-1}(E)).$$

It is well known (see [35]) that (3.6) defines a measure.

Suppose $\hat{\mu}$ exists such that (3.4) holds. Let ξ be a bounded \mathfrak{B} -measurable function on X . Then

$$\int_X \xi \circ r \, d\mu_{n+1} = \int_{X_\infty(r)} \xi \circ r \circ \theta_{n+1} \, d\hat{\mu} = \int_{X_\infty(r)} \xi \circ \theta_n \, d\hat{\mu} = \int_X \xi \, d\hat{\mu} \circ \theta_n^{-1} = \int_X \xi \, d\mu_n,$$

which proves the identity (3.5).

Conversely, suppose $(\mu_n)_{n \in \mathbb{N}_0}$ satisfies (3.5). Note that the \mathfrak{B}_n -measurable functions are of the form $f \circ \theta_n$ with f \mathfrak{B} -measurable. The \mathfrak{B}_n -measurable functions are also \mathfrak{B}_{n+1} -measurable and this inclusion is given by

$$(3.7) \quad f \circ \theta_n = (f \circ r) \circ \theta_{n+1}.$$

The relations (2.2) and (2.3) imply that the union of the algebras of bounded \mathfrak{B}_n -measurable functions is $L^1(\hat{\mu})$ -dense in the algebra of bounded \mathfrak{B}_∞ -measurable functions on $X_\infty(r)$.

Now set

$$\hat{\mu}(f \circ \theta_n) := \mu_n(f), \text{ } f \text{ assumed } \mathfrak{B}_n\text{-measurable and bounded.}$$

To see that this is consistent, use (3.7) and (3.5) to check that

$$\hat{\mu}((f \circ r) \circ \theta_{n+1}) = \mu_{n+1}(f \circ r) = \mu_n(f) = \hat{\mu}(f \circ \theta_n).$$

The existence and uniqueness of $\hat{\mu}$ follows from Kolmogorov's theorem; see [17] and [27]. \square

3.3. Quasi-invariance. In this subsection we state our theorem regarding the scaling equation (1.1) from the Introduction.

In this section we will study the problem (2.6) via our correspondence $\hat{\mu} \leftrightarrow (\mu_n)_{n \in \mathbb{N}_0}$ from Lemma 3.2.

The questions are two: (1) Which measures μ_0 on X admit “extensions” $\hat{\mu}$ to measures on $X_\infty(r)$ satisfying (2.6)?

(2) Understand the measures μ_0 which admit solutions $\hat{\mu}$ to (2.6).

Our first theorem answers the question (1).

Theorem 3.3. *Let $(\mu_n)_{n \in \mathbb{N}_0}$ be a sequence of measures as in Lemma 3.1, i.e., satisfying $\mu_{n+1} \circ r^{-1} = \mu_n$, and let $V: X \rightarrow [0, \infty)$ be \mathfrak{B} -measurable. Then the “extended” measure $\hat{\mu}$ satisfies*

$$(3.8) \quad \frac{d(\hat{\mu} \circ \hat{r})}{d\hat{\mu}} = V \circ \theta_0 \text{ a.e. on } (X_\infty(r), \mathfrak{B}_\infty)$$

if and only if

$$(3.9) \quad d\mu_0 = (V d\mu_0) \circ r^{-1}, \text{ and } d\mu_{n+1} = (V \circ r^n) d\mu_n \quad (n \in \mathbb{N}_0).$$

Proof. Note first that (3.9) is equivalent to the following

$$\int_X f d\mu_0 = \int_X f \circ r V d\mu_0, \text{ and } \int_X f d\mu_{n+1} = \int_X f V \circ r^n d\mu_n \quad (n \in \mathbb{N}_0)$$

for all bounded \mathfrak{B} -measurable functions f on X .

Let $(\mu_n)_{n \in \mathbb{N}_0}$ be a sequence of measures that satisfy

$$(3.10) \quad \mu_{n+1} \circ r^{-1} = \mu_n,$$

and let $\hat{\mu}$ be the corresponding measure on $X_\infty(r)$. Suppose first that $\hat{\mu}$ satisfies (3.8) for some \mathfrak{B} -measurable function $V: X \rightarrow [0, \infty)$; in other words, we have the identity

$$(3.11) \quad \int_{X_\infty(r)} f \circ \hat{r}^{-1} d\hat{\mu} = \int_{X_\infty(r)} f V \circ \theta_0 d\hat{\mu}$$

for all bounded \mathfrak{B}_∞ -measurable functions f on $X_\infty(r)$.

Now specializing (3.11) to $f = \xi \circ \theta_n$ with ξ some bounded \mathfrak{B} -measurable function on X , we get

$$\begin{aligned} \int_{X_\infty(r)} \xi \circ \theta_{n+1} d\hat{\mu} &= \int_{X_\infty(r)} \xi \circ \theta_n V \circ \theta_0 d\hat{\mu} \\ &= \int_{X_\infty(r)} \xi \circ \theta_n V \circ r^n \circ \theta_n d\hat{\mu} = \int_X \xi V \circ r^n d\mu_n, \end{aligned}$$

and therefore

$$d\mu_{n+1} = (V \circ r^n) d\mu_n$$

by (3.1).

We take a closer look at the case $n = 0$. Substitute $f = \xi \circ \theta_0$ into (3.11). We get

$$\int_{X_\infty(r)} \xi \circ \theta_0 d\hat{\mu} = \int_{X_\infty(r)} \xi \circ \theta_0 \circ \hat{r} V \circ \theta_0 d\hat{\mu} = \int_{X_\infty(r)} \xi \circ r \circ \theta_0 V \circ \theta_0 d\hat{\mu} = \int_X \xi \circ r V d\mu_0$$

and therefore $d\mu_0 = (V d\mu_0) \circ r^{-1}$ which is (3.9).

Conversely, suppose $(\mu_n)_{n \in \mathbb{N}_0}$ satisfies (3.10) and (3.9). Lemma 3.2 yields the existence of a $\hat{\mu}$ on $(X_\infty(r), \mathfrak{B}_\infty)$ such that

$$\mu_n = \hat{\mu} \circ \theta_n^{-1}.$$

We claim that $\hat{\mu}$ is V -quasi-invariant, i.e., that (3.8) holds. This amounts to the identity

$$(3.12) \quad \int_{X_\infty(r)} f \circ \hat{r}^{-1} d\hat{\mu} = \int_{X_\infty(r)} f V \circ \theta_0 d\hat{\mu}.$$

By Lemma 3.1, this is equivalent to

$$d\mu_{n+1} = V \circ r^n d\mu_n.$$

To see this, substitute $f = \xi \circ \theta_n$ into (3.12) for ξ a bounded \mathfrak{B} -measurable function on X , and $n \in \mathbb{N}_0$. But this holds by (3.9), and (3.8) follows. \square

4. A FIXED-POINT PROBLEM

In Theorem 3.3, we saw that our condition (2.6), V -quasi-invariance, on measures $\hat{\mu}$ on $(X_\infty(r), \mathfrak{B}_\infty)$ entails a *fixed-point property* (4.1) for the corresponding measure $\mu_0 := \hat{\mu} \circ \theta_0^{-1}$. In this section, we turn the problem around. We show that this fixed-point property characterizes the measures $\hat{\mu}$ satisfying the quasi-invariance, i.e., satisfying (2.6).

If the function $V: X \rightarrow [0, \infty)$ is given, we define

$$V^{(n)}(x) := V(x)V(r(x)) \cdots V(r^{n-1}(x)),$$

and set $d\mu_n := V^{(n)}d\mu_0$. Our result states that the corresponding measure $\hat{\mu}$ on $X_\infty(r)$ is V -quasi-invariant if and only if

$$(4.1) \quad d\mu_0 = (V d\mu_0) \circ r^{-1}.$$

Theorem 4.1. *Let $V: X \rightarrow [0, \infty)$ be \mathfrak{B} -measurable, and let μ_0 be a measure on X satisfying the following fixed-point property:*

$$(4.2) \quad d\mu_0 = (V d\mu_0) \circ r^{-1}.$$

Then there exists a unique measure $\hat{\mu}$ on $X_\infty(r)$ such that

$$(4.3) \quad \frac{d(\hat{\mu} \circ \hat{r})}{d\hat{\mu}} = V \circ \theta_0$$

and

$$\hat{\mu} \circ \theta_0^{-1} = \mu_0.$$

Proof. Let μ_0 be a measure satisfying the fixed-point property (4.2), and set

$$(4.4) \quad d\mu_{n+1} := (V \circ r^n) d\mu_n.$$

This defines a sequence $\mu_0, \mu_1, \mu_2, \dots$ of measures on X . To establish the existence of the measure $\hat{\mu}$ on $X_\infty(r)$, we appeal to Lemma 3.2. We claim that (3.5) holds, or equivalently that

$$(4.5) \quad \int_X f \circ r \, d\mu_{n+1} = \int_X f \, d\mu_n$$

for all bounded \mathfrak{B} -measurable functions f on X , and for all $n \in \mathbb{N}_0$.

Substitution of (4.4) into (4.5) yields

$$\int_X f \circ r \, d\mu_{n+1} = \int_X f \circ r \, V \circ r^n \, d\mu_n,$$

and we will prove (4.5) by induction:

First

$$\int_X f \circ r \, d\mu_1 = \int_X f \circ r \, V \, d\mu_0 = \int_X f \, (V \, d\mu_0) \circ r^{-1} = \int_X f \, d\mu_0,$$

where we used (4.2) in the last step. So (4.5) holds for $n = 0$.

Assume that (4.5) holds for μ_m and $m < n$. Then

$$\int_X f \circ r \, d\mu_{n+1} = \int_X f \circ r \, V \circ r^{n-1} \circ r \, d\mu_n = \int_X f \, V \circ r^{n-1} \, d\mu_{n-1} = \int_X f \, d\mu_n,$$

where we used the definition (4.4) in the last step. The induction is completed.

This proves (3.5). An application of Lemma 3.2 yields the existence of a unique measure $\hat{\mu}$ on $(X_\infty(r), \mathfrak{B}_\infty)$ such that

$$\hat{\mu} \circ \theta_n^{-1} = \mu_n \quad (n \in \mathbb{N}_0).$$

Since the sequence $(\mu_n)_{n \in \mathbb{N}_0}$ satisfies the pair of conditions (3.9) by construction, we conclude from Theorem 3.3 that $\hat{\mu}$ must satisfy (4.3), i.e., $\hat{\mu}$ is V -quasi-invariant as claimed. \square

The proof of Theorem 4.1 also yields the following corollary:

Corollary 4.2. *Let $V: X \rightarrow [0, \infty)$ be a bounded \mathfrak{B} -measurable function, and let μ_0 be a measure on (X, \mathfrak{B}) . Suppose*

$$d\mu_0 = (V \, d\mu_0) \circ r^{-1}.$$

Set

$$(4.6) \quad d\mu_n := V^{(n)} \, d\mu_0, \text{ for } n \in \mathbb{N}.$$

Then

$$(4.7) \quad \mu_{n+1} \circ r^{-1} = \mu_n,$$

and

$$(4.8) \quad \mu_n(X) = \int_X V^{(n)} \, d\mu_0 = \mu_0(X), \text{ for all } n \in \mathbb{N}.$$

Proof. Indeed, by (4.6) we have the recursive formula $d\mu_{n+1} = (V \circ r^n) \, d\mu_n$, and as in the proof of Theorem 4.1, we get that (4.7) is satisfied. Applying (4.6) to the constant function 1, we get that

$$\mu_n(X) = \int_X V^{(n)} \, d\mu_0,$$

and, with (4.7)

$$\mu_{n+1}(X) = \int_X 1 d\mu_{n+1} = \int_X 1 \circ r d\mu_{n+1} = \int_X 1 d\mu_n = \mu_n(X).$$

Now (4.8) follows by induction. \square

5. TRANSFORMATIONS OF MEASURES

Let X be a compact Hausdorff space, \mathfrak{B} the sigma-algebra of all Borel subsets of X , and $C(X)$ the continuous functions on X .

For $a \in \mathbb{R}_+$ denote by

$$M_a(X) := \{\mu \mid \mu \text{ is a measure on } (X, \mathfrak{B}), \mu(X) = a\}.$$

Note that $M_a(X)$ is equipped with the weak*-topology coming from the duality

$$(f, \mu) \mapsto \int_X f d\mu =: \mu(f)$$

where $f \in C(X)$. The neighborhoods are generated by the sets $N_{\mu_0}(f_1, \dots, f_k, \epsilon)$ where $\epsilon \in \mathbb{R}_+$, $f_1, \dots, f_k \in C(X)$ and

$$N_{\mu_0}(f_1, \dots, f_k, \epsilon) = \{\mu \in M(X) \mid |\mu(f_i) - \mu_0(f_i)| < \epsilon, i = 1, \dots, k\}.$$

It is known [35] that each $M_a(X)$ is weak*-compact, i.e., $M_a(X)$ is a compact convex set in the topology determined by the neighborhoods $N_{\mu_0}(f_1, \dots, f_k, \epsilon)$.

Let $V: X \rightarrow [0, \infty)$ be bounded and \mathfrak{B} -measurable, and define

$$T_V(\mu) := (V d\mu) \circ r^{-1} \quad (\mu \in M(X)).$$

Lemma 5.1. *Let V be as above, and assume V is also continuous. Then $T_V: M(X) \rightarrow M(X)$ is continuous.*

Proof. Let $\mu \in M(X)$, $k \in \mathbb{Z}_+$ and $f_i \in C(X)$, $i = 1, \dots, k$. Then

$$\langle T_V(\mu) \mid f_i \rangle = \langle \mu \mid V f_i \circ r \rangle.$$

Setting $g_k := V f_i \circ r$, we see that

$$T_V(N_{\mu_0}(g_1, \dots, g_k, \epsilon)) \subset N_{T_V(\mu_0)}(f_1, \dots, f_k, \epsilon)$$

and the conclusion follows if V is assumed continuous. \square

Definition 5.2. Let $V: X \rightarrow [0, \infty)$ be bounded and \mathfrak{B} -measurable. We use the notation

$$M^V(X) := \{\mu \in M(X) \mid d\mu = (V d\mu) \circ r^{-1}\}.$$

For measures $\hat{\mu}$ on $(X_\infty(r), \mathfrak{B}_\infty)$ we introduce

$$M_{qi}^V(X_\infty(r)) := \left\{ \hat{\mu} \in M(X_\infty(r)) \mid \hat{\mu} \circ \hat{r} \ll \hat{\mu} \text{ and } \frac{d(\hat{\mu} \circ \hat{r})}{d\hat{\mu}} = V \circ \theta_0 \right\}.$$

The results of the previous section may be summarized as follows:

Theorem 5.3. *Let V be as in Definition 5.2. For measures $\hat{\mu}$ on $X_\infty(r)$ and $n \in \mathbb{N}_0$, define*

$$C_n(\hat{\mu}) := \hat{\mu} \circ \theta_n^{-1}.$$

Then C_0 is a bijective affine isomorphism of $M_{qi}^V(X_\infty(r))$ onto $M^V(X)$ that preserves the total measure, i.e., $C_0(\hat{\mu})(X) = \hat{\mu}(X_\infty(r))$ for all $\hat{\mu} \in M_{qi}^V(X_\infty(r))$.

Proof. We showed in Theorem 3.3 that C_0 maps $M_{qi}^V(X_\infty(r))$ onto $M^V(X)$. The inverse mapping

$$C_0^{-1}: M^V(X) \rightarrow M_{qi}^V(X_\infty(r))$$

may be realized using Theorem 4.1 and Lemma 4.2. \square

Remark 5.4. The intuitive idea behind $C_0^{-1}: M^V(X) \rightarrow M_{qi}^V(X_\infty(r))$ is as follows:

Let $C_0^{-1}(\mu_0) = \hat{\mu}$. Then $\hat{\mu}$ is a measure on $X_\infty(r)$, and we view points \hat{x} in $X_\infty(r)$ as infinite paths. Recall, if $\hat{x} = (x_0, x_1, \dots) \in X_\infty(r)$, then $r(x_{n+1}) = x_n$. So in a random walk we choose $x_{n+1} \in r^{-1}(x_n)$, and the function V assigns the probabilities in each step. At step n , there are $\#r^{-1}(x_n)$ choices.

The assertion in the theorem is that the measure $\hat{\mu}$ is completely determined by the prescribed measure μ_0 at the starting point x_0 of the path, and by the function V .

The two measures $\hat{\mu}$ and μ_0 are normalized so that

$$\mu_0(X) = \hat{\mu}(X_\infty(r)).$$

In a special case, an explicit formula for $C_0^{-1}(\mu_0)$ is known [17, Proposition 8.2].

Theorem 5.5. *Let $V: X \rightarrow [0, \infty)$ be continuous. Assume that there exist some measure ν on (X, \mathfrak{B}) and two numbers $0 < a < b$ such that*

$$(5.1) \quad a \leq \nu(X) \leq b, \text{ and } a \leq \int_X V^{(n)} d\nu \leq b \text{ for all } n \in \mathbb{N}.$$

Then there exists a measure μ_0 on (X, \mathfrak{B}) that satisfies

$$d\mu_0 = (V d\mu_0) \circ r^{-1},$$

and there exists a V -quasi-invariant measure $\hat{\mu}$ on $(X_\infty(r), \mathfrak{B}_\infty)$.

Proof. Condition (5.1) guarantees that the set

$$M_{ab}^V(X) := \left\{ \mu \in M(X) \mid a \leq \mu(X) \leq b \text{ and } a \leq \int_X V^{(n)} d\mu \leq b, \text{ for all } n \in \mathbb{N} \right\}$$

is non-empty. Moreover, by the Banach-Alaoglu theorem [35], this set is compact in the weak*-topology. It is clear also that the set is convex.

We claim that the operator T_V maps $M_{ab}^V(X)$ into itself. Indeed, if $\mu \in M_{ab}^V(X)$, then

$$T_V(\mu)(X) = \int_X V \mathbf{1} \circ r d\mu \in [a, b]$$

and, using $V V^{(n)} \circ r = V^{(n+1)}$,

$$T_V(\mu)(V^{(n)}) = \int_X V V^{(n)} \circ r d\mu = \int_X V^{(n+1)} d\mu \in [a, b].$$

Lemma 5.1 shows that T_V is continuous so we can apply the Markov-Kakutani fixed-point theorem (see [36]) to conclude that there exists a $\mu_0 \in M_{ab}^V$ such that $T_V(\mu_0) = \mu_0$. The V -quasi-invariant measure $\hat{\mu}$ can be obtained from μ_0 using Theorem 4.1. \square

Definition 5.6. We now connect the above discussion with Section 2.2. Let (X, \mathfrak{B}) and $r: X \rightarrow X$ be as described above. Suppose in addition that (X, \mathfrak{B}) carries a strongly r -invariant probability measure, ρ ; see (2.4) for the definition.

Set

$$M(\rho, X) := \{\mu \in M(X) \mid \mu \ll \rho\}.$$

Lemma 5.7. *Let $V: X \rightarrow [0, \infty)$ be a bounded measurable function. Then T_V leaves $M(\rho, X)$ invariant; and μ_0 in $M(\rho, X)$ satisfies*

$$T_V(\mu_0) = \mu_0$$

if and only if there is an $h \in L^1(X, \rho)$ such that $d\mu_0 = h d\rho$ and $R_V h = h$.

Proof. Let $\mu \in M(\rho, X)$, and write $d\mu = f d\rho$, $f \in L^1(X, \rho)$. Let ξ be a bounded measurable function on X . Then

$$\begin{aligned} T_V(\mu)(\xi) &= \int_X \xi \circ r V f d\rho = \int_X \xi(x) \frac{1}{\#r^{-1}(x)} \sum_{r(y)=x} V(y) f(y) d\rho(x) \\ &= \int_X \xi(x) (R_V f)(x) d\rho(x). \end{aligned}$$

Stated differently,

$$T_V(f d\rho) = R_V(f) d\rho.$$

Hence $T_V(f d\rho) = f d\rho$ if and only if $R_V f = f$ as claimed. \square

Before stating and proving our next result we need one more lemma.

Lemma 5.8. *Let (X, \mathfrak{B}) , $r: X \rightarrow X$, and $V: X \rightarrow [0, \infty)$ be as above. Suppose (X, \mathfrak{B}) carries a strongly r -invariant probability measure ρ . Then*

$$(5.2) \quad \int_X V^{(n)} f d\rho = \int_X R_V^n(f) d\rho$$

for all bounded measurable functions f and $n \in \mathbb{N}$.

Proof. We prove (5.2) by induction, starting with $n = 1$.

For $f \in L^\infty(X)$,

$$\int_X V f d\rho = \int_X \frac{1}{\#r^{-1}(x)} \sum_{r(y)=x} V(y) f(y) d\rho(x) = \int_X R_V f d\rho$$

where we used the definition (2.4) of strong r -invariance.

Suppose (5.2) holds up to n . Then

$$\begin{aligned} \int_X V^{(n+1)} f d\rho &= \int_X (V^{(n)} \circ r) V f d\rho = \int_X V^{(n)}(x) \frac{1}{\#r^{-1}(x)} \sum_{r(y)=x} V(y) f(y) d\rho(x) \\ &= \int_X V^{(n)} R_V f d\rho = \int_X R_V^n R_V f d\rho = \int_X R_V^{n+1} f d\rho, \end{aligned}$$

where we used the induction hypothesis in the last step.

This completes the proof of the lemma. \square

Our next result is an infinite-dimensional analogue of the (finite-dimensional) Perron-Frobenius theorem for non-negative matrices.

Theorem 5.9. *Let (X, \mathfrak{B}) , and $r: X \rightarrow X$, be as described above. Suppose $V: X \rightarrow [0, \infty)$ is measurable,*

$$\frac{1}{\#r^{-1}(x)} \sum_{r(y)=x} V(y) \leq 1,$$

and that some probability measure ν_V on X satisfies

$$(5.3) \quad \nu_V \circ R_V = \nu_V.$$

Assume also that (X, \mathfrak{B}) carries a strongly r -invariant probability measure ρ , such that

$$\rho(\{x \in X \mid V(x) > 0\}) > 0.$$

Then

- (i) $T_V^n(d\rho) = R_V^n(\mathbf{1}) d\rho$, for $n \in \mathbb{N}$, where $\mathbf{1}$ denotes the constant function one.
 - (ii) [Monotonicity] $\dots \leq R_V^{n+1}(\mathbf{1}) \leq R_V^n(\mathbf{1}) \leq \dots \leq \mathbf{1}$, pointwise on X .
 - (iii) The limit $\lim_{n \rightarrow \infty} R_V^n(\mathbf{1}) = h_V$ exists, $R_V h_V = h_V$, and
- (5.4) $\rho(\{x \in X \mid h_V(x) > 0\}) > 0.$

- (iv) The measure $d\mu_0^{(V)} = h_V d\rho$ is a solution to the fixed-point problem

$$T_V(\mu_0^{(V)}) = \mu_0^{(V)}.$$

- (v) The sequence $d\mu_n^{(V)} = V^{(n)} h_V d\rho$ defines a unique $\hat{\mu}^{(V)}$ (as in Theorem 4.1 and Lemma 4.2); and
- (vi) $\mu_n^{(V)}(f) = \int_X R_V^n(fh_V) d\rho$ for all bounded measurable functions f on X , and all $n \in \mathbb{N}$.

Finally,

- (vii) the measure $\hat{\mu}^{(V)}$ on $X_\infty(r)$ satisfying $\hat{\mu}^{(V)} \circ \theta_n^{-1} = \mu_n^{(V)}$ has total mass

$$\hat{\mu}^{(V)}(X_\infty(r)) = \rho(h_V) = \int_X h_V(x) d\rho(x).$$

Proof. Part (i) follows from Lemma 5.7.

- (ii) It is clear that

$$R_V(\mathbf{1}) = \frac{1}{\#r^{-1}(x)} \sum_{r(y)=x} V(y) \leq 1, \text{ for } x \in X.$$

Starting the induction, suppose

$$R_V^n(\mathbf{1}) \leq R_V^{n-1}(\mathbf{1}).$$

Then

$$\begin{aligned} R_V^{n+1}(\mathbf{1}) &= \frac{1}{\#r^{-1}(x)} \sum_{r(y)=x} V(y) R_V^n(\mathbf{1})(y) \\ &\leq \frac{1}{\#r^{-1}(x)} \sum_{r(y)=x} V(y) R_V^{n-1}(\mathbf{1})(y) = R_V^n(\mathbf{1})(x). \end{aligned}$$

An induction now proves (ii).

It follows that the limit h_V in (iii) exists, and that

$$R_V(h_V) = h_V.$$

Using (5.3), we get that

$$\nu_V(R_V^n(\mathbf{1})) = \nu_V(\mathbf{1}) = 1, \text{ for all } n \in \mathbb{N},$$

and therefore

$$\nu_V(h_V) = \int_X h_V d\nu_V = 1.$$

The conclusion (5.4) follows from [22, Chapter 3].

The assertion in (iv) is that $\mu_0^{(V)} := h_V d\rho$ is a fixed point, i.e., that

$$T_V(h_V d\rho) = h_V d\rho.$$

This follows in turn from Lemma 5.7. Hence, the measures

$$d\mu_n^{(V)} = V^{(n)} h_V d\rho$$

extend to a unique measure $\hat{\mu}^{(V)}$ on $X_\infty(r)$ and, by Theorem 5.3,

$$\hat{\mu}^{(V)} \in M_{qi}^V(X_\infty(r)).$$

Moreover, using now Lemma 5.8, we get

$$\mu_n^{(V)}(f) = \int_X f V^{(n)} h_V d\rho = \int_X R_V^n(f h_V) d\rho,$$

which yields the desired conclusion (vi) and (vii).

Indeed

$$\hat{\mu}^{(V)}(X_\infty(r)) = \int_{X_\infty(r)} \mathbf{1} \circ \theta_0 d\hat{\mu}^{(V)} = \int_X \mathbf{1} d\mu_0 = \int_X h_V d\rho.$$

Comparing with (5.1) in Theorem 5.5, notice that

$$\int_X h_V d\rho = \lim_{n \rightarrow \infty} \int_X R_V^n(\mathbf{1}) d\rho = \lim_{n \rightarrow \infty} \rho(V^{(n)}).$$

□

The conclusions of Theorem 5.9 hold in the following more general form.

Corollary 5.10. *Let $V, W: X \rightarrow [0, \infty)$ be measurable, and suppose the assumptions (i) and (ii) are satisfied:*

- (i) $\frac{1}{\#r^{-1}(x)} \sum_{r(y)=x} V(y)W(y) \leq 1, x \in X.$
- (ii) *There exists a non-negative function h_W on X , and a measure ρ_W such that $\rho_W R_W = \rho_W$, $R_W h_W = h_W$ and $\rho_W(h_W) = 1.$*

Then

$$(5.5) \quad T_V^n(f d\rho_W) = (R_{VW}^n f) d\rho_W,$$

and the limit

$$\lim_{n \rightarrow \infty} R_{VW}^n(\mathbf{1}) = h$$

exists, and satisfies

$$(5.6) \quad T_V(h d\rho_W) = h d\rho_W.$$

Proof. We first show (5.5) for $n = 1$. Let f and ξ be bounded, measurable functions on X . Then

$$\begin{aligned} T_V(f d\rho_W)(\xi) &= \int_X \xi \circ r V f d\rho_W = \int_X R_W(\xi \circ r V f) d\rho_W \\ &= \int_X \xi R_W(V f) d\rho_W = \int_X \xi R_{VW}(f) d\rho_W. \end{aligned}$$

Hence (5.5) holds for $n = 1$, and the general case follows by iteration.

The argument from the proof of the theorem shows that

$$\cdots \leq R_{VW}^{n+1}(\mathbf{1}) \leq R_{VW}^n(\mathbf{1}) \leq \cdots \leq R_{VW}(\mathbf{1}) \leq \mathbf{1},$$

so the limit

$$h := \lim_{n \rightarrow \infty} R_{VW}^n(\mathbf{1})$$

exists, and satisfies (5.6) by the same argument. \square

Remark 5.11. Stated in Ruelle's thermodynamical formalism [37], the data ρ_W (measure) and h_W (eigenfunction) in part (ii) of Corollary 5.10 represent an equilibrium distribution where W is related to a potential function. Under mild conditions on (X, r) and W , it is known that solutions (ρ_W, h_W) exist, and we say that $\lambda_W = 1$ is the Perron-Frobenius eigenvalue of the Ruelle operator R_W . In that case

$$1 = \sup\{|\lambda| \mid \lambda \in \text{spectrum}(R_W)\}.$$

The reader is referred to [37], [8], [32] for further details regarding the spectral theory of R_W .

Notice further that the conclusion of Ruelle's Perron-Frobenius theorem is a generalization of the classical Perron-Frobenius theorem for matrices with non-negative entries.

6. EXTREME POINTS

Theorem 5.3 shows that the map $\hat{\mu} \mapsto \mu_0 := \hat{\mu} \circ \theta_0^{-1}$ establishes a bijective affine correspondence between the following two sets:

$$M_{qi,1}^V(X_\infty(r)) := \{\hat{\mu} \in M(X_\infty(r)) \mid \hat{\mu}(X_\infty(r)) = 1, \hat{\mu} \text{ is } V\text{-quasi-invariant}\}$$

and

$$M_1^V(X) := \{\mu_0 \in M(X) \mid \mu_0(X) = 1, T_V(\mu_0) = \mu_0\}.$$

It is easy to see that both these sets are convex, and an application of the Banach-Alaoglu theorem [36] shows that they are compact in the weak*-topology. Then, using the Krein-Milman theorem [36], we conclude that each of these sets is the convex weak*-closure of their extreme points. Moreover, since the correspondence is affine it preserves the extreme points.

This section is devoted to an analysis of the extreme points. Before we state and prove our main result on extreme points, we need to define the concepts of conditional expectation E_{μ_0} , and relative ergodicity.

Proposition 6.1. *Let $(X, \mathfrak{B}), r$ and V be as above. Let μ_0 be a measure in $M_1^V(X)$. Then for each bounded measurable function g on X there exists a bounded $r^{-1}(\mathfrak{B})$ -measurable function $E_{\mu_0}(Vg)$ such that*

$$(6.1) \quad \int_X V g f \circ r \, d\mu_0 = \int_X E_{\mu_0}(Vg) f \circ r \, d\mu_0,$$

for all bounded \mathfrak{B} -measurable functions f on X . Moreover, this is unique up to $\mu_0 \circ r^{-1}$ -measure zero.

Proof. The positive linear functional

$$\Lambda_g : f \circ r \mapsto \int_X V g f \circ r \, d\mu_0$$

defines a measure on $(X, r^{-1}(\mathfrak{B}))$ which is absolutely continuous with respect to μ_0 . Indeed, if $\mu(E) = 0$, then

$$0 = \int_X \chi_E \, d\mu_0 = \int_X V \chi_E \circ r \, d\mu_0$$

so

$$\int_X V g \chi_E \circ r \, d\mu_0.$$

Therefore, by the Radon-Nikodym theorem, there exists some $r^{-1}(\mathfrak{B})$ -measurable function $E_{\mu_0}(Vg)$ such that (6.1) holds. Since

$$\left| \int_X E_{\mu_0}(Vg) f \circ r \, d\mu_0 \right| \leq \|Vg\|_\infty \int_X |f \circ r| \, d\mu_0$$

it follows that $|E_{\mu_0}(Vg)| \leq \|Vg\|_\infty$ holds $\mu_0 \circ r^{-1}$ a.e.

The uniqueness is also clear from the definition (6.1). \square

Definition 6.2. A measure $\mu_0 \in M_1^V(X)$ is called *relatively ergodic* with respect to (r, V) if the only non-negative, bounded \mathfrak{B} -measurable functions f on X satisfying

$$E_{\mu_0}(Vf) = E_{\mu_0}(V)f \circ r, \quad \text{pointwise } \mu_0 \circ r^{-1}\text{-a.e.,}$$

are the functions which are constant μ_0 -a.e.

Theorem 6.3. Let $V: X \rightarrow [0, \infty)$ be bounded and measurable. Let $\hat{\mu} \in M_{qi,1}^V(X_\infty(r))$, and $\mu_0 := \hat{\mu} \circ \theta_0^{-1} \in M_1^V(X)$. The following affirmations are equivalent:

- (i) $\hat{\mu}$ is an extreme point of $M_{qi,1}^V(X_\infty(r))$;
- (ii) $V \circ \theta_0 d\hat{\mu}$ is ergodic with respect to \hat{r} ;
- (iii) μ_0 is an extreme point of $M_1^V(X)$;
- (iv) μ_0 is relatively ergodic with respect to (r, V) .

Proof. The arguments in the beginning of this section (mainly Theorem 5.3) show that (i) and (iii) are equivalent.

We now prove (i) \Rightarrow (ii). Suppose $\hat{\mu}$ is not ergodic. Then there exists two measurable subsets A and B of $X_\infty(r)$ such that

- $\int_A V \circ \theta_0 \, d\hat{\mu} > 0$ and $\int_B V \circ \theta_0 \, d\hat{\mu} > 0$
(note that then $\hat{\mu}(A) > 0$ and $\hat{\mu}(B) > 0$),
- $A \cup B = X_\infty(r)$, and
- A and B are \hat{r} -invariant.

Define then the measures

$$\hat{\mu}_A(E) = \frac{\hat{\mu}(A \cap E)}{\hat{\mu}(A)}, \quad \hat{\mu}_B(E) = \frac{\hat{\mu}(B \cap E)}{\hat{\mu}(B)}$$

for all $E \in \mathfrak{B}$.

Note that

$$(6.2) \quad \hat{\mu} = \hat{\mu}(A)\mu_A + \hat{\mu}(B)\mu_B.$$

We prove next that μ_A, μ_B are in $M_{qi,1}^V(X_\infty(r))$. Clearly they have total mass equal to 1, so we only have to prove the V -quasi-invariance. For f bounded, measurable

function on $X_\infty(r)$, we have:

$$\begin{aligned}
\int_{X_\infty(r)} f \circ \hat{r}^{-1} d\mu_A &= \frac{1}{\hat{\mu}(A)} \int_{X_\infty(r)} f \circ \hat{r}^{-1} \chi_A d\hat{\mu} \\
&= \frac{1}{\hat{\mu}(A)} \int_{X_\infty(r)} f \circ \hat{r}^{-1} \chi_A \circ \hat{r}^{-1} d\hat{\mu} \quad (\hat{r}(A) = A) \\
&= \frac{1}{\hat{\mu}(A)} \int_{X_\infty(r)} V \circ \theta_0 f \chi_A d\hat{\mu} \\
&= \int_{X_\infty(r)} V \circ \theta_0 f d\mu_A.
\end{aligned}$$

Hence μ_A is V -quasi-invariant. The same argument works for μ_B . Therefore μ_A and μ_B are both in $M_{qi,1}^V(X_\infty(r))$. Now, equation (6.2) contradicts the fact that $\hat{\mu}$ is extreme.

(ii) \Rightarrow (i): Suppose there are some measures $\mu_1, \mu_2 \in M_{qi,1}^V(X_\infty(r))$ and some $\lambda \in [0, 1]$ such that

$$\hat{\mu} = \lambda\mu_1 + (1 - \lambda)\mu_2.$$

Then μ_1 and μ_2 are absolutely continuous with respect to $\hat{\mu}$. Therefore there exist f_1 and f_2 in $L^1(X_\infty(r), \hat{\mu})$ such that

$$d\mu_1 = f_1 d\hat{\mu}, \quad d\mu_2 = f_2 d\hat{\mu}.$$

Since μ_1 and $\hat{\mu}$ are V -quasi-invariant, we have, for all bounded measurable functions f on $X_\infty(r)$,

$$\begin{aligned}
\int_{X_\infty(r)} f f_1 \circ \hat{r} V \circ \theta_0 d\hat{\mu} &= \int_{X_\infty(r)} f \circ \hat{r}^{-1} f_1 d\hat{\mu} \\
&= \int_{X_\infty(r)} f \circ \hat{r}^{-1} d\mu_1 \\
&= \int_{X_\infty(r)} V \circ \theta_0 f d\mu_1 \\
&= \int_{X_\infty(r)} V \circ \theta_0 f f_1 d\hat{\mu}.
\end{aligned}$$

Therefore $f_1 = f_1 \circ r$ pointwise $V \circ \theta_0 d\hat{\mu}$ -almost everywhere. The hypothesis implies then that f_1 is constant $V \circ \theta_0 d\hat{\mu}$ -a.e. The same argument shows that f_2 is constant $V \circ \theta_0 d\hat{\mu}$ -a.e.

Since $\hat{\mu}$ and μ_1 are V -quasi-invariant we have also that

$$1 = \int_{X_\infty(r)} \mathbf{1} \circ \hat{r}^{-1} d\mu_1 = \int_{X_\infty(r)} V d\mu_1 = \int_{X_\infty(r)} V f_1 d\hat{\mu}.$$

It follows that $f_1 = 1$, pointwise $V \circ \theta_0 d\hat{\mu}$ -a.e. The same is true for f_2 . Then

$$\begin{aligned}
\int_{X_\infty(r)} f d\mu_1 &= \int_{X_\infty(r)} V \circ \theta_0 f \circ \hat{r} d\mu_1 = \int_{X_\infty(r)} V \circ \theta_0 f \circ \hat{r} f_1 d\hat{\mu} \\
&= \int_{X_\infty(r)} V f \circ \hat{r} d\hat{\mu} = \int_{X_\infty(r)} f d\hat{\mu}.
\end{aligned}$$

Hence $\mu_1 = \hat{\mu} = \mu_2$ and $\hat{\mu}$ is extreme.

(iii) \Rightarrow (iv): Suppose μ_0 is not relatively ergodic. Then there exists a bounded measurable function $f_1 \geq 0$ on X such that $E_{\mu_0}(V f_1) = E_{\mu_0}(V) f_1 \circ r$, pointwise

$\mu_0 \circ r^{-1}$ -a.e., and f_1 is not constant μ_0 -a.e. We may assume that $\int_X f_1 d\mu_0 = 1$. Define the measure $d\mu_1 := f_1 d\mu_0$. We check that μ_1 is in $M_1^V(X)$.

$$\begin{aligned} \int_X f d\mu_1 &= \int_X f f_1 d\mu_0 = \int_X V f \circ r f_1 \circ r d\mu_0 = \int_X E_{\mu_0}(V) f_1 \circ r f \circ r d\mu_0 \\ &= \int_X E_{\mu_0}(V f_1) f \circ r d\mu_0 = \int_X V f_1 f \circ r d\mu_0 = \int_X V f \circ r d\mu_1. \end{aligned}$$

Now choose some $0 < \lambda < 1$ such that $\lambda f_1 \leq 1$ and define

$$f_2 := \frac{1 - \lambda f_1}{1 - \lambda}, \quad d\mu_2 = f_2 d\mu_0.$$

Then $E_{\mu_0}(V f_2) = E_{\mu_0}(V) f_2 \circ r$, and the same calculation as before shows that $T_V(\mu_2) = \mu_2$. Note also that $\mu_2(X) = 1$. Since $\mu_0 = \lambda \mu_1 + (1 - \lambda) \mu_2$ and f_1 is not constant, we have that $\mu_1 \neq \mu_0$. It follows that μ_0 is not an extreme point, thus contradicting the hypothesis and proving (iv).

(iv) \Rightarrow (iii): Suppose

$$(6.3) \quad \mu_0 = \lambda \mu_1 + (1 - \lambda) \mu_2$$

for some $\mu_1, \mu_2 \in M_1^V(X)$, $\lambda \in [0, 1]$. Then μ_1 and μ_2 are absolutely continuous with respect to μ_0 . Let f_1, f_2 be the corresponding Radon-Nikodym derivatives. The relation (6.3) implies that

$$1 = \lambda f_1 + (1 - \lambda) f_2, \quad \mu_0\text{-a.e.}$$

In particular, f_1 and f_2 are bounded.

We know that μ and μ_1 have the fixed-point property. Then for all bounded measurable functions f on X ,

$$\begin{aligned} \int_X E_{\mu_0}(V) f_1 \circ r f \circ r d\mu_0 &= \int_X V f \circ r f_1 \circ r d\mu_0 = \int_X f f_1 d\mu_0 = \int_X f d\mu_1 \\ &= \int_X V f \circ r d\mu_1 = \int_X V f \circ r f_1 d\mu_0 = \int_X E_{\mu_0}(V f_1) f \circ r d\mu_0. \end{aligned}$$

Therefore $E_{\mu_0}(V) f_1 \circ r = E_{\mu_0}(V f_1)$, $\mu_0 \circ r^{-1}$ -a.e.

The hypothesis implies that f_1 is constant μ_0 -a.e. Since $\mu_1(X) = \mu_0(X) = 1$, it follows that $f_1 = 1$ and $\mu_1 = \mu_0$, and therefore μ_0 is extreme. \square

ACKNOWLEDGMENTS

The authors are pleased to thank Professors Richard Gundy, and Roger Nussbaum for helpful discussions and suggestions, the referee for useful suggestions, and Brian Treadway for help with revisions in a first version. This work was supported at the University of Iowa by a grant from the National Science Foundation (NSF-USA) under a Focused Research Program, DMS-0139473 (FRG).

REFERENCES

- [1] L. Accardi, A. Frigerio, J.T. Lewis, Quantum stochastic processes, *Publ. Res. Inst. Math. Sci.* **18** (1982), 97–133. MR0660823 (84m:82031)
- [2] A. Aldroubi, D. Larson, W.-S. Tang, E. Weber, Geometric aspects of frame representations of abelian groups, *Trans. Amer. Math. Soc.* **356** (2004), 4767–4786. MR2084397
- [3] J. Ashley, B. Markus, S. Tuncel, The classification of one-sided Markov chains, *Ergodic Theory Dynam. Systems* **17** (1997), 269–295. MR1444053 (98k:28021)
- [4] A. Aldroubi, Q. Sun, W.-S. Tang, Nonuniform average sampling and reconstruction in multiply generated shift-invariant spaces, *Constr. Approx.* **20** (2004), 173–189. MR2036639

- [5] L.W. Baggett, A. Carey, W. Moran, P. Ohring, General existence theorems for orthonormal wavelets, an abstract approach, *Publ. Res. Inst. Math. Sci.* **31** (1995), 95–111. MR1317525 (96c:42060)
- [6] L.W. Baggett, P.E.T. Jorgensen, K.D. Merrill, J.A. Packer, An analogue of Bratteli-Jorgensen loop group actions for GMRA's, in *Wavelets, Frames and Operator Theory* (College Park, 2003), eds. C. Heil et al., Contemp. Math., vol. 345, American Mathematical Society, 2004, pp. 11–25. MR2066819
- [7] L.W. Baggett, K.D. Merrill, Abstract harmonic analysis and wavelets in \mathbb{R}^n , in *The Functional and Harmonic Analysis of Wavelets and Frames* (San Antonio, 1999), eds. L.W. Baggett, D.R. Larson, Contemp. Math., vol. 247, American Mathematical Society, 1999, pp. 17–27. MR1735967 (2001b:42043)
- [8] V. Baladi, *Positive Transfer Operators and Decay of Correlations*, Advanced Series in Non-linear Dynamics, vol. 16, World Scientific, River Edge, NJ, Singapore, 2000. MR1793194 (2001k:37035)
- [9] L. Barreira, Ya. B. Pesin, *Lyapunov exponents and smooth ergodic theory*, University Lecture Series, vol. 23, American Mathematical Society, Providence, RI, 2002. MR1862379 (2003a:37040)
- [10] B. Brenken, P.E.T. Jorgensen, A family of dilation crossed product algebras, *J. Operator Theory* **25** (1991), 299–308. MR1203035 (94m:46103)
- [11] H. Brolin, Invariant sets under iteration of rational functions, *Ark. Mat.* **6** (1965), 103–144. MR0194595 (33:2805)
- [12] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conf. Ser. in Appl. Math., vol. 61, Society for Industrial and Applied Mathematics, Philadelphia, 1992. MR1162107 (93e:42045)
- [13] V. Deaconu, P.S. Muhly, C^* -algebras associated with branched coverings, *Proc. Amer. Math. Soc.* **129** (2001), 1077–1086. MR1814145 (2002a:46075)
- [14] J.L. Doob, *Classical Potential Theory and Its Probabilistic Counterpart*, Grundlehren der Mathematischen Wissenschaften, vol. 262, Springer-Verlag, 1984. MR0731258 (85k:31001)
- [15] D.E. Dutkay, The wavelet Galerkin operator, *J. Operator Theory* **51** (2004), 49–70. MR2055804 (2005c:47041)
- [16] D.E. Dutkay, P.E.T. Jorgensen, Wavelets on fractals, to appear in *Rev. Mat. Iberoamericana*.
- [17] D.E. Dutkay, P.E.T. Jorgensen, *Martingales, endomorphisms, and covariant systems of operators in Hilbert space*, preprint 2004, arxiv math.CA/0407330.
- [18] R. Gohm, B. Kummerer, T. Lang, *Non-commutative symbolic coding*, Preprint 2005, Greifswald University.
- [19] R.F. Gundy, Two remarks concerning wavelets: Cohen's criterion for low-pass filters and Meyer's theorem on linear independence, in *The Functional and Harmonic Analysis of Wavelets and Frames* (San Antonio, 1999), eds. L.W. Baggett, D.R. Larson, Contemp. Math., vol. 247, American Mathematical Society, 1999, pp. 249–258. MR1738093 (2001b:42051)
- [20] J.E. Hutchinson, Fractals and self-similarity, *Indiana Univ. Math. J.* **30** (1981), no. 5, 713–747. MR0625600 (82h:49026)
- [21] P.E.T. Jorgensen, Ruelle operators: Functions which are harmonic with respect to a transfer operator, *Mem. Amer. Math. Soc.* **152** (2001), no. 720. MR1837681 (2002c:46117)
- [22] P.E.T. Jorgensen, *Analysis and Probability*, preprint 2004.
- [23] P.E.T. Jorgensen, D.W. Kribs, Wavelet representations and Fock space on positive matrices, *J. Funct. Anal.* **197** (2003), 526–559. MR1960424 (2004e:42055)
- [24] P.E.T. Jorgensen, P.S. Muhly, Selfadjoint extensions satisfying the Weyl operator commutation relations, *J. Analyse Math.* **37** (1980), 46–99. MR0583632 (82k:47058)
- [25] E. Katsoulis, D.W. Kribs, Isomorphisms of algebras associated with directed graphs, *Math. Ann.* **330** (2004), 709–728. MR2102309
- [26] D.W. Kribs, On bilateral weighted shifts in noncommutative multivariable operator theory, *Indiana Univ. Math. J.* **52** (2003), 1595–1614. MR2021049 (2004k:47004)
- [27] A.N. Kolmogoroff, *Grundbegriffe der Wahrscheinlichkeitsrechnung* (German), Springer-Verlag, 1977 (reprint of the 1933 original); *Foundations of the Theory of Probability* (English), Chelsea, 1950. MR0494348 (58:13242)
- [28] B. Kummerer, Markov dilations on W^* -algebras, *J. Funct. Anal.* **63** (1985), 139–177. MR0803091 (87b:46070)

- [29] S. G. Mallat, Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbb{R})$, *Trans. Amer. Math. Soc.* **315** (1989), 69–87. MR1008470 (90e:42046)
- [30] P.S. Muhly, B. Solel, Quantum Markov processes (correspondences and dilations), *Internat. J. Math.* **13** (2002), no. 8, 863–906. MR1928802 (2003j:46100)
- [31] J. Neveu, *Discrete-Parameter Martingales*, North-Holland, 1975. MR0402915 (53:6729)
- [32] R.D. Nussbaum, S.M. Verduyn Lunel, Generalizations of the Perron-Frobenius theorem for nonlinear maps, *Mem. Amer. Math. Soc.* **138** (1999), no. 659. MR1470912 (99i:58125)
- [33] G. Popescu, Commutant lifting, tensor algebras, and functional calculus, *Proc. Edinb. Math. Soc. (2)* **44** (2001), 389–406. MR1880399 (2003f:47125)
- [34] G. Popescu, Central intertwining lifting, suboptimization, and interpolation in several variables, *J. Funct. Anal.* **189** (2002), 132–154. MR1887631 (2002m:47013)
- [35] W. Rudin, *Real and Complex Analysis*, third edition, McGraw-Hill, 1987. MR0924157 (88k:00002)
- [36] W. Rudin, *Functional Analysis*, second edition, International Series in Pure and Applied Mathematics, McGraw-Hill, 1991. MR1157815 (92k:46001)
- [37] D. Ruelle, The thermodynamic formalism for expanding maps, *Comm. Math. Phys.* **125** (1989), 239–262. MR1016871 (91a:58149)
- [38] D. Ruelle, Application of hyperbolic dynamics to physics: some problems and conjectures, *Bull. Amer. Math. Soc. (N.S.)* **41** (2004), 275–278. MR2058287 (2005b:37038)
- [39] M. Su, The information topology and true laminations for diffeomorphisms, *Conform. Geom. Dyn.* **8** (2004), 36–51. MR2060377
- [40] D. Szasz, T. Varju, Local limit theorem for the Lorentz process and its recurrence in the plane, *Ergodic Theory Dynam. Systems* **24** (2004), 257–278. MR2041271

DEPARTMENT OF MATHEMATICS, HILL CENTER-BUSCH CAMPUS, RUTGERS UNIVERSITY, PISCATAWAY, NEW JERSEY 08845-8019

E-mail address: `ddutkay@math.rutgers.edu`

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF IOWA, 14 MACLEAN HALL, IOWA CITY, IOWA 52242-1419

E-mail address: `jorgen@math.uiowa.edu`