CONTEMPORARY MATHEMATICS

567

Dynamical Systems and Group Actions

Lewis Bowen
Rostislav Grigorchuk
Yaroslav Vorobets
Editors

American Mathematical Society

Dynamical Systems and Group Actions

CONTEMPORARY MATHEMATICS

567

Dynamical Systems and Group Actions

Lewis Bowen
Rostislav Grigorchuk
Yaroslav Vorobets
Editors

American Mathematical Society
Providence, Rhode Island

2010 *Mathematics Subject Classification.* Primary 37Axx, 37Bxx, 37Dxx, 20Exx, 20Pxx, 20Nxx, 22Fxx.

Library of Congress Cataloging-in-Publication Data

Dynamical systems and group actions / Lewis Bowen, Rostislav Grigorchuk, Yaroslav Vorobets, editors.

p. cm. — (Contemporary mathematics ; v. 567)

Includes bibliographical references.

ISBN 978-0-8218-6922-2 (alk. paper)

1. Ergodic theory. 2. Topological dynamics. I. Bowen, Lewis, 1972- II. Grigorchuk, R. I. III. Vorobets, Yaroslav, 1971-

QA611.5.D96 2012

515′.39—dc23 2011051433

---

10 9 8 7 6 5 4 3 2 1   17 16 15 14 13 12

Dedicated with the greatest respect and admiration to Anatoli Stepin
on the occasion of his 70th birthday.

Contents

Preface

On July 20, 2010 Anatoli Mikhailovich Stepin, Professor of Mathematics at Moscow (Lomonosov) State University, celebrated his seventieth birthday. Anatoli Stepin studied mathematics at Moscow State University, then worked there for over 40 years. His supervisors were Feliks Berezin and Yakov Sinai and that shaped the range of research interests of Stepin: dynamical systems and ergodic theory as well as related problems of functional analysis and probability theory. A research direction of special interest has been the study of group actions on measured spaces, orbit equivalence, and σ-algebras.

It would take too long to describe all of the areas and problems to which Stepin contributed. We are going to mention very briefly just a few (for a more detailed account of his work, see [**1**]). Anatoli Stepin began his scientific career studying approximations of dynamical systems by periodic transformations. The pioneering results in this direction obtained jointly with Anatole Katok formed the basis of his Ph.D. thesis defended in 1968. Those results are now classical, they are widely and commonly used in ergodic theory and its applications. In 1967 Stepin received the Prize of the Moscow Mathematical Society, one of the major awards in Soviet mathematics.

Another one of the early results of Stepin was solution of a problem, due to Kolmogorov, on the group property of spectra of dynamical systems. His construction of dense measures with mutually singular convolutions found its applications in harmonic analysis and ergodic theory. The paper [**2**], written jointly with Anatole Katok, among other things laid the foundation of the theory of interval exchange transformations, which is now a rapidly developing direction in one-dimensional dynamics. Stepin proposed a cohomological approach to the study of orbit equivalence in [**3**]. Of great importance are results of Stepin and his students on the theory of billiard systems. He established and studied mechanisms of creating periodic trajectories in billiard systems without focusing and scattering. Stepin clarified the fundamental role of the Lyapunov exponents in the spectral theory of the weighted shift operators. He has results on the von Neumann algebras and C^*-algebras associated with dynamical systems. Yet another important contribution is the work of Stepin and his students on integrability of dynamical systems. Stepin was one of those who started the study of models of statistical physics, in particular, the Ising model on graphs and groups.

Stepin obtained a complete solution of the problem of calculating unitary invariants of induced flows on homogeneous spaces of semisimple and soluble groups, the problem having been posed as a consequence of the famous paper by Gelfand and Fomin.

The group actions and the use of symmetries were fundamental principles in Stepin's approach to many problems of contemporary mathematics. As early as 1970 he obtained results on the classification of decreasing sequences of homogeneous partitions and invented an entropy-type invariant for such sequences. These methods were further developed by Stepin and Vershik along with the idea of a local isomorphism that led to the introduction of the classes of hyperfinite groups and sofic groups, as well as to the use of various topologies in spaces of groups and subgroups to solve problems in group theory.

The paper [**4**] initiated the study of isomorphisms of Bernoulli shifts on non-commutative groups. The notion of an Ornstein group was introduced there, which is the subject of one of the papers in this volume. Stepin pioneered the use of invariant means on groups in the study of dynamical systems and statistical physics. In the early 70s, he was one of those who came up with an idea of a random ordering on a group and used it to prove variational type theorems for actions of non-commutative groups. The work of Stepin and his students led to a much better understanding of the class of amenable groups and amenable actions.

During his 40 years at the Department of Mathematics and Mechanics of Moscow State University, Anatoli Stepin supervised more than 30 Ph.D. students. Since 1978, he has been organizing (together with Dmitri Anosov and, from 1985 to 2002, Rostislav Grigorchuk) the research seminar on dynamical systems and ergodic theory at Moscow State University. The colleagues and students of Anatoli Stepin enjoyed working with him and praised him for his positive attitude, supportiveness and work ethic. The students can recall numerous field trips and sports tournaments involving seminars of Sinai, Kirillov and Stepin.

This collection of papers conceived and organized by the undersigned (of whom two are Anatoli Stepin's students) reflects to a large extent the research interests of Stepin. Many of the papers further develop or discuss ideas pioneered by Anatoli Stepin and it is a great pleasure for us to dedicate to him this volume of Contemporary Mathematics.

Lewis Bowen
Rostislav Grigorchuk
Yaroslav Vorobets

Texas A&M University
College Station, TX

Bibliography

[1] D. V. Anosov, A. A. Bolibrukh, A. M. Vershik et al., *Anatolii Mikhailovich Stepin (on the occasion of his sixtieth birthday)*. Russ. Math. Surveys **56** (2001), no. 1, 201–205.

[2] A. B. Katok, A. M. Stepin, *Approximations in ergodic theory* (Russian). Uspehi Mat. Nauk **22** (1967), no. 5(137), 81–106.

[3] A. M. Stepin, *The cohomology of automorphism groups of a Lebesgue space*. Functional Anal. Appl. **5** (1971), 167–168.

[4] A. M. Stepin, *Bernoulli shifts on groups*. Soviet Math. Dokl. **16** (1975), no. 4, 886–889.

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11230

Hyperfinite actions on countable sets and probability measure spaces

Miklós Abért and Gábor Elek

Abstract. We introduce the notion of hyperfiniteness for permutation actions of countable groups and give a geometric and analytic characterization, similar to the known characterizations for amenable actions. We also answer a question of van Douwen on actions of the free group on two generators on countable sets.

1. Introduction

Let Γ be a countable group acting on a countably infinite set X by permutations. An invariant mean on X is a Γ-invariant, finitely additive map μ from the set of subsets of X to $[0,1]$ satisfying $\mu(X) = 1$. Von Neumann [**NEU**] initiated the study of invariant means of group actions.

We say that a group action of Γ on X is *amenable* if X admits a Γ-invariant mean. The group Γ is amenable, if the right action of Γ on itself is amenable. Over the decades, amenability of groups has become an important subject, with connections to various areas in mathematics, like combinatorial group theory, probability theory, ergodic theory and harmonic analysis.

All actions of amenable groups are amenable and for free actions, this trivially goes the other way round as well, but in general, one has to assume certain faithfulness conditions to make the notion meaningful. Even when making the natural assumption that the action is transitive, the general picture is that for most sets of conditions, one can construct corresponding amenable actions of groups that are very far from being amenable themselves.

In particular, van Douwen [**DOU**] constructed a transitive amenable action of the free group on two generators such that every nontrivial element of the group fixes only finitely many points. We call this condition *almost freeness*. Further examples of amenable actions of non-amenable groups were given by Glasner and Monod [**GN**] and by Moon [**MO1**], [**M02**].

For probability measure preserving (p.m.p.) actions, the notion that mostly corresponds to amenability is hyperfiniteness. Let Γ act on a probability measure space, preserving the measure. The action is called *hyperfinite* if the measurable equivalence relation generated by the action is up to measure zero an ascending

2010 *Mathematics Subject Classification.* Primary 37A15.
Key words and phrases. Hyperfinite equivalence relations, amenability.
Supported by OTKA Grant 69062 and TAMOP Grant 4.2.1/B-09/1/KMR-2010-003.

union of finite measurable equivalence relations. As before, all p.m.p. actions of amenable groups are hyperfinite, and for free actions, there is equivalence, but in general, very large groups can act hyperfinitely. In particular, Grigorchuk and Nekhrashevych [**GRIG**] constructed an ergodic, faithful, hyperfinite p.m.p. action of a non-amenable group. More generally, for a hyperfinite p.m.p. action of a group Γ, the action of Γ on almost all orbits is amenable. For the other direction, Kaimanovich [**KAI**] presented a counterexample. The main goal of this paper is to introduce and analyze the notion of hyperfiniteness for permutation actions of countable groups. If Γ acts on a countably infinite set X, then the action always extends to the Stone-Cech compactification βX. This connection establishes a bijection between invariant means on X and invariant measures on βX. In particular, an action is amenable if and only if the extended action preserves a regular Borel-probability measure. This suggests the following definition.

Definition 1.1. Let the countable group Γ act on the set X by permutations. We say that the action is *hyperfinite* if βX admits a regular Borel probability measure that is invariant under the extended action and for which this action is hyperfinite.

In particular, every hyperfinite action is automatically amenable. Our first result gives a combinatorial and a geometric characterization of hyperfiniteness for actions of finitely generated groups. Let G_n be a sequence of graphs with an absolute bound on the degrees of vertices in G_n. We say that (G_n) is *hyperfinite*, if for all $\varepsilon > 0$, there exists $Y_n \subseteq V(G_n)$ and $K > 0$ such that $|Y_n| < \varepsilon\,|G_n|$ and every connected component of the subgraph induced on $V(G_n)\backslash Y_n$ has size at most K $(n \geq 1)$. This notion was introduced in [**ECOST**].

Theorem 1. *Let Γ be a group generated by the finite symmetric set S, acting on the countably infinite set X by permutations. Let S_Γ denote the Schreier graph of this action with respect to S. Then the following are equivalent:*
1) The action is hyperfinite;
2) There exists a hyperfinite Følner-sequence in S_Γ;
3) There exists an invariant mean μ on X, such that for all $\varepsilon > 0$, there exists $Y \subseteq X$ with $\mu(Y) < \varepsilon$ and $K > 0$ such that the connected components of the induced subgraph of S_Γ on $X\backslash Y$ have size at most K.

In [**DOU**] van Douwen asked the following question. Let H be a countable infinite amenable group. Is there an almost free transitive action of F_2, the free group of two generators, on H such that every invariant mean on H is F_2-invariant? We will show that the answer is negative, however, it is true if we change the almost freeness condition to faithfulness.

Theorem 2.

(1) *There exists no almost-free transitive action of F_2 on a finitely generated amenable group H which preserves all H-invariant means.*
(2) *For any finitely generated amenable group H, there exists a faithful, transitive action of F_2 on H which preserves all the H-invariant means.*

Finally, we show the following.

Theorem 3. *There exists an ergodic, faithful p.m.p. profinite action of a non-amenable group that is hyperfinite but topologically free.*

Note that this answers a question of Grigorchuk, Nekrashevich and Sushchanskii [**GRNS**]. Note that Bergeron and Gaboriau [**BG**] also constructed an ergodic, faithful p.m.p profinite action which is not free, but topologically free.

Acknowledgement

The authors thank Nicolas Monod for helpful comments on a previous version of this paper.

2. The Stone-Cech compactification

Let X be a countably infinite set and βX be its Stone-Cech compactification. The elements of βX are the ultrafilters on X and the set X itself is identified with the principal ultrafilters. For a subset $A \subseteq X$, let U_A be the set of ultrafilters $\omega \in \beta X$ such that $A \in \omega$. Then $\{U_A\}_{A \subseteq X}$ forms a base of the compact Hausdorff topology of βX. It is well-known that the Banach-algebra of continuous functions $C(\beta X)$ can be identified with the Banach-algebra $l^\infty(X)$. Let μ be a finitely additive measure on X. Then one can associate a regular Borel measure $\widehat{\mu}$ on βX, by taking

$$\widehat{\mu}(U_A) = \mu(A)\,.$$

Indeed, let $f \in l^\infty(X)$ be a bounded real function on X. Then the continuous linear transformation

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

is well-defined. Thus, by the Riesz representation theorem

$$T(f) = \int_{\beta X} f\, d\widehat{\mu}$$

for some regular Borel-measure $\widehat{\mu}$. Since $T(\chi_{U_A}) = \mu(A)$, the equality $\mu(A) = \widehat{\mu}(U_A)$ holds.

In fact, there is a one-to-one correspondance between the regular Borel-measures and the finitely additive measures on X, since the integral $\int_X f\, d\mu$ is completely defined by μ. If $s : X \to X$ is a bijection, then it extends to a map $\widehat{s} : \beta X \to \beta X$ by

$$\widehat{s}(\omega) = \bigcup_{A \in \omega} s(A)\,.$$

Since $\widehat{s}(U_A) = U_{s(A)}$, the map $\widehat{s}$ is a continuous bijection.

Thus if Γ is a countable group acting on X, then we have an extended action on βX. The following lemma is due to Blümlinger [**BLU**, Lemma 1]

LEMMA 2.1. *There is a one-to-one correspondance between the* Γ*-invariant means on* X *and the* Γ*-invariant regular Borel probability measures on* X*.*

Proof. Observe that the set of Γ-invariant regular measures is the annihilator of the set

$$\{f - \gamma(f) \mid f \in C(\beta X), \gamma \in \Gamma\}\,,$$

and the set of Γ-invariant means is the annihilator of the set

$$\{f - \gamma(f) \mid f \in l^\infty(X)\}\,.$$

Therefore an action of Γ is amenable if and only if the corresponding action on βX has an invariant probability measure.

3. Geometrically hyperfinite actions

Let Γ be a finitely generated group acting on X preserving the mean μ. Let S be a finite, symmetric generating set for Γ and S_Γ be the Schreier graph of the action. That is

- $V(S_\Gamma) = X$.
- $(x, y) \in E(S_\Gamma)$ if $x \neq y$ and there exists $s \in S$ such that $s(x) = y$

Note that we do not draw loops in our Schreier-graphs.

Let T be a subgraph of S_Γ. The edge measure of T is defined as

$$\mu_E(T) = \frac{1}{2} \int_X \deg_T(x)\, d\mu(x)\,,$$

where $\deg_T(x)$ is the degree of x in T.

We say that the action is *geometrically hyperfinite* if for any $\varepsilon > 0$, there exists $K_\varepsilon > 0$ and a subgraph $G_\varepsilon \in S_\Gamma$ such that $V(G_\varepsilon) = X$ and

$$\mu_E(S_\Gamma \backslash G_\varepsilon) < \varepsilon$$

and all the components of G_ε have size at most K_ε. It is easy to see that geometrical hyperfiniteness does not depend on the choice of the generating system. Note however, that the geometric hyperfiniteness and the hyperfiniteness of an action do depend on the choice of the invariant measure.

It is possible that for some invariant mean μ_1 the action is hyperfinite and for another invariant mean μ_2 the action is not hyperfinite, only amenable. The hyperfiniteness of a family of finite graphs was introduced in [**ECOST**]. Let $\mathcal{G} = \{G_n\}$ be a family of finite graphs with vertex degree bound d. Then $\mathcal{G}$ is called hyperfinite if for any $\varepsilon > 0$ there exists $K_\varepsilon > 0$ such that for any $n \geq 1$ one can delete $\varepsilon|V(G_n)|$ edges from G_n to obtain a graph of maximum component size at most K_ε.

Proposition 3.1. *Let S_Γ be the Schreier graph of an action of the finitely generated group Γ on X. Then the following two statements are equivalent.*

1. *S_Γ contains a hyperfinite Følner-sequence.*
2. *The action is geometrically hyperfinite with respect to some Γ-invariant mean μ.*

Recall that a Følner-sequence of S_Γ is sequence of induced subgraphs $\{F_n\}_{n=1}^\infty$, where the isoperimetric constant $i(F_n)$ tends to zero as n tends to infinity. The isoperimetric constant of a finite subgraph is the number of outgoing edges divided by the number of vertices.

Proof. Suppose that S_Γ has a hyperfinite Følner-sequence $\{F_n\}_{n=1}^\infty$. Let $G_n \subseteq F_n$ be induced subgraphs such that $\lim_{n\to\infty} \frac{|V(G_n)|}{|V(F_n)|} = 1$. Then clearly $\{G_n\}$ is a hyperfinite Følner- sequence as well. Therefore, we can suppose that $\{F_n\}_{n=1}^\infty$ are vertex-disjoint subgraphs. Indeed, let F_{n_1} be an element of the Følner- sequence such that

$$\frac{|V(F_{n_1} \backslash F_1)|}{|V(F_{n_1})|} > 1 - \frac{1}{10}\,.$$

Then let F_{n_2} be an element such that

$$\frac{|V(F_{n_1} \backslash (F_1 \cup F_{n_1})|}{|V(F_{n_2}|} > 1 - \frac{1}{100}\,.$$

Inductively, we can construct a hyperfinite Følner-sequence consisting of vertex-disjoint subgraphs. Now let ω be an ultrafilter on $\mathbb{N}$ and $\lim_\omega$ be the corresponding ultralimit $\lim_\omega : l^\infty(\mathbb{N}) \to \mathbb{R}$. Let

$$\mu(A) := \lim_\omega \frac{|A \cap V(F_n)|}{|V(F_n)|}\,.$$

Then μ is an invariant mean and the action is geometrically hyperfinite with respect to μ. Now let us suppose that μ is a Γ-invariant mean on X and the action is geometrically hyperfinite with respect to μ.

Let $\{G_\varepsilon\}_{\varepsilon>0}$ be the subgraphs of S_Γ as in the definition of hyperfiniteness. We need the following lemma.

LEMMA 3.2. *Let $R \subseteq S_\Gamma$ be a subgraph of components of size at most C. Suppose that the edge-density (number of edges divided by the number of vertices) in each component is at least α. Then $\alpha\mu(V(R)) \leq \mu_E(R)$.*

Proof. We can write R as a vertex-disjoint union $R = \cup_{i=1}^k R_i$, where all the components of R_i are isomorphic, having l_i vertices and m_i edges. Let $S_i \subset V(R_i)$ be a set containing exactly one vertex from each component. We can even suppose that under the isomorphisms of the components we always choose the same vertex.

Thus by the invariance of the mean, we have a partition

$$V(R_i) = \cup_{j=1}^{l_i} S_i^j\,,$$

where $S_i^1 = S_i$, $\mu(S_i^j) = \frac{1}{l_i}V(R_i)$, and S_i^j also has the property that it contains one vertex from each component and under the isomorphisms of the components, it always contains the same vertex. Then

$$\mu_E(R_i) = \frac{1}{2}\sum_{j=1}^{l_i} d_i^j \mu(S_i^j)\,,$$

where d_i^j is the degree in a component of R_i at a vertex of S_i^j. This yields $\mu_E(R_i) = m_i\mu(V(R_i))/l_i$. Therefore

$$\mu_E(R) = \sum_{i=1}^k \mu_E(R_i) = \sum_{i=1}^k \frac{m_i}{l_i}\mu(V(R_i)) \geq \alpha\mu(V(R))\,.$$

Now, pick a sequence $\varepsilon_1 \geq \varepsilon_2 \geq \ldots$ such that

$$\sum_{i=1}^\infty \sqrt{\varepsilon_i} < 1\,. \tag{3.1}$$

Let $\delta > 0$ be a real number and G_δ be a subgraph as above.

Let S_i^δ be the union of components of G_δ in which the edge density of $S_\Gamma \backslash G_{\varepsilon_i}$ is at least $\sqrt{\varepsilon_i}$. By the previous lemma, we have

$$\mu(V(S_i^\delta))\sqrt{\varepsilon_i} \leq \mu_E(S_\Gamma \backslash G_{\varepsilon_i})\,.$$

Hence $\mu(V(S_i^\delta)) \leq \sqrt{\varepsilon_i}$. By (3.1), for any $n \geq 1$, there exists $G'_\delta \subset G_\delta$, having the following properties.

- G'_δ is a union of components of G_δ.
- $\mu(V(G'_\delta)) > 0$.

- If Z is a component of G'_δ then the edge-density of $S_\Gamma \backslash G_{\varepsilon_i}$ inside Z is less than $\sqrt{\varepsilon_i}$, for any $1 \leq i \leq n$. That is, we can remove $\sqrt{\varepsilon_i}|V(Z)|$ edges from Z to obtain a graph of maximum component size K_{ε_i}.

For $\varepsilon > 0$ let $W^\varepsilon_\delta \subset G_\delta$ be the union of components H such that the isoperimetric constant of H is less than ε. By our previous lemma, it is easy to see that for any fixed $\varepsilon > 0$ we have

$$\lim_{\delta \to 0} \mu(W^\varepsilon_\delta) = 1\,.$$

Therefore, for any $n \geq 1$ there exists δ_n and a component H_n of G_{δ_n} such that

- the isoperimetric constant of H_n is less than $\frac{1}{n}$,
- for any $1 \leq i \leq n$ one can remove $\sqrt{\varepsilon_i}|V(H_n)|$ edges from H_n to obtain a graph of maximum component size K_{ε_i}.

This implies that $\{H_n\}_{n=1}^\infty$ is a hyperfinite Følner-sequence in S_Γ.

4. Graphs and graphings

Let T be a countable graph of vertex degree bound d, such that $V(T) = X$. Then there exists an action of a finitely generated group Γ such that T is the (loopless) Schreier graph of the action. Indeed, one can label the edges of T with finitely many labels $\{c_1, c_2, \ldots, c_n\}$ in such a way that incident edges are labeled differently. This way we obtain the Schreier graph of the n-fold free product of C_2. If μ is a Γ-invariant mean on X such that T is the Schreier graph of the action with respect to a finite symmetric generating set $S \subset \Gamma$, then μ is an H-invariant mean for any other action by a finitely generated group H with the same Schreier graph. Indeed, if $h \in H$ is a generator of H and $A \subseteq X$, then A can be written as a disjoint union

$$A = \bigcup_{s \in S} A_s \cup A_1\,,$$

where $h(x) = s(x)$ on A_s and $h(x) = x$ on A_1. Therefore,

$$\mu(h(A)) = \sum_{s \in S} \mu(s(A_s)) + \mu(A_1) = \mu(A)\,.$$

Thus if T is a graph on X with bounded vertex degree, we can actually talk about T-invariant means on X.

Let us consider a Γ-action on X preserving the mean μ and the extended Γ-action on βX preserving the associated probability measure $\widehat{\mu}$.

Let $\mathcal{G}$ be the graphing of this action on βX (see [**KECH**]) associated to a finite symmetric generating set S, that is the Borel graph on βX such that $(x, y) \in E(\mathcal{G})$ if $x \neq y$ and $s(x) = y$ for some generator s.

Lemma 4.1. *The graphing $\mathcal{G}$ depends only on T, assuming T is the Schreier graph for the action.*

Proof. Let H be another finitely generated group with generating system S', such that the Schreier graph of this action is T as well. It is enough to prove that for any $\omega \in \beta X$ and $s' \in S'$ either $s'(\omega) = \omega$ or $s'(\omega) = s(\omega)$ for some $s \in S$. Observe that there exists $s \in S$ or $s = 1$ such that

$$B = \{n \in X \mid s(n) = s'(n)\} \in \omega\,.$$

Let $A \in \omega$. Then $s(A \cap B) = s'(A \cap B)$. Hence $s(A) \in s'(\omega)$. Thus, $s(\omega) = s'(\omega)$.

We denote the graphing associated to T by $\mathcal{G}(T)$. Note that if S is a subgraph of T, such that $V(S) = A \subseteq X$ then $\mathcal{G}(S) \subseteq \mathcal{G}(T)$ and all the edges of $\mathcal{G}(S)$ are in between points of U_A.

Let us briefly recall the local statistics for graphings [**EFIN**]. A rooted, finite graph of radius r is a graph H, with a distinguished vertex x such that

$$\max_{y \in V(H)} d_H(x,y) = r\,.$$

Let U_d^r denote the finite set of rooted finite graphs of radius r with vertex degree bound d (up to rooted isomorphism).

If T is a countable graph as above, let $A(T,H) \subseteq X$ be the set of points x such that the r-neighborhood of x on T is rooted isomorphic to H.

Similarly, let $A(\mathcal{G}(T),H) \in \beta X$ be the set of points $\omega \in \beta X$ such that the r-neighbourhood of $\omega \in \mathcal{G}(T)$ is rooted isomorphic to H.

We need the labeled version of the above setup as well. Let $U_d^{r,n}$ denote the finite set of rooted finite graphs of radius r with vertex degree bound d, edge-labeled by the set $[n] = \{1,2,\dots,n\}$. Now let us label the edges of T by the set $[n]$ in such a way that incident edges are labeled differently. Then this labeling induces a Schreier graph of T and thus a labeling of $\mathcal{G}(T)$ as well. Again, for $\tilde{H} \in U_d^{r,n}$ let $A(T,\tilde{H}) \subseteq X$ be the set of points n such that the r-neighborhood of n on T is rooted isomorphic to $\tilde{H}$.

Similarly, let $A(\mathcal{G}(T),\tilde{H}) \in \beta X$ be the set of points $\omega \in \beta X$ such that the r-neighbourhood of $\omega \in \mathcal{G}(T)$ is rooted isomorphic to $\tilde{H}$. For $\tilde{H} \in U_d^{r,n}$, we denote by $[H]$ the underlying unlabeled rooted graph in U_d^r.

The following proposition states that the local statistics of T and $\mathcal{G}(T)$ coincide.

Proposition 4.2. *For any $r \geq 1$ and $H \in U_d^r$*

$$\mu(A(T,H)) = \widehat{\mu}(A(\mathcal{G}(T),H))\,.$$

Proof. Let us partition $A(T,H)$ into finitely many parts

$$A(T,H) = \cup_{\tilde{H}\,,[\tilde{H}]=H} A(T,\tilde{H})$$

Lemma 4.3. *The r-neighborhood of any ω in $U_{A(T,\tilde{H})}$ is rooted-labeled isomorpic to $\tilde{H}$.*

Proof. Let γ and δ be elements in the n-fold free product of C_2 with word-length at most r.

Suppose that $\gamma(x) = \delta(x)$ if $x \in A(T,\tilde{H})$. Then $\gamma(A) = \delta(A)$ if $A \subset A(T,\tilde{H})$, thus $\gamma(\omega) = \delta(\omega)$. Now suppose that $\gamma(x) \neq \delta(x)$ if $x \in A(T,\tilde{H})$. Then $\gamma\delta^{-1}(A(T,\tilde{H})) \cap A(T,\tilde{H}) = \emptyset$. Therefore, $\gamma(\omega) \neq \delta(\omega)$. Therefore, the rooted-labeled r-ball around ω is isomorphic to $\tilde{H}$.

By the lemma,

$$\widehat{\mu}(A(\mathcal{G}(T),H)) \geq \sum_{\tilde{H}\,,[\tilde{H}]=H} \widehat{\mu}(U_{A(T,\tilde{H})}) = \sum_{\tilde{H}\,,[\tilde{H}]=H} \mu(A(T,\tilde{H})) = \mu(A(T,H))\,.$$

Since

$$\sum_{H \in U_d^r} \mu(A(T,H)) = \sum_{H \in U_d^r} \widehat{\mu}(A(\mathcal{G}(T),H)) = 1$$

the proposition follows.

The following lemma is an immediate consequence of our proposition.

LEMMA 4.4. *Let $W \subseteq T$ be a subgraph. Then for any $l > 0$ the μ-measure of points that are in some components of size l is exactly the $\widehat{\mu}$-measure of points that are in some components of $\mathcal{G}(W)$ of size l.*

Now we prove the main result of this section.

THEOREM 4.5. *Let Γ be a finitely generated group acting on X preserving the mean μ. Then the Schreier graph S_Γ is hyperfinite if and only if the extended action on βX is hyperfinite. That is, the notions of geometric hyperfiniteness and hyperfiniteness of Γ-actions are equivalent.*

Proof. First let us suppose that S_Γ is hyperfinite. Then $S_\Gamma = W \cup Z$, where all the components of W have size at most K and $\mu_E(Z) \leq \varepsilon$. Then $\mathcal{G}(S_\Gamma) = \mathcal{G}(W) \cup \mathcal{G}(Z)$, where all the components of $\mathcal{G}(W)$ have size at most K and $\widehat{\mu}_E(\mathcal{G}(Z)) \leq \varepsilon$. Hence if S_Γ is hyperfinite, then the extended action is hyperfinite. Now let us suppose that the extended action on βX is hyperfinite. Let T be the Schreier graph of the Γ-action on X and $\varepsilon > 0$ be a real number. A basic subgraph of T is a graph (A, B, t), where

- A and B are disjoint subsets of X,
- t is a bijection between A and B with graph contained in T,
- $V(A, B, t) = X$,
- $E(A, B, t) = \bigcup_{x \in A}(x, t(x))$.

Clearly, T can be written as an edge-disjoint union

$$T = \bigcup_{i=1}^{n} (A_i, B_i, t_i)\,.$$

Then $\mathcal{G}(T) = \bigcup_{i=1}^{n}(U_{A_i}, U_{B_i}, \widehat{t_i})$ where $\widehat{t_i}$ is the extension of t_i to βX. Since $\mathcal{G}(T)$ is hyperfinite, there exists $W \subset \mathcal{G}(T)$, such that

$$W = \bigcup_{i=1}^{n} (L_i, M_i, \widehat{t_i})\,,$$

where

- $L_i \subset U_{A_i}$ is Borel for any $1 \leq i \leq n$,
- $\sum_{i=1}^{n} \widehat{\mu}(U_{A_i} \backslash L_i) \leq \frac{\varepsilon}{2}$,
- $\widehat{t}(L_i) = M_i$.
- all the components of W have size at most K .

LEMMA 4.6. *For any $\delta > 0$, there exist sets $N_i \subseteq A_i$ such that*

$$\sum_{i=1}^{n} \widehat{\mu}(L_i \triangle U_{N_i}) < \delta\,.$$

Proof. Since $\widehat{\mu}$ is regular, there exist compact sets $\{C_i\}_{i=1}^n$ and open sets $\{U_i\}_{i=1}^n$ such that

- $C_i \subset L_i \subset U_i \subset U_{A_i}$,
- $\sum_{i=1}^{n} \widehat{\mu}(U_i \backslash C_i) < \delta$.

Cover C_i by finitely many base sets $U_{A_{i,j}}$ that are contained in U_i. Since the finite union of base sets is still a base set the lemma follows.

The following lemma is straightforward.

LEMMA 4.7. *Let W be as above and let $\{W_k\}_{k=1}^\infty$ be a sequence of subgraphings of $\mathcal{G}(T)$ such that $\lim_{k\to\infty} \widehat{\mu}_E(W_k \triangle W) = 0$. Then $\lim_{k\to\infty} \widehat{\mu}(Bad_K^{W_k}) = 0$, where $Bad_K^{W_k}$ is the set of points that are in a component of W_k of size larger than K.*

By the two lemmas, we have a sequence of subgraphings

$$W'_k = \bigcup_{i=1}^n (U_{N_{i,k}}, U_{t_i(N_{i,k})}, \widehat{t_i})$$

such that

$$\lim_{k\to\infty} \widehat{\mu}(Bad_K^{W'_k}) = 0$$

and

$$\lim_{k\to\infty} \sum_{i=1}^n \widehat{\mu}(L_i \triangle U_{N_i,k}) = 0\,.$$

Now let us consider the subgraphs $H_k \subset T$

$$H_k = \bigcup_{i=1}^n (N_{i,k}, t_i(N_{i,k}), t_i)\,.$$

Then by Lemma 4.4,

$$\lim_{k\to\infty} \mu(Bad_K^{H_k}) = 0\,.$$

and

$$\limsup_{k\to\infty} \mu_E(T \backslash H_k) \leq \varepsilon\,.$$

This immediately shows that T is geometrically hyperfinite.

We finish this section with a proposition that further underlines the relation between hyperfinite p.m.p actions and hyperfinite actions on countable sets.

PROPOSITION 4.8. *Let Γ be a finitely generated group acting hyperfinitely p.m.p on a probability measure space. Then for almost all orbits, the corresponding actions are hyperfinite as well.*

Proof. We use the same idea as in the proof of Proposition 3.1. Let $\varepsilon_1 > \varepsilon_2 > \dots$ be real numbers such that

$$\sum_{n=1}^\infty \sqrt{\varepsilon_n} < \frac{1}{2}\,. \tag{4.1}$$

We may suppose that the graphing $\mathcal{G}$ of our action on the probability measure space (Z,μ) is the ascending union of subgraphings $\mathcal{G} = \cup_{n=1}^\infty \mathcal{G}_n$ such that for all $n \geq 1$

- $\mu_E(\mathcal{G} \backslash \mathcal{G}_n) < \varepsilon_n$
- all the components of $\mathcal{G}_n$ are finite.

We can also suppose that the action is ergodic, since in the ergodic decomposition $\mu = \int \mu_t d\nu(t)$, almost all the μ_t are hyperfinite actions as well [**KECH**]. Let $X_n \subset Z$ be the set of points z such that the isoperimetric constant of the component of z is greater than $\sqrt{\varepsilon_n}$. Then using the same estimate as in Lemma 3.2, one can immediately see that

$$\mu(X_n) \leq \sqrt{\varepsilon_n}\,. \tag{4.2}$$

Now let $Y^k_n \subset Z$ be the set of points y in Z such that the edge density of $\mathcal{G}\backslash\mathcal{G}_k$ in the component of $\mathcal{G}_n$ containing y is greater than $\sqrt{\varepsilon_k}$. Then

$$\mu(Y^k_n) \leq \sqrt{\varepsilon_k}\,. \tag{4.3}$$

For $k \geq 1$ we define the set $A_k \subset Z$ the following way. The point z is in A_k if

- The component of $\mathcal{G}_{k+1}$ containing z has isoperimetric constant not greater than $\sqrt{\varepsilon_{k+1}}$;
- For any $1 \leq i \leq k$ the edge-density of $\mathcal{G}\backslash\mathcal{G}_i$ in the component of $\mathcal{G}_{k+1}$ containing z is not greater than $\sqrt{\varepsilon_i}$.

By (4.1), the measure of A_k is not zero. Therefore by ergodicity, almost every point of Z has an orbit containing a point from each A_k. Let $z \in Z$ be such a point. Let F_k be the component of G_{k+1} of a point in the orbit of z. Then by the two conditions above $\{F_k\}$ is a hyperfinite Følner-sequence.

5. On faithfulness

If Γ acts on the countably infinite set preserving the mean μ, faithfulness means that for any $1 \neq \gamma \in \Gamma$ the fixed point set of γ is not the whole set. Glasner and Monod proved that for any countable group Γ the free product $\Gamma \star \mathbb{Z}$ can act on a countable set in an amenable, transitive and faithful manner. One can see however, that in their construction, if an element γ is in the group Γ, then the fixed point set of γ has mean one. We call a group action preserving a mean μ *strongly faithful* if the fixed point set of any non-unit element has μ-measure less than one.

Proposition 5.1. *If a countable group* Γ *admits an amenable, strongly faithful action on countable set, then the group is sofic.*

(see [**EHYP**] for the definition of soficity). *Proof.* Recall that such an action is called *essentially free*, if the fixed point set of any non-unit element has μ-measure zero. It is proved in [**EHYP**, Corollary 4.2] that any countable group with an amenable essentially free action is sofic. Hence, the only thing remaining is to show the following lemma.

Lemma 5.2. *Let* Γ *be a countable group acting amenably and strongly faithfully on a countably infinite set* X*, preserving the mean* μ*. Then* Γ *admits an amenable, essentially-free action on a countably infinite set.*

Proof. Let $K = \bigcup_{n=1}^\infty X^n$. Let us consider the product action of Γ on X^n. Define the mean μ_2 the following way. If $A \subset X \times X$ let

$$\mu_2(A) = \int_X \mu(\pi_1((A \cap (X,z)))d\mu(z)\,,$$

where π_1 is the projection to the first coordinate.

Clearly, μ_2 is preserved by the Γ-action. Inductively, we can construct invariant means $\{\mu_n\}_{n=1}^\infty$ on the sets $\{X^n\}_{n=1}^\infty$. Now, let ω be a non-principal ultrafilter on $\mathbb{N}$. Let us define a mean on K the following way.

$$\nu(B) = \lim_\omega \mu_n(B \cap X^n)\,.$$

Then μ is a Γ-invariant mean on K.

Let $1 \neq \gamma \in \Gamma$ and F be the fixed point set of γ in X. The fixed point set of γ in X^n is exactly F^n, and obviously,

$$\mu_n(F^n) = (\mu(F))^n .$$

Hence, $\nu(\cup_{n=1}^{\infty} F^n) = 0$ and the lemma follows.

This finishes the proof of our proposition.

6. On a problem of van Douwen

In [**DOU**] van Douwen asked the following question [Question 1.4]: If H is any countable infinite amenable group, then is there an almost free transitive action of F_2 (the free group of two generators) on H such that every invariant mean on H is F_2-invariant ?

THEOREM 6.1.

(1) *There exists no almost-free transitive action of F_2 on a finitely generated amenable group H which preserves all H-invariant means.*
(2) *For any finitely generated amenable group H, there exists a faithful, transitive action of F_2 on H which preserves all the H-invariant means.*

Proof. Let $Cay(H, S)$ be the Cayley-graph of the finitely generated group H with respect to a symmetric generating system S. Suppose that F_2 acts almost freely on H. We separate two cases for the action. **Case 1** There exists a Følner-sequence $F_1, F_2, \dots$ in $Cay(H, S)$ for which the following holds.

- $\{sF_n \cup tF_n \cup s^{-1}F_n \cup t^{-1}F_n \cup F_n\}_{n=1}^{\infty}$ are disjoint subsets, where s, t are generators of F_2.
- There exists $\varepsilon > 0$ such that for any $n \geq 1$, $\frac{|(sF_n \cup tF_n \cup s^{-1}F_n \cup t^{-1}F_n)\backslash F_n|}{|F_n|} > \varepsilon$.

We define the H-invariant mean μ by

$$\mu(A) := \lim_{\omega} \frac{|A \cap F_n|}{|F_n|} ,$$

where ω is a nonprincipal ultrafilter on $\mathbb{N}$ and $\lim_\omega$ is the corresponding ultralimit.

We claim that μ is not preserved by the F_2-action. Observe that for any $n \geq 1$ at least one of the following four inequalities hold: $|sF_n \backslash F_n| \geq \frac{\varepsilon}{4}|F_n|, |tF_n \backslash F_n| \geq \frac{\varepsilon}{4}|F_n|, |s^{-1}F_n \backslash F_n| \geq \frac{\varepsilon}{4}|F_n|, |t^{-1}F_n \backslash F_n| \geq \frac{\varepsilon}{4}|F_n|$. Hence we can assume that for the set A defined by

$$A = \{n \mid |sF_n \backslash F_n| \geq \frac{\varepsilon}{4}|F_n|\} ,$$

$A \in \omega$. Therefore

$$\mu(s \cup_{n=1}^{\infty} F_n) < 1 \quad \text{and} \quad \mu(\cup_{n=1}^{\infty} F_n) = 1 .$$

Therefore μ is not preserved by the F_2-action. **Case 2** If Følner-sequences described in **Case 1** do not exist, then any Følner-sequence is almost invariant under the F_2-action, that is

$$\lim_{n \to \infty} \frac{|(sF_n \cup tF_n \cup s^{-1}F_n \cup t^{-1}F_n)\backslash F_n|}{|F_n|} = 0 . \tag{6.1}$$

Now let us fix a Følner-sequence $\{G_n\}_{n=1}^{\infty}$ in $Cay(H, S)$.

By [**ESZA**, Proposition 4.1], $\{G_n\}_{n=1}^{\infty}$ is a hyperfinite sequence. That is, for any $\varepsilon > 0$ there exists $K_\varepsilon > 0$ such that one can remove $\varepsilon|V(G_n)|$ vertices and the incident edges in such a way that in the resulting graph G'_n, all components have size at most K_ε. By the counting argument applied in the proof of Proposition 3.1, it is easy to see that one can even suppose that all the remaining components have isoperimetric constant at most $\sqrt{\varepsilon}$ in G_n.

Now let us consider the following graph sequence $\{T_n\}_{n=1}^\infty$ edge-labeled by the set $\{s, t, s^{-1}, t^{-1}\}$

- $V(T_n) = V(G_n)$
- (x, y) is a directed edge labeled by s (resp. by t, s^{-1} or t^{-1}) if $s(x) = y$ (resp. $t(x) = y$, $s^{-1}(x) = y$ or $t^{-1}(x) = y$).

By almost-freeness and (6.1) it is clear that $\{T_n\}_{n=1}^\infty$ is a sofic approximation of F_2 (see [**ESZA**] for definition).

LEMMA 6.2. *$\{T_n\}_{n=1}^\infty$ is a hyperfinite graph sequence.*

Proof. By (6.1) there exists a function $f : \mathbb{R} \to \mathbb{R}$ such that $\lim_{x\to 0} f(x) = 0$ satisfying

$$|(sL \cup tL \cup s^{-1}L \cup t^{-1}L)\backslash L| < f(\delta)|L|$$

for any finite set $L \subset H$ with isoperimetric constant less than δ. Let $\{G'_n\}_{n=1}^\infty$ be the subgraphs of $\{G_n\}_{n=1}^\infty$ obtained by removing $\varepsilon|V(G_n)|$ vertices and the incident edges such that all the components of G'_n have size at most K_ε and G_n-isoperimetric constant at most $\sqrt{\varepsilon}$.

The number of edges of T_n that are in between the components of G'_n is less than $4f(\sqrt{\varepsilon})|V(T_n)|$. Hence by removing $4f(\sqrt{\varepsilon})|V(T_n)|$ edges from T_n and all the edges that are incident to a vertex in $V(T_n)\backslash V(G'_n)$ we can obtain a graph with maximum component size at most K_ε. Therefore $\{T_n\}_{n=1}^\infty$ is hyperfinite.
Since by [**ESZA**, Proposition 4.1], F_2 has no hyperfinite sofic approximation, we obtain a contradiction.

Therefore F_2 has no almost-free action on H that preserves all the H-invariant means.

Now we construct a faithful and transitive F_2-action on H that preserves all the H-invariant means.

First, fix a subset $\{i_n\}_{n=-\infty}^\infty \subset \mathbb{Z}$ such that $i_n < i_{n+1}$ for all $n \in \mathbb{Z}$. Then fix a function $f : \mathbb{Z} \to \{1, -1\}$. The action α of F_2 on $\mathbb{Z}$ is defined the following way. Let s and t be the generators of F_2.

- If n is odd and $i_n < j < i_{n+1}$, then let $s(j) = j$.
- If n is even and $f(i_n) = 1$, then if $i_n \le j < i_{n+1}$, let $s(j) = j + 1$. If $j = i_{n+1}$, let $s(j) = i_n$.
- If n is even and $f(i_n) = -1$, then if $i_n < j \le i_{n+1}$, let $s(j) = j - 1$. If $j = i_n$, let $s(j) = i_{n+1}$.
- If n is even and $i_n < j < i_{n+1}$, then let $t(j) = j$.
- If n is odd and $f(i_n) = 1$, then if $i_n \le j < i_{n+1}$, let $t(j) = j + 1$. If $j = i_{n+1}$, let $t(j) = i_n$.
- If n is odd and $f(i_n) = -1$, then if $i_n < j \le i_{n+1}$, let $t(j) = j - 1$. If $j = i_n$, let $t(j) = i_{n+1}$.

Note that the orbits of the F_2-action α generated by s (we call these orbits s-orbits) resp. by t are finite cycles.

Clearly, one can define $\{i_n\}_{n=-\infty}^\infty$ and $f : \mathbb{Z} \to \{1, -1\}$ in such a way that for any $1 \neq \gamma \in F_2$ there exists $n \in \mathbb{Z}$ such that $\gamma(n) \neq n$. Now let $\phi : \mathbb{Z} \to H$ be a bijection and $K > 0$ such that $d(\phi(n), \phi(n+1)) \leq K$ for any $n \in \mathbb{Z}$, where $d(x, y)$ is the shortest path distance in the Cayley graph of H.

Such bijection always exists from $\mathbb{Z}$ to an infinite connected bounded degree graph G if G has one or two ends [**SEW**]. On the other hand, the Cayley graph of an amenable group always has one or two ends [**MV**]. The F_2-action on H is given by

$$\gamma(x) := \phi(\alpha(\gamma)\phi^{-1}(x)).$$

Let μ be an H-invariant mean. We need to prove that μ is invariant under the F_2-action above.

Lemma 6.3. *Let $n \geq 1$ and let*

- $\Omega_n^s := \{x \in H \mid d(s(x), x) \geq Kn\}$
- $\Omega_n^{s^{-1}} := \{x \in \mathbb{Z} \mid d(s^{-1}(x), x) \geq Kn\}$
- $\Omega_n^t := \{x \in \mathbb{Z} \mid d(t(x), x) \geq Kn\}$
- $\Omega_n^{t^{-1}} := \{x \in \mathbb{Z} \mid d(t^{-1}(x), x) \geq Kn\}$

Then the μ-measure of any of these sets is less than $\frac{1}{n}$.

Proof. Clearly, each s-orbit of size at least $n+1$ contains at most one element of Ω_n^s. The other s-orbits are disjoint from Ω_n^s. We need to show that the union of s-orbits of the F_2-action on H of size at least $n+1$ has measure at least $n\mu(\Omega_n^s)$.

Let C be such an s-orbit of size t and let C_p be the unique vertex such that

$$d(s^i(C_p), s^{i+1}(c_p)) \leq K$$

for $i \leq t-2$. Consider the set

$$\bigcup_{C, |C| \geq n+1} C_p = L.$$

It suffices to prove that

$$\mu(L) = \mu(s(L)) = \cdots = \mu(s^n(L)) \tag{6.2}$$

Define $h_p \in H$ by $s(C_p) = h_p C_p$. Then h_p is in the K-ball around the unit element in the Cayley-graph of H. Let $L = \bigcup_{h \in B_K(1)} L^h$, where $C_p \in L^h$ if $s(C_p) = hC_p$. Then

$$\mu(s(L)) = \sum_{h \in B_k(1)} \mu(s(L^h)) = \sum_{h \in B_k(1)} \mu(L^h) = \mu(L).$$

Similarly, $\mu(s^i(L)) = \mu(L)$ if $i \leq t-1$, therefore (6.2) holds.

Now we finish the proof of the second part of our theorem. Let $A \subseteq X$. Then

$$A = \bigcup_{h \in H} A_h \quad \text{where} \quad A_h = \{x \in A \mid s(x) = hx\}$$

Obviously, $\mu(A_h) = \mu(s(A_h))$. Also, by our previous lemma,

$$\mu(\bigcup_{h, h \notin B_{Kn}(1)} s(A_h)) \leq \frac{1}{n}.$$

Therefore $\mu(s(A)) \leq \mu(A) + 1/n$. for any $n \geq 1$. Hence $\mu(s(A)) \leq \mu(A)$. However, the same way we can see that $\mu(s^{-1}(A)) \leq \mu(A)$ as well. That is, $\mu(s(A)) = \mu(A)$. Similarly, $\mu(t(A)) = \mu(A)$.

7. A topologically free, hyperfinite action of a nonamenable group

Answering a question of Grigorchuk, Nekhrashevich and Sushschanskii [**GRNS**] Gaboriau and Bergeron [**BG**] constructed a profinite, faithful, ergodic action of a nonamenable group that is not essentially free, but topologically free. Note that topological freeness of an action means that the set of points that are not in the fixed point set of any nonunit element of the group is comeager.

On the other hand, Grigorchuk and Nekrashevich constructed a profinite, ergodic action of a nonamenable group that is faithful and hyperfinite. Their construction is very far from being topologically free, in fact the set of points that are not in the fixed point set of any nonunit element is meager.

However, we prove that the two results can be combined.

Theorem 7.1. *There exists a finitely generated non-amenable group with an ergodic, faithful, profinite action that is hyperfinite and topologically free.*

Proof. An *amoeba* is a finite connected graph G (with loops) having edge-labels A, B, C, D satisfying the following properties:

- G is the union of simple cycles $\{C_i\}_{i=1}^{n_G}$. Some of the cycles might be loops. We call these cycles the basic cycles.
- Any two of the basic cycles intersect each other in at most one vertex.
- Let us consider the graph T where the vertex set of T is the set of basic cycles and two vertices are connected if and only if the corresponding cycles have non-empty intersection. Then T is a tree.
- For each vertex $x \in V(G)$ and for any label A, B, C, D there exists exactly one edge (maybe a loop) incident to x having that label. Hence the degree of any vertex is 4. Note that the contribution of a loop in the degree of a vertex is 1.
- Each loop is labeled by C or D. For any vertex x there are 0 or 2 loops incident to x.

Clearly, any amoeba is a planar graph. The minimal amoeba M has 2 vertices. The edge set of M consists of a cycle of length two labeled by A and B respectively and two loops for each of the two vertices labeled by C and D. Let G be an amoeba. A *doubling* of G is a two-fold topological covering $\phi : H \to G$ by an amoeba H constructed the following way. Let $\{C_i\}_{i=1}^{n_G}$ be the set of basic cycles of G. One way to construct H is to pick a basic cycle C_i which is not a loop and consider a two-fold covering $\psi : C_i' \to C_i$. Obviously, ψ extends to a two-fold covering $\phi : H \to G$ uniquely, and H is an amoeba as well. The second way to construct H is to pick a vertex x incident to two loops C_i and C_j. Then we cover their union by a cycle of length two. Again, this covering extends to a two-fold covering in a unique way.

7.1. The cycle-elimination tower. Let start with a minimal amoeba G_1. A *cycle-elimination tower* is a sequence of doublings

$$G_1 \overset{\phi_1}{\leftarrow} G_2 \overset{\phi_2}{\leftarrow} G_3 \overset{\phi_3}{\leftarrow} \dots$$

such that there exists a sequence of vertices $\{p_n \in V(G_n)\}_{n=1}^{\infty}$, $\phi_n(p_{n+1}) = p_n$ with the following property. For any $k \geq 1$, there exists an integer n_k such that the k-neighborhood of p_{n_k} in the graph G_{n_k} is a tree. It is easy to see that by succesively eliminating cycles, such a tower can be constructed. Let Γ be the group $C_2 \star C_2 \star C_2 \star C_2$ with free generators A, B, C, D of order 2. Note that Γ acts on the vertices of an amoeba. Indeed, the generator A maps the vertex x to the unique vertex y such that the edge (x,y) is labeled by A. If $x=y$, that is the edge is a loop, then A fixes x. Since the covering maps commute with the Γ-actions, one can extend the Γ-action to the inverse limit space $X_\Gamma = \lim_{\leftarrow} V(G_n)$. Recall that there is a natural probability measure μ on X_Γ induced by the normalized counting measures on the vertex sets $V(G_n)$.

The Γ-action on μ preserves the measure μ and in fact this is the only Borel probability measure preserved by the action. The ergodicity of the Γ-action follows from the fact that Γ acts transitively on each vertex set $V(G_n)$.

7.2. Topological freeness. In this subsection we show that the action of Γ on X_Γ is topologially free. Let us introduce some notation. If $m>n$, let ϕ_n^m be the covering map from G_m to G_n. Also, let $\Phi_n : X_\Gamma \to G_n$ be the natural covering map from the inverse limit space. We need to prove that if $1 \neq \gamma \in \Gamma$, then the fixed point set of γ has empty interior. Let $q \in V(G_n)$. Then $\Phi_n^{-1}(q)$ is an basic open set in X_Γ. It is enough to prove that there exists $z \in \Phi_n^{-1}(q)$ such that $\gamma(z) \neq z$. Let $d = dist_{G_n}(q, p_n)$, where $dist$ is the shortest path distance and $\{p_n\}_{n=1}^\infty$ is the sequence of vertices as above.

By the properties of graph coverings, for any element x in $(\phi_n^m)^{-1}(p_n)$ there exists $r \in (\phi_n^m)^{-1}(q)$ such that $dist_{G_m}(x,r) = d$. Now let $w(\gamma)$ be the wordlength of γ and consider the vertex $x = p_{n_{d+|w(\gamma)|}}$. Clearly, if $r \in G_{n_{d+|w(\gamma)|}}$ and $dist(r,x) = d$ then $\gamma(r) \neq r$. Therefore $\Phi_n^{-1}(q)$ contains a point $z \in X_\Gamma$ that is not fixed by γ.

7.3. Hyperfiniteness. Now we finish the proof of Theorem 7.1 by showing that the action of Γ on X_Γ is hyperfinite. Fix $\varepsilon > 0$. Let us recall [**BSS**] that planar graphs with bounded vertex degree form a hyperfinite family. Hence there exists $K>0$ such that for each G_n one can remove $\frac{\varepsilon}{10}|V(G_n)|$ edges in such a way that in the remaining graph G'_n the maximal component size is at most K. Since X_Γ is the inverse limit of $\{V(G_n)\}_{n=1}^\infty$ for any $p \in X_\Gamma$ there exists $m(p) \in \mathbb{N}$ such that if $l \geq m(p)$ then the $K+1$-neighborhood of p in its Γ-orbit graph and the $K+1$-neighborhood of $\Phi_l(p)$ in G_l are isomorphic. Hence we have $m>0$ such that the Haar-measure of the set A of the points in X_Γ for which the $K+1$-neighborhood of x is not isomorphic to the $K+1$-neighborhood of $\Phi_m(x) \in G_m$ is less than $\frac{\varepsilon}{10}$.

Let $\mathcal{G}$ denote the graphing of the Γ-action on X_Γ. We remove the edges from $\mathcal{G}$ that are incident to a point in X. Also, we remove the edges that are inverse images of an edge removed from $E(G_n)$. Then the edge-measure of the edges removed from $\mathcal{G}$ is less than ε and in the remaining graphing all the components have size at most K.

References

[BSS] I. Benjamini, O. Schramm and A. Shapira, *Every minor-closed property of sparse graphs are testable.* Adv. Math **223** (2010) 2200–2218. MR2601012 (2011m:05292)

[BLU] M. Blümlinger, *Lévy group action and invariant measures on $\beta\mathbb{N}$.* Trans. Amer. Math. Soc. **348** (1996) no. 12 5087–5111. MR1390970 (97d:54063)

[BG] N. Bergeron and D. Gaboriau, *Asymptotique des nombres de Betti, invariants l^2 et laminations.* Comment. Math. Helv. **79** (2004) no. 2 362–395. MR2059438 (2005d:58041)

[EHYP] G. Elek and E. Szabó, *Hyperlinearity, essentially free actions and L^2-invariants. The sofic property. Math. Ann.* **332** (2005) 421–441. MR2178069 (2007i:43002)

[ECOST] G. Elek, *The combinatorial cost. Enseign. Math* **53** (2007) no. 3-4 225–235. MR2455943 (2009m:05083)

[EFIN] G. Elek, *On limits of finite graphs. Combinatorica* **27** (2007) no. 4 503–507. MR2359831 (2008h:05106)

[ESZA] G. Elek and E. Szabó, *Sofic representations of amenable groups. to appear in Proc. of the AMS* http://arxiv.org/pdf/1010.3424 MR2823074

[GN] Y. Glasner and N. Monod, *Amenable actions, free products and a fixed point property. Bull. London Math. Soc.* **39** (2007) no. 1 138–150. MR2303529 (2008h:43001)

[GRIG] R. I. Grigorchuk and V. Nekrashevych, *Amenable actions of nonamenable groups. Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **326** (2005) 85–96. MR2183217 (2006j:43005)

[GRNS] R.I. Grigorchuk, V.V. Nekrashevich and V.I. Sushchanskii, *Automata, dynamical systems, and groups.* (Russian) Tr. Mat. Inst. Steklova 231 (2000), Din. Sist., Avtom. i Beskon. Gruppy, 134–214; translation in Proc. Steklov Inst. Math. 2000, no. 4 (231), 128–203 MR1841755 (2002m:37016)

[KAI] V. Kaimanovich, *Amenability, hyperfiniteness, and isoperimetric inequalities*, Comptes Rendus de l'Académie des Sciences - Series I - Mathematics, Volume 325, Issue 9, November 1997, Pages 999-1004 MR1485618 (98j:28014)

[KECH] A. Kechris and B. Miller, *Topics in Orbit Equivalence Theory*, Lecture Notes in Mathematics 1852, Springer (2004). MR2095154 (2005f:37010)

[MV] S. Moon and A. Valette, *Non-properness of amenable actions on graphs with infinitely many ends.* Ischia Group Theory, World Sci. Publ. Hackensack (2007) 227–233. MR2405942 (2009g:20096)

[MO1] S. Moon, *Amenable actions of amalgamated free products*, Groups, geometry and dynamics **4** (2010) 309–332. MR2595094 (2011g:43001)

[M02] S. Moon, *Permanence properties of amenable, transitive and faithful actions,* Bull. Belg. Math. Soc. Simon Stevin **18** (2011), no. 2, 287–296. MR2848804

[SEW] B. Seward, *Burnside's problem, spanning trees and tilings.* preprint http://arxiv.org/pdf/1104.1231v1

[DOU] E. K. van Douwen, *Measures invariant under actions of F_2*, Topology Appl. **34** (1990), no. 1 53–68. MR1035460 (91b:43003)

[NEU] Johan von Neumann, *Zur allgemeinen Theorie des Maßes*, Fund. Math. 13 (1929), 73–116.

Alfred Rényi Institute of the Hungarian Academy of Sciences, Realtanoda u. 13-15, 1053 Budapest, Hungary
E-mail address: `abert@renyi.hu`

Alfred Rényi Institute of the Hungarian Academy of Sciences, Realtanoda u. 13-15, 1053 Budapest, Hungary
E-mail address: `elek@renyi.hu`

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11252

A road to the spectral radius of transfer operators

A. B. Antonevich, V. I. Bakhtin, and A. V. Lebedev

Abstract. The paper is a survey of the variational principles for evaluation of the spectral radii of transfer and weighted shift operators associated with a dynamical system. These variational principles have been the matter of numerous investigations and the principal results have been achieved in the situation when the dynamical system is either reversible or it is a topological Markov chain. As the main summands these principles contain the integrals over invariant measures and the Kolmogorov–Sinai entropy. In this survey we also discuss the variational principle for an *arbitrary* dynamical system. It gives an explicit description of the Legendre dual object to the spectral potential. In general this principle contains not the Kolmogorov–Sinai entropy but a new invariant of entropy type — the t-entropy.

Contents

2010 *Mathematics Subject Classification.* Primary 37A35, 47B37; Secondary 47C15.

Key words and phrases. Transfer operator, weighted shift, functional operator, variational principle, spectral potential, ergodic measure, Legendre transform, entropy, t-entropy.

The first and the third authors were supported by the Grant of Polish Minister of Science and Higher Education.

1. Introduction

The article is a review of the spectral radius of transfer operators investigation life story.

Let us consider a compact space X, and let $\alpha : X \to X$ be a continuous mapping. This mapping generates a dynamical system with discrete time which we will denote by (X, α).

Among the operators whose spectral analysis is of prime importance in dynamical systems theory are

a) the *shift* operators, that is the operators of the form

$$T_\alpha f(x) := f(\alpha(x)), \qquad f \in F(X), \tag{1.1}$$

where $F(X)$ is a certain functional space,

b) *weighted shift* operators

$$aT_\alpha f(x) := a(x)f(\alpha(x)), \qquad f \in F(X), \tag{1.2}$$

where a is a fixed function (weight), (operators (1.2) are also called *evolution operators*), and

c) *transfer* operators (associated with the adjoint operators to weighted shift operators (see Definition 5.1)) among which the most popular one is the classical Perron–Frobenius operator, that is the operator acting in the space $C(X)$ of continuous functions on X and having the form

$$Af(x) := \sum_{y \in \alpha^{-1}(x)} \psi(y) f(y), \tag{1.3}$$

where $\psi \in C(X)$ is fixed. This operator is well defined when α is a local homeomorphism (cf. Proposition 5.3).

These operators will be the main heroes of our story.

Apart from 'pure' dynamical systems theory the operators introduced above have numerous applications in mathematical physics and in particular in thermodynamics, stochastic processes and information theory, investigations of zeta functions and Fredholm determinants, operator algebras theory, where they serve as an inexhaustible source of important examples, counterexamples and key constructive elements of the crossed product type algebras, in the theory of solvability of functional differential equations, wavelet analysis etc. We refer to the books [**72, 69, 74, 4, 15, 46, 12, 13, 24**] and recent papers [**35, 70, 44, 11, 33, 34, 9, 31**] and the bibliographies therein.

It is also worth mentioning that operators (1.1) – (1.3) are generators of the class of *functional operators*. And in connection with the latter ones it is reasonable to note that in the whole variety of really investigated linear operators one could recognize three major classes, namely, differential, integral and functional operators. In university mathematical programmes the fundamental courses are devoted to presentation of the first two classes, while the third one appears as a rule from time to time in the form of various counter examples thus manifesting special complicated properties of such operators. On the other hand it is clear that differential and integral operators are limits of functional operators and therefore the principal analysis of the first two classes of operators can be provided on this base. Thus it could be speculated that in due course the theory of functional operators will occupy an appropriate place in the university education.

2. Reversible dynamics. Spectral radius and ergodic measures

The life story of the spectral radius of operators under consideration originates in the second half of the 20-th century and we start its presentation with a discussion of the spectral radius of weighted shift operators with invertible shifts.

As a model example here one can take X to be a compact topological space and $\alpha : X \to X$ to be a continuous invertible mapping. As $F(X)$ we take the space $C(X)$ of continuous functions or the spaces $L^p(X,m)$, $1 \le p \le \infty$, where m is an α-invariant measure on X whose support coincides with X (that is any open set has a non zero measure). Note that the assumption of α-invariance of the measure m here is not essential, it is presumed only for simplicity of presentation — under this assumption shift operator T_α is isometric in each of the spaces $L^p(X,m)$, $1 \le p \le \infty$. For the weight a in the model example we assume that $a \in C(X)$.

It seems that the first to examine the spectral properties of weighted shift operators (1.2) was Parrot [**68**], who, in particular, described certain parts of spectra of weighted shift operators in $L^p(X,m)$ under assumption that m is α-ergodic. A lot of authors have studied the spectral radius for concrete examples of X and α by means of different methods. A weighted shift operator in $C(S^1)$ (S^1 is the unit circle) in the case when a homeomorphism α has a finite number of periodic points was investigated by Kravchenko [**47**] and the corresponding results for $L^p(S^1)$, $1 \le p < \infty$ were obtained by Karlovich and Kravchenko [**41**]. Here it is reasonable to note that the spectra of functional operators in $C(S^1)$ in the situation when α is periodic mapping $\alpha^n = 1$ were investigated even by Babbage [**20**] and the corresponding results for $L^p(S^1)$, $1 \le p < \infty$ are due to Litvinchuk [**61**] (see also Antonevich [**4**] for a general finite group action discussion). The spectra of weighted shift operators generated by irrational rotation of the circle was investigated by Karapetyants and Samko [**39, 40**] and Antonevich and Ryvkin [**16**]. In connection with this it is worth mentioning the paper by Mukhamadiev, Sadovski [**63**], where an estimate for the spectral radius of an irrational rotation weighted shift operator was obtained, and this estimate appeared to be sharp. The spectrum of weighted shift operators in the situation when α is an intransitive diffeomorphism of the circle was studied by Lebedev [**57**]. Multidimentional generalizations of the Karlovich and Kravchenko results, namely, spectra of weighted shift operators generated by diagonal shifts of the torus and Morse-Smale diffeomorphisms were considered by Antonevich, Volpert [**17**] and Lebedev [**56**]. Spectra of weighted shift operators where the action α is included in an action on X of a compact group were investigated by Antonevich [**2**].

The general result on the spectral radius of weighted shift operator for the model example is given by the next theorem which in essence can be extracted from the results presented in [**4, 15**].

THEOREM 2.1. *If $\alpha\colon X \to X$ is a continuous* invertible *mapping and m is an α-invariant measure on X whose support coincides with X, and $a \in C(X)$, then in the space $F(X) = L^p(X,m)$, $1 \le p \le \infty$, as well as in the space $C(X)$ the following formulae for the spectral radius of weighted shift operator* (1.2) *are valid*:

$$\ln r(aT_\alpha) = \max_{\mu \in M_\alpha} \int_X \ln |a(x)|\, d\mu, \tag{2.1}$$

$$\ln r(aT_\alpha) = \max_{\mu \in EM_\alpha} \int_X \ln |a(x)|\, d\mu. \tag{2.2}$$

Here M_α is the set of all Borel probability α-invariant measures on X and EM_α is the subset of M_α consisting of all ergodic measures.

From the 'ideological' point of view this result shows that the spectral properties of weighted shift operators depend on *all* α-invariant measures while the properties of shift operator T_α depend only on the measure m.

The statements of this type are called in dynamical systems theory and related fields of analysis *variational principles* and we will come across a number of them in this survey.

Abstract weighted shift with invertible shift. The conditions that X is a compact topological space and α is a continuous mapping in fact are not restrictive which can be approved by the following argument. Let us consider a probability measure space (Y, m) and a measure preserving invertible mapping $\gamma : Y \to Y$. Routine computation shows that for any $a \in L^\infty(Y, m)$ we have

$$\left[\left(T_\gamma a T_\gamma^{-1}\right) f\right](y) = a(\gamma(y)) f(y), \qquad f \in L^p(Y, m), \quad 1 \le p \le \infty \tag{2.3}$$

which means that the operator T_γ generates an *automorphism* $\hat{T}_\gamma$ of the algebra $\mathcal{A} = L^\infty(Y, m)$: $\hat{T}_\gamma(a) = T_\gamma a T_\gamma^{-1}$. The Gelfand transform establishes an isomorphism between $\mathcal{A}$ and $C(X)$, where X is the maximal ideal space of $L^\infty(Y, m)$ and under this isomorphism the automorphism $\hat{T}_\gamma$ has the form

$$\hat{T}_\gamma(a)(x) = a(\alpha(x)) \tag{2.4}$$

where α is a homeomorphism of X (in this formula and henceforth in this section we identify an element $a \in \mathcal{A}$ with its Gelfand transform — the element of $C(X)$).

Therefore in an axiomatic way it is reasonable to consider the following object.

Definition 2.2. Let $\mathcal{A} \subset L(H)$ be a subalgebra of the algebra $L(H)$ (of all linear continuous operator acting on a Banach space H) that is isometrically isomorphic to $C(X)$ for a certain compact X; and let $T : H \to H$ be an invertible isometry such that $T\mathcal{A}T^{-1} = \mathcal{A}$. Then an operator of the form aT, $a \in \mathcal{A}$ will be called an (abstract) *weighted shift operator* (with an invertible shift). Algebra $\mathcal{A}$ is called the algebra of (abstract) *weights* and T is an (abstract) *shift* operator (with an invertible shift).

For the spectral radius of an abstract weighted shift operator we have the following generalization of Theorem 2.1.

Theorem 2.3. *Let aT be an abstract weighted shift operator described in Definition* 2.2 *and α be a homeomorphism of the maximal ideal space X generated by the automorphism $\hat{T}$ (see* (2.4)*). Then*

$$\ln(r(aT)) = \max_{\mu \in M_\alpha} \int_X \ln|a(x)|\, d\mu = \max_{\mu \in EM_\alpha} \int_X \ln|a(x)|\, d\mu. \tag{2.5}$$

The variational principle (2.5) was stated by Antonevich and for a number of concrete situations it has been proved for example in [**1, 2**] where one can also find the corresponding explicit calculation of the set M_α. In the general form (for an arbitrary homeomorphism α) the principle was established by Lebedev [**55**] and Kitover [**45**]. The applications of formula (2.5) to the calculation of the spectral radii of various weighted shift operators are given in [**14, 4, 15**].

In view of (2.4), the isometric property of T and the equality

$$\|a\|_{L(H)} = \max_{x\in X} |a(x)|$$

we have the next simple formula for the norms of the powers of aT:

$$\|(aT)^n\|_{L(H)} = \max_{x\in X} \prod_{k=0}^{n-1} \bigl|a\bigl(\alpha^k(x)\bigr)\bigr|. \tag{2.6}$$

Namely the usage of this formula is one of the key moments in the proof of Theorem 2.3. As we shall see in the subsequent sections in the situations when the shift operator is generated by an *irreversible* mapping the expression for the norm of the powers of the weighted shift operator is essentially more complicated (it contains some mean values of sums of analogous products) and at this moment transfer operators come into play. As a result the arising variational principles will take into account also the entropy and stochastic nature of α.

REMARK 2.4. It is clear from the observations preceding Theorem 2.3 that as the algebra $\mathcal{A}$ one can take for example any closed subalgebra of $L^\infty(Y,m)$ which is invariant with respect to T_γ and T_γ^{-1}. If we take $\mathcal{A} = L^\infty(Y,\mu)$ then its maximal ideal space is extremely complicated and so there is no hope to find an explicit description of the set M_α. Thus Theorem 2.3 takes into account the complexity of the algebra $\mathcal{A}$: 'simpler' algebra $\mathcal{A}$ implies 'simpler' maximal ideal space structure and therefore one obtains an easier description of the set M_α.

Spectral radius and the Lyapunov exponents. If one considers weighted shift operators in the spaces of vector-valued functions in the situation when the algebra $\mathcal{A}$ of weights is not a commutative algebra of operator-valued functions then the corresponding variational principle for the spectral radius takes into account not integrals by ergodic measures but their generalizations, the so called Lypunov exponents. This variational principle was established by Latushkin and Stepin [**53, 54**] and here we present in brief their result in the form of a model example.

Let X be a compact metrizable space and $\alpha\colon X \to X$ be a homeomorphism. Consider a separable Hilbert space H and the Hilbert space of vector-valued functions $B := L^2((X,m),H))$ where m is an α-invariant measure on X whose support coincides with X. Let $L(B)$ be the algebra of linear bounded operators in B and $\mathcal{A} \subset L(B)$ be the subalgebra of operators of multiplication by continuous operator-valued functions, that is $\mathcal{A} := C(X, L(H))$, and let $T_\alpha f(x) := f(\alpha(x)),\ \ f \in B$. Clearly T_α is a unitary operator and $[T_\alpha a {T_\alpha}^*](x) = a(\alpha(x)),\ \ a \in \mathcal{A}$.

For every $a \in \mathcal{A}$ one can define the cocycle $\boldsymbol{a} : X \times \mathbb{Z}_+ \to L(H)$ by means of the formulae

$$\boldsymbol{a}(x,n) := a(\alpha^n(x)) \times \cdots \times a(\alpha(x)), \qquad \boldsymbol{a}(x,0) := I.$$

If $a \in \mathcal{A}$ is a continuous operator-valued function having compact values then by the multiplicative ergodic theorem (see, for example Ruelle [**73**]) we have that *for any α-ergodic measure μ there exists an α-invariant set $X_\mu \subset X,\ \ \mu(X_\mu) = 1$, such that for any $x \in X_\mu$ there exists the limit*

$$\lambda_\mu^1 = \lim_{n\to\infty} \ln \|\boldsymbol{a}(x,n)\|.$$

This limit λ_μ^1 is called the (first) Lyapunov (characteristic) exponent (here we used the upper index 1 simply to note that there are also other Lyapunov exponents, $\lambda_\mu^2, \dots$).

Observe that if $\mathcal{A}$ is a commutative algebra (that is $\mathcal{A} = C(X)$) and $a(x) \neq 0$ then by the Birkhoff–Hinchin ergodic theorem for every $x \in X_\mu$ we have

$$\lim_{n\to\infty} \ln \|\boldsymbol{a}(x,n)\| = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \ln\bigl|a\bigl(\alpha^i(x)\bigr)\bigr| = \int_X \ln|a(x)|\, d\mu$$

which means that in this case we have

$$\lambda_\mu^1 = \int_X \ln|a(x)|\, d\mu.$$

So the statement of Theorem 2.3 in this situation has the form

$$\ln r(aT) = \max_{\mu\in EM_\alpha} \{\lambda_\mu^1\}.$$

Precisely in this form it was generalized by Latushkin and Stepin to the variational principle for the spectral radius of a weighted shift operator with an operator-valued weight a. Their result looks as follows.

Theorem 2.5 ([**53, 54**]). *Let X, α, $\mathcal{A}$ and T_α be the objects mentioned above, and let $a \in \mathcal{A}$ be a continuous operator-valued function having compact values. Then*

$$\ln r(aT_\alpha) = \sup_{\mu\in EM_\alpha} \{\lambda_\mu^1\}.$$

Remark 2.6. 1) In [**53, 54**] Latushkin and Stepin not only established the variational principle for the spectral radius but also estimated other components of the spectrum by means of the corresponding Lyapunov exponents.

2) In [**15**] there was obtained a different variational principle for the spectral radius of operators mentioned in Theorem 2.5, namely, it exploits not the Lyapunov exponents but an appropriate extension of X and the corresponding version of Theorem 2.3.

3. Irreversible-reversible dynamics

If $\alpha\colon X \to X$ is *not invertible* the results related to the corresponding formulae for the spectral radius can be divided into a number of classes. In this section we present the situations where the variational principles for the spectral radius are of the same form as for reversible dynamical systems considered in Section 2.

First of all one should mention here the case when the weighted shift operators (1.2) act in the spaces of $C(X)$ or $L^\infty(X,m)$ type. In this case formulae (2.1), (2.2) preserve their form (see, for example, [**62**]).

Reversible extensions of irreversible dynamical systems. The second situation is much more sophisticated. This is the case when $\alpha : X \to X$ is *not invertible*, the space where the weighted shift operator (1.2) acts is of $L^p(X,m)$, $1 \le p < \infty$ type, and at the same time its spectral radius calculation can be reduced to that for another weighted shift operator which is associated with a certain *invertible* mapping $\tilde\alpha\colon \tilde X \to \tilde X$. An example of such situation is the classical *unilateral* weighted shift operator acting in $l^p(\mathbb{N})$. Its spectral radius calculation can be easily reduced to that for the corresponding *bilateral* weighted shift operator acting in $l^p(\mathbb{Z})$.

A general construction of this sort (for *arbitrary* $\alpha : X \to X$) is based on the procedure of extension of the algebra $A = C(X)$ of weights up to a certain algebra $C(\tilde{X}) = \tilde{A} \supset A$ in such a way that the mapping $\alpha\colon X \to X$ is 'extended' up to an *invertible* mapping $\tilde{\alpha}\colon \tilde{X} \to \tilde{X}$. This construction is based on a method of calculation of the maximal ideal space of the corresponding algebras and the arising reversible extensions of endomorphisms. It was developed by Kwasniewski and Lebedev [**50, 48**]. In particular, as a result one obtains a unified method of construction of fractals of Smale's horse shoe type, Browder–Janishewski–Knaster continua and so on.

Thus in the situation mentioned the variational principle in essence preserves its form (2.5) while the (extended) space $\tilde{X}$, algebra $C(\tilde{X})$ and homeomorphism $\tilde{\alpha}$ differ drastically from the initial space X, algebra $C(X)$ and the mapping $\alpha\colon X \to X$. We are not going here into details (see in this connection [**48, 49, 50**]) since our main goal is the presentation of a variety of variational principles arising in essentially irreversible dynamical systems. This is the theme of the subsequent sections.

4. Irreversible dynamics. Weighted shifts and Perron–Frobenius operators. Spectral radius and entropy

Now we proceed to consideration of operators associated with essentially irreversible dynamics in the spaces $L^p(X, m)$, $1 \le p < \infty$, that is we will discuss operators whose spectral properties can not be reduced to the study of certain reversible extensions as in the preceding section.

As we will see the situation in this case changes drastically. At this point the deep 'entropy' and 'stochastic' nature of the spectrum of weighted shift and transfer operators springs out.

Variational principle for weighted shifts associated with topological Markov chains. The starting principal results for irreversible dynamics have been achieved by Latushkin and Stepin [**51, 52, 54, 76**] under a rather special assumption on the nature of the mapping α. Namely, in the case when α is a *topological Markov chain* (in particular α can be an expanding k-sheeted cover of a manifold X). To get a hint on the arising phenomena let us consider the following

EXAMPLE 4.1. Let $X = S^1 = \{z : |z| = 1\} = \{e^{i2\pi t} : t \in [0, 1)\} = \mathbb{R} \pmod 1$, and $\alpha(t) = 2t$.

Let us consider the operator aT_α, $a \in C(X)$ in the space $L^p(X, m)$, $1 \le p < \infty$, where m is the normalized Lebesgue measure. For this operator we have

$$\|aT_\alpha\|_{L^p(X,m)} = \max_{t\in[0,2\pi]} \left(\frac{1}{2} \sum_{y\in\alpha^{-1}(t)} |a(y)|^p \right)^{1/p} = \|A\|^{1/p}_{C(X)},$$

where

$$Af(t) := \frac{1}{2} \sum_{y\in\alpha^{-1}(t)} |a(y)|^p f(y)$$

is a transfer operator (see (1.3)). Calculating the norms of the powers of aT_α one gets

$$\left\|(aT_\alpha)^n\right\|_{L^p(X,m)} = \left\|A^n\right\|^{1/p}_{C(X)}.$$

In general in the case when α is a topological Markov chain (in particular α can be an expanding k-sheeted cover of a manifold X) one arrives at the following formula

$$\big\|(aT_\alpha)^n\big\|_{L^p(X,m)} = \big\|A^n\big\|_{C(X)}^{1/p}, \tag{4.1}$$

where $A\colon C(X)\to C(X)$ is a transfer (Perron–Frobenius) operator of the form

$$Af(x) = \sum_{y\in\alpha^{-1}(x)} \big(|a(y)|^p\rho(y)\big)f(y); \tag{4.2}$$

here ρ is a certain continuous nonnegative function defined by this mapping and such that

$$\sum_{y\in\alpha^{-1}(x)} \rho(y)\equiv 1.$$

Calculation of the norm of aT_α by means of formulae (4.1), (4.2) shows, in particular, that in $L^p(X,m)$ the norm of the weighted shift operator aT_α with noninvertible shift is *not* equal to the maximum of the weight a as for invertible shift (see (2.6)) but it is equal to the maximum of the weight *averaged* over inverse images and is related to the norm of the corresponding transfer operator in $C(X)$.

Equality (4.1) implies the equality

$$r(aT_\alpha) = r(A)^{1/p}. \tag{4.3}$$

Latushkin and Stepin proved the following variational principle for the spectral radius of operator (1.2) in $L^p(X,m)$, $1\le p<\infty$ [**51, 52, 54**].

THEOREM 4.2. *If α is the shift on a unilateral topological Markov chain* (*in particular, if α is an expanding m-sheeted cover of a manifold X*) *then in the space $L^p(X,m)$, $1\le p<\infty$ we have*

$$\ln r(aT_\alpha) = \sup_{\mu\in M_\alpha}\left\{\int_X \ln|a(x)|\,d\mu + \frac{1}{p}\left(\int_X \ln\rho(x)\,d\mu + h(\mu)\right)\right\}, \tag{4.4}$$

where $h(\mu)$ is the Kolmogorov–Sinai entropy of the measure μ with respect to the mapping α.

Here, namely the speed of averaging (mixing) under the powers of α is reflected in appearance of the entropy summand in the right-hand part of the variational principle (4.4).

REMARK 4.3. **Variational principles: what does it mean calculation.** In connection with the variational principles as presented above, so also with those that will follow it is resonable to discuss what does it mean to calculate a certain value by means of such formulae. Here the ‘calculation’ is realized as derivation of the expression for the spectral radius in terms of other quantities — in this case the dynamical invariants. Originally the problem has been understood precisely as calculation — in order to apply say formula (2.2) one has to know the α-invariant measures and the problem was assumed to be solved when the description of these measures was obtained. At the next step of generalization the problem is divided into two problems: 1) derivation of a general formulae (like (2.2), (4.4)); and 2) calculation of invariant measures and dynamical invariants involved in it. It should be noted that in many examples the second problem does not have an *explicit* solution but once one obtains the complete answer to the first mentioned part of the problem the problem on the calculation of the spectral radius is agreed to be solved.

Spectral radius and the Lyapunov exponents. Just as in Section 2 one can considers weighted shift operators in the situation under study in the spaces of vector-valued functions in the case when the algebra A of weights is not a commutative algebra of operator-valued functions. Here the corresponding variational principle (counterpart of Theorem 2.5) was also proved by Latushkin and Stepin. The precise description of the situation that has been investigated looks as follows.

Let X be a compact metrizable space and $\alpha : X \to X$ be a topological Markov chain. Consider the Hilbert space H of vector-valued functions $H := L^2((X,m),\mathbb{C}^n)$ where m is an α-invariant measure on X whose support coincides with X. Let $L(H)$ be the algebra of linear bounded operators in H and $A \subset L(B)$ be the subalgebra of operators of multiplication by continuous matrix-valued functions, that is $A := C(X, L(\mathbb{C}^n))$, and let $T_\alpha f(x) := f(\alpha(x)),\ \ f \in H$. The variational principle in this case has the form

Theorem 4.4 ([**54**]). *Let α be the shift on a unilateral topological Markov chain, and let T_α and A be the objects mentioned above, and let $a \in A$ be a continuous operator-valued function having compact values. Then in the space $L^p(X,m)$, $1 \le p < \infty$ we have*

$$\ln r(aT_\alpha) = \sup_{\mu \in EM_\alpha} \left\{ \lambda_\mu^1 + \frac{1}{p}\left(\int_X \ln \rho(x)\, d\mu + h(\mu) \right) \right\}.$$

Perron–Frobenius operators and the Lyapunov exponents. We have already remarked the relation between the spectral properties of weighted shift and Perron–Frobenius operators. Theorem 4.4 has its natural counterpart for Perron–Frobenius operators acting in the spaces of vector-valued functions. The corresponding variational principle was obtained by Campbell and Latushkin [**29**], and here we present their result.

Let α be an expanding m-sheeted cover of a compact connected manifold X. Consider the space $C(X,\mathbb{C}^n)$ of continuous functions with values in $\mathbb{C}^n$ and a (matrix) Perron–Frobenius operator $A\colon C(X,\mathbb{C}^n) \to C(X,\mathbb{C}^n)$ given by the formula

$$Af(x) = \sum_{y \in \alpha^{-1}(x)} a(y) f(y) \tag{4.5}$$

where $a(\,\cdot\,)$ is a continuous matrix-valued function. For this operator we have the next variational principle.

Theorem 4.5 ([**29**]). *Let A be the Perron-Frobenius operator* (4.5)*. Then*

$$\ln r(A) = \sup_{\mu \in EM_\alpha} \left\{ \lambda_\mu^1 + h(\mu) \right\}.$$

Remark 4.6. This result can also be extracted from [**54**] (Theorem 4.15).

Spectral radius and topological pressure. In connection with formula (4.4) it is reasonable to recall the variational principle for the *topological pressure* established by Ruelle [**71**] and Walters [**77**]:

$$P(\alpha, c) = \sup_{\mu \in M_\alpha} \left\{ \int_X c(x) d\mu + h(\mu) \right\}, \tag{4.6}$$

where $c \in C(X)$ is a nonnegative function and $P(\alpha,c)$ is the topological pressure defined by α and c. We would like to stress that in contrast to (4.4) the mapping $\alpha\colon X \to X$ in (4.6) is an *arbitrary* continuous mapping.

Comparing formulae (4.4), (4.3), and (4.6) we see that in the case when α is a topological Markov chain the following relation is valid

$$\ln r(aT_\alpha) = \frac{1}{p}\,\ln r(A) = \frac{1}{p}\,P(\alpha, \ln\psi), \tag{4.7}$$

where $\psi = |a|^p\rho$.

The equality in the right-hand part of (4.7), namely,

$$\ln r(A) = P(\alpha, \ln\psi), \tag{4.8}$$

has been known, probably, since [**72**]. In fact, the establishment of the relation between the spectral radius of weighted shift operators and topological pressure was the essence of the Latushkin–Stepin work. This link along with the observed relation between the spectral radii of weighted shift and transfer operators serves as a basis for numerous applications of these operators and also inspires the investigation of their spectral properties in various functional spaces and, in particular, in the spaces of smooth functions and smooth vector-valued functions (see, for example, [**30, 26, 29, 37, 25, 38**]). We have to stress again that all the mentioned sources deal only with the case when α is a topological Markov chain.

Remark 4.7. 1) In general (that is for an arbitrary continuous mapping α, and even when α is a local homeomorphism) the equalities $\ln r(aT_\alpha) = \frac{1}{p}P(\alpha, \ln\psi)$ and $\ln r(A) = P(\alpha, \ln\psi)$ (see (4.7), (4.8)) are not true and (4.4) is *not* a generalization of the variational principle (2.1). For example, let us consider an invertible mapping α. Then $\rho \equiv 1$. Let us set $a \equiv 1$, thus $\psi \equiv 1$ and we have

$$\ln r(T_\alpha) = 0,$$

while

$$P(\alpha, 0) = h(\alpha),$$

where $h(\alpha)$ is the topological entropy of the dynamical system (X, α), and in general $h(\alpha)$ could be equal to any nonnegative number.

2) A different proof of the Latushkin–Stepin formulae was obtained by Maslak and Lebedev [**59, 60**] by means of newly introduced topological invariants that also gave a number of estimates for the spectral radius. In addition it was shown in [**59, 60**] that the variational principle (2.1) and the Latushkin–Stepin result are in a way the 'extreme points' of the situations one could come across when dealing with calculation of the spectral radius $r(aT_\alpha)$.

3) Note that if $p \to \infty$ then formula (4.4) transforms into (2.1) and this agrees with the fact (that has been already noted in Section 3) that in the spaces of $C(X)$ and $L^\infty(X)$ type the variational principle (2.1) preserves its form.

4) Observe that in a general situation (when $\alpha^{-1}(x)$ may be an uncountable set) there is no function ρ at all (recall that $\rho(y),\ y \in \alpha^{-1}(x)$ is defined by a normalized measure on $\alpha^{-1}(x)$). Therefore in general not even all of the components of the variational principle (4.4) exist.

The goal of the subsequent sections is a description the variational principles for the spectral radii of transfer and weighted shift operators for an *arbitrary* dynamical system (X, α). It will be shown that in general these variational principles contain *not* the Kolmogorov–Sinai entropy $h(\mu)$ as in (4.4) but a *new* dynamical characteristic which we call *t-entropy*. The dynamical and stochastic meanings of t-entropy will be uncovered as well.

5. Spectral potential of a transfer operator. Convexity, Legendre duality and thermodynamics

In the previous section we have noted an essential role played by transfer (Perron–Frobenius) operators in the spectral analysis of weighted shift operators. This and subsequent sections will show that namely transfer operators are the key heroes of the whole story.

Recall that by X we denote a compact space, and in this section $\alpha\colon X \to X$ is an arbitrary continuous mapping, and $C(X)$ is the algebra of continuous real-valued functions on X.

DEFINITION 5.1. A linear operator $A : C(X) \to C(X)$ is called a *transfer operator* for the dynamical system (X, α) if

a) A is positive (that is it maps nonnegative functions to nonnegative) and

b) it satisfies the *homological identity*

$$A\big(f \circ \alpha \cdot g\big) = f A g, \qquad f, g \in C(X).$$

If in addition this operator maps $\mathbf{1}$ into $\mathbf{1}$ we will call it a *conditional expectation operator.*

REMARK 5.2. 1) A conditional expectation operator satisfies the equality

$$A(f \circ \alpha) = A\big((f \circ \alpha)\mathbf{1}\big) = f A\mathbf{1} = f \quad \text{for every} \quad f \in C(X).$$

Thus this operator is a positive left inverse to the mapping $f \mapsto f \circ \alpha$.

2) If $A\colon C(X) \to C(X)$ is a transfer operator and $\frac{1}{A\mathbf{1}} \in C(X)$ then $\frac{1}{A\mathbf{1}} A$ is a conditional expectation operator.

The next proposition shows that Perron–Frobenius operators could be considered as ‘typical’ transfer operators and in addition it shows that in reversible dynamics (Section 2) the whole story was also about transfer operators.

PROPOSITION 5.3 ([**10**]). *Let X be a compact space and $\alpha\colon X \to X$ be a local homeomorphism. Then any transfer operator $A\colon C(X) \to C(X)$ has the form*

$$Af(x) = \sum_{y \in \alpha^{-1}(x)} a(y) f(y), \tag{5.1}$$

where $a \in C(X)$ is a certain fixed nonnegative function (*so A is a Perron–Frobenius operator*)*. In particular, if $\alpha : X \to X$ is a homeomorphism, then any transfer operator $A\colon C(X) \to C(X)$ has the form*

$$\big[Af\big](x) = a(x) f\big(\alpha^{-1}(x)\big),$$

where $a \in C(X)$ is a certain nonnegative function (*and so A is a weighted shift operator*)*.*

The next model example gives a general idea of transfer operators. In this example we exploit not $C(X)$ but the spaces of measurable functions, although this is not essential (see remark, following the example).

Model example. Let $(Y, \mathfrak{A}, m)$ be a measurable space with a σ-finite measure m, and let β be a measurable mapping such that for all measurable sets $G \in \mathfrak{A}$ the following estimate holds

$$m\big(\beta^{-1}(G)\big) \le C m(G),$$

where the constant C does not depend on G. For example, if the measure m is β-invariant one can set $C = 1$. Let us consider the space $L^1(Y, m)$ of real-valued integrable functions and the shift operator that takes every function $f \in L^1(Y, m)$ to $f \circ \beta$. Clearly the norm of this operator does not exceed C. The mapping

$$\delta f := f \circ \beta \tag{5.2}$$

acts also on the space $L^\infty(Y, m)$ and it is an endomorphism of this space. As is known, the dual space to $L^1(Y, m)$ coincides with $L^\infty(Y, m)$. Define the linear operator $B\colon L^\infty(Y, m) \to L^\infty(Y, m)$ by the identity

$$\int_Y f \cdot g \circ \beta \, dm \equiv \int_Y (Bf) g \, dm, \qquad g \in L^1(Y, m).$$

In other words B is the adjoint operator to the shift operator in $L^1(Y, m)$. If one takes as g the index functions of measurable sets $G \subset Y$, then the latter identity takes the form

$$\int_{\beta^{-1}(G)} f \, dm \equiv \int_G Bf \, dm.$$

Therefore Bf is nothing else than the Radon–Nikodim density of the additive set function $\mu_f(G) = \int_{\beta^{-1}(G)} f \, dm$. Evidently, the operator B is positive and satisfies the homological identity

$$B\big((\delta f) g\big) = f Bg, \qquad f, g \in L^\infty(X, m).$$

We see that B is a transfer operator. And in the case when m is β-invariant measure it is a conditional expectation operator.

Remark 5.4. 1) In the example presented we exploited the space $L^\infty(Y, m)$ to define a transfer operator, but in fact the whole picture could be given in terms of continuous functions as well. Indeed, by means of Gelfand transform one can identify the algebra $L^\infty(Y, m)$ with the corresponding algebra $C(X)$, where X is the maximal ideal space of $L^\infty(Y, m)$. Under this identification endomorphism (5.2) of $L^\infty(Y, m)$ is transformed into endomorphism of $C(X)$ that has the form

$$C(X) \ni f \mapsto f \circ \alpha \in C(X), \tag{5.3}$$

where $\alpha : X \to X$ is a continuous mapping. And in this setting operator B is transformed into a transfer operator $A\colon C(X) \to C(X)$ for the dynamical system (X, α). We used $L^\infty(Y, m)$ instead of $C(X)$ simply to make the idea of transfer operators more visible.

2) A general description of transfer operators for an arbitrary dynamical system (X, α) is given in [**10**].

Spectral potential. Now we proceed to the introduction of the objects playing the central role in our story.

Given a transfer operator A we define a family of operators $A_\varphi\colon C(X) \to C(X)$ depending on the functional parameter $\varphi \in C(X)$ by means of the formula

$$A_\varphi f = A(e^\varphi f).$$

Evidently, all the operators of this family are transfer operators as well. Let us denote by $\lambda(\varphi)$ the logarithm of the spectral radius of A_φ, that is

$$\lambda(\varphi) = \lim_{n\to\infty} \frac{1}{n} \ln \|A_\varphi^n\|.$$

The positivity of transfer operator implies that

$$\lambda(\varphi) = \lim_{n\to\infty} \frac{1}{n} \ln \|A_\varphi^n \mathbf{1}\|, \tag{5.4}$$

where $\mathbf{1}$ is the unit function on X, and $\|f\|$ denotes the uniform norm of the function $f \in C(X)$. The functional $\lambda(\varphi)$ is called the *spectral potential* or the *spectral exponent* of the transfer operator A (depending on whether we have in mind dynamical or spectral associations). In this paper when dealing with the objects associated with $\lambda(\varphi)$ we are staying on the platform of the dynamical (entropy, thermodynamics, information, stochastics) point of view and therefore throughout the paper $\lambda(\varphi)$ will be called the *spectral potential*.

Our goal is the investigation of $\lambda(\varphi)$.

The next starting proposition gives a list of its principal elementary properties.

PROPOSITION 5.5 ([**10**]). *The spectral potential $\lambda(\varphi)$ is either identically equal to $-\infty$ on the whole of $C(X)$ or takes only finite values on $C(X)$ and possesses the following properties:*

a) (monotonicity) *if $\varphi \le \psi$, then $\lambda(\varphi) \le \lambda(\psi)$;*

b) (additive homogeneity) *$\lambda(\varphi + t) = \lambda(\varphi) + t$ for any $t \in \mathbb{R}$;*

c) (strong α-invariance) *$\lambda(\varphi + \psi) = \lambda(\varphi + \psi \circ \alpha)$ for all $\varphi, \psi \in C(X)$;*

d) (convexity) *$\lambda(t\varphi + (1-t)\psi) \le t\lambda(\varphi) + (1-t)\lambda(\psi)$ for all $\varphi, \psi \in C(X)$ and $t \in [0, 1]$;*

e) (Lipschitz property) *$\lambda(\varphi) - \lambda(\psi) \le \|\varphi - \psi\|$ for all $\varphi, \psi \in C(X)$.*

In particular, the spectral potential is continuous.

Legendre analysis. Proposition 5.5 shows in particular that the spectral potential is a convex functional on $C(X)$. As is known among the standard instruments of investigation of convex functionals is the Legendre transform. Here we recall the principal notions and facts related to this transform (in essence they are borrowed from [**32**]).

Let f be a functional on a real Banach space L with the values in the extended real straight line $\bar{\mathbb{R}} = [-\infty, +\infty]$. The set $D(f) = \{\varphi \in L \mid f(\varphi) < +\infty\}$ is called the *effective domain* of the functional f. The functional f is called *convex*, if for all $\varphi, \psi \in D(f)$ and $t \in [0, 1]$ the following inequality holds

$$f\bigl(t\varphi + (1-t)\psi\bigr) \le tf(\varphi) + (1-t)f(\psi).$$

The functional f is called *lower semicontinuous* if the set $\{\varphi \in L \mid f(\varphi) > c\}$ is open for any $c \in \bar{\mathbb{R}}$. One can speak about lower semicontinuity with respect to the norm topology or with respect to the weak topology on L, but for convex functionals these properties are equivalent.

Let L^* be the dual space to L. The functional $f^* \colon L^* \to \bar{\mathbb{R}}$ that is defined on the dual space by the equality

$$f^*(\mu) = \sup_{\varphi\in L} \bigl\{\mu(\varphi) - f(\varphi)\bigr\} = \sup_{\varphi\in D(f)} \bigl\{\mu(\varphi) - f(\varphi)\bigr\}, \qquad \mu \in L^*, \tag{5.5}$$

is called the *Legendre dual* to the functional f (or the Legendre transform of f). For a functional g on the dual space the Legendre transform is defined as the functional on the initial space given by the similar formula:

$$g^*(\varphi) = \sup_{\mu\in L^*} \bigl\{\mu(\varphi) - g(\mu)\bigr\}, \qquad \varphi \in L. \tag{5.6}$$

PROPOSITION 5.6. *Let a functional $f\colon L \to (-\infty, +\infty]$ be not identically equal to $+\infty$. Then*

*a) the dual functional f^* is convex and lower semicontinuous with respect to *-weak topology on the dual space*;

b) if the functional f is convex and lower semicontinuous then $f = (f^*)^*$ (the Legendre transform is involutory);

c) in general $(f^*)^*$ is the maximal convex lower semicontinuous functional that does not exceed f (the convex hull of f).

The analogous statements are valid for functionals $g\colon L^* \to (-\infty, +\infty]$.

We have already noted that the spectral potential $\lambda(\varphi)$ is convex and continuous (see Proposition 5.5). Therefore it can be represented as the Legendre transform of its Legendre dual on the dual space $C(X)^*$. However we will slightly modify the form of the record of this duality. The matter is that in thermodynamics, information theory and ergodic theory there is a tradition to change the sign of the dual to $\lambda(\varphi)$ and the result obtained is called the *entropy.* Following this tradition (see the discussion of Thermodynamic Formalism below and, in particular, Proposition 5.12) we define the *dual entropy* $S(\mu)$ of the spectral potential $\lambda(\varphi)$ by means of the formula

$$S(\mu) := \inf_{\varphi\in C(X)} \bigl\{\lambda(\varphi) - \mu(\varphi)\bigr\}, \qquad \mu \in C(X)^*. \tag{5.7}$$

Since the dual entropy differs from the dual to $\lambda(\varphi)$ functional only by sign it follows that $S(\mu)$ is concave and upper semicontinuous (with respect to *-weak topology). As the Legendre transform is involutory the next equality holds true

$$\lambda(\varphi) = \sup_{\mu\in C(X)^*} \bigl\{\mu(\varphi) + S(\mu)\bigr\}. \tag{5.8}$$

This equality is in fact *the simplest form* of the further subject of the review — the *Variational Principle* for $\lambda(\varphi)$ (the main one is Theorem 7.3)

REMARK 5.7. In particular, the Kolmogorov–Sinai entropy $h(\mu)$ arising in the right-hand parts of (4.4) and (4.6) is nothing else than manifestation of the fact that $h(\mu)$ is the Legendre dual object to the logarithm of the spectral radius of weighted shift operator associated with the topological Markov chain and it is the Legendre dual object to the topological pressure for an arbitrary α respectively.

Thermodynamic formalism. Before proceeding to a deeper discussion of the objects introduced above we would like to recall that by means of the Legendre analysis one could uncover the thermodynamical nature of convex objects. The material is borrowed from [**11**].

Throughout this subsection $\lambda(\varphi)$ is an arbitrary convex lower semicontinuous function on a real Banach space L with the values in the extended real straight line $\bar{\mathbb{R}} = [-\infty, +\infty]$, and we will make use of a natural notation of inner product of a vector and functional

$$\langle \varphi, \mu\rangle := \mu(\varphi). \tag{5.9}$$

In this setting the corresponding analogue to formulae (5.7) and (5.8) have the form

$$S(\mu) := \inf_{\varphi\in L} \bigl\{\lambda(\varphi) - \langle \varphi, \mu\rangle\bigr\}, \qquad \mu \in L^*, \tag{5.10}$$

and

$$\lambda(\varphi) = \sup_{\mu \in L^*} \big\{ \langle \varphi, \mu \rangle + S(\mu) \big\}. \tag{5.11}$$

Definition (5.10) implies the Young inequality

$$S(\mu) \le \lambda(\varphi) - \langle \varphi, \mu \rangle. \tag{5.12}$$

Obviously the Young inequality becomes an equality precisely for those vectors φ at which the infimum in (5.10) is attained and precisely for those functionals μ at which the supremum in (5.11) is attained. We define a Lagrangian manifold $\mathcal{L} \subset L \times L^*$ as the set of pairs (φ, μ) at which the Young inequality becomes an equality. In other words,

$$\mathcal{L} := \big\{ (\varphi, \mu) \bigm| S(\mu) = \lambda(\varphi) - \langle \varphi, \mu \rangle \big\}. \tag{5.13}$$

Proposition 5.8. *The following three conditions are equivalent*:
a) a pair (φ, μ) belongs to $\mathcal{L}$;
b) the linear functional μ is a subgradient of the functional λ at the point φ;
c) the vector φ is a subgradient of the function $-S$ at the point μ.

Corollary 5.9. *The Lagrangian manifold $\mathcal{L}$ is the graph of the subgradient of the function $\lambda(\varphi)$.*

It is useful to consider the Lagrangian manifold $\mathcal{L}$ as a natural domain of definition for the function $\lambda(\varphi)$. The point is that the subgradient of $\lambda(\varphi)$ may take many values, or it may not exist at all. Nevertheless, Proposition 5.8 allows us to define it as a single-valued function on $\mathcal{L}$. Namely, we define the subgradient of a function $\lambda(\varphi)$ at the point $(\varphi, \mu) \in \mathcal{L}$ to be linear functional $\lambda'(\varphi) = \mu$ and its differential to be $d\lambda := \langle \delta\varphi, \mu \rangle$. Similarly, the differential of a function $S(\mu)$ at the point $(\varphi, \mu) \in \mathcal{L}$ is defined as $dS := -\langle \varphi, \delta\mu \rangle$. These subgradients and differentials coincide with the usual derivatives and differentials at all points $(\varphi, \mu) \in \mathcal{L}$ at which the functions $\lambda(\varphi)$ and $S(\mu)$ are Gateaux differentiable. But even at the points where they are not differentiable, the formal equalities $d\lambda := \langle \delta\varphi, \mu \rangle$ and $dS := -\langle \varphi, \delta\mu \rangle$ have a meaning which is clarified by the following assertion.

Proposition 5.10. *If points (φ, μ) and $(\varphi + \delta\varphi, \mu + \delta\mu)$ belong to $\mathcal{L}$, then there exists a $\Theta \in [0, 1]$ such that*

$$\lambda(\varphi + \delta\varphi) - \lambda(\varphi) = \langle \delta\varphi, \mu \rangle + \Theta \langle \delta\varphi, \delta\mu \rangle, \tag{5.14}$$

$$S(\mu) - S(\mu + \delta\mu) = \langle \varphi, \delta\mu \rangle + (1 - \Theta) \langle \delta\varphi, \delta\mu \rangle. \tag{5.15}$$

Proposition 5.11. *For any vector $h \in L \setminus \{0\}$ and a number $U \in \mathbb{R}$*
a) the inequality

$$\sup_{\langle h, \mu \rangle} S(\mu) \le \inf_{b \in \mathbb{R}} \{ \lambda(bh) - bU \} \tag{5.16}$$

holds;
b) if the infimum on the right-hand side of (5.16) *is attained at some $b = b_0$ and the function $\lambda(\varphi)$ is continuous at the point $b_0 h$, then* (5.16) *becomes an equality, the supremum on the left-hand side of* (5.16) *is attained at some functional $\mu_0 \in L^*$, and every such functional μ_0 satisfies the condition $(b_0 h, \mu_0) \in \mathcal{L}$.*

Now having in mind the foregoing observations we can pass to Thermodynamic Formalism. We refer to the elements $\varphi \in L$ as *phase potentials* and to the elements $\mu \in L^*$ as *states.* The function $U(\varphi, \mu) := \langle \varphi, \mu \rangle$ is called *internal energy.* On the Cartesian product $L \times L^*$ there are two natural first order differential forms, the *heat form* $\delta Q := \langle \varphi, \delta\mu \rangle$ and the *work form* $\delta W := -\langle \delta\varphi, \mu \rangle$. Obviously,

$$dU(\varphi, \mu) = \delta Q - \delta W.$$

This identity is usually called the energy preservation law, or *the first principle of thermodynamics.* The convex lower semicontinuous functional $\lambda(\varphi)$ is called *thermodynamic potential* and its dual functional $S(\mu)$ defined by formula (5.10) is called *entropy.* We call the subgradients of the thermodynamic potential $\lambda(\varphi)$ *equilibrium states.*

Let us fix a phase potential $h \in L$. The *isolated thermodynamic system* (TDS) with energy U is the set of pairs (h, μ) for which $\langle h, \mu \rangle = U$. An *equilibrium* TDS is a pair (h, μ) with energy U at which the entropy attains a conditional maximum.

Consider an equilibrium TDS with the same phase potential h but with different energy levels U. Let I denote the set of all numbers $b \in \mathbb{R}$ for which the function $\lambda(\varphi)$ is continuous at the point bh. It is known that the set of continuity points of any convex function is open and convex. Therefore, I is an interval. Suppose that it is nonempty. Consider the convex function $\lambda(bh)$ of scalar argument $b \in I$. Let $U(b)$ denote an arbitrary subgradient of this function at the point h. Obviously, $U(b)$ is a (set-valued) nondecreasing function on I taking all values between $\inf_I U(b)$ and $\sup_I U(b)$. By virtue of Proposition 5.10 b), for each energy level U between $\inf_I U(b)$ and $\sup_I U(b)$, there exists an equilibrium thermodynamic system (h, μ) with energy U. There must exist a number $b \in I$ such that the pair (bh, μ) belongs to Lagrangian manifold $\mathcal{L}$. By Proposition 5.8, μ is the equilibrium state corresponding to the potential bh. The number

$$T := -\frac{1}{b}$$

is called the *temperature* of the equilibrium thermodynamic system (h, μ). The monotonicity of the function $U(b)$ implies that the energy of an equilibrium system does not decrease with increasing temperature. The multivalence points of $U(b)$ correspond to *phase transitions* of the first kind in physics (i.e., to the presence of equilibrium states with different energies at the same temperature, eg., in ice melting).

We define the equilibrium manifold $\mathcal{E} \subset L \times L^* \times \mathbb{R}$ as the set of triples (h, μ, T) for which the pair (h, μ) is an equilibrium TDS with temperature T, or, equivalently, for which the pair $\bigl(-\frac{h}{T}, \mu\bigr)$ belongs to the Lagrangian manifold $\mathcal{L}$. Since $\mathcal{L}$ is specified by the equation $S(\mu) = \lambda(\varphi) - \langle \varphi, \mu \rangle$, the equilibrium manifold $\mathcal{E}$ is determined by the formula

$$\mathcal{E} = \left\{ (h, \mu, T) \in L \times L^* \times \mathbb{R} \;\middle|\; S(\mu) = \lambda\left(-\frac{h}{T}\right) + \left\langle \frac{h}{T}, \mu \right\rangle \right\}. \tag{5.17}$$

Proposition 5.12 (The second principle of thermodynamics). *On the equilibrium manifold $\mathcal{E}$, the identity*

$$\delta Q = T dS$$

holds. More precisely, if points (h, μ, T) *and* $(h + \delta h, \mu + \delta\mu, T + \delta T)$ *belong to* $\mathcal{E}$, *then, for some* $\Theta \in [0, 1]$,

$$\langle h, \delta\mu \rangle = T\big(S(\mu + \delta\mu) - S(\mu)\big) + \frac{\Theta}{T + \delta T}\big\langle h\delta T - T\delta h, \delta\mu\big\rangle. \tag{5.18}$$

REMARK 5.13. Classical thermodynamics considers families of phase potentials $h = h(a)$ differentiably depending on a finite-dimensional parameter $a = (a_1, a_2, \ldots, a_n)$. The components a_i are called external parameters. For thermodynamical systems of the form $(h(a), \mu)$, the numbers

$$A_i = -\frac{\partial U(h(a), \mu)}{\partial a_i}$$

are called thermodynamic forces, or internal parameters. In this notation the work form has the standard thermodynamic expression

$$\delta W = -\langle \delta h(a), \mu \rangle = \sum_i A_i da_i.$$

For equilibrium TDSs with potentials $h(a)$ it is desirable to represent all quantities as functions of external parameters and temperature. For instance, a dependence $U = U(a, T)$ is called a caloric equation, and dependencies $A_i = A_i(a, T)$ are thermal equations.

6. Spectral potentials and dual entropy

The discussion presented in the previous section implies the existence of the dual entropy, and its explicit construction by means of formula (5.7) is possible provided that the spectral potential is known. However, the spectral potential (which is the object of our study) itself is a rather hard object to investigate. The principal goal of this and subsequent sections is the *independent* derivation of the *explicit* formula for the dual entropy, *not leaning* on the spectral potential. This formula allows, in particular, to impart a more effective character to the variational principles for transfer and weighted shift operators.

REMARK 6.1. We would like to stress that the definition of the Kolmogorov–Sinai entropy $h(\mu)$ (which is the dual entropy for the topological pressure $P(\alpha, c)$ (see (4.6))) does not lean on the Legendre transform duality and one can consider it as an explicit calculation of the dual entropy for $P(\alpha, c)$.

In this section we present the starting properties of the dual entropy, while the main result and the corresponding variational principle will be given in Section 7.

Henceforth $\lambda(\varphi)$ is the spectral potential (5.4).

An *equilibrium measure*, corresponding to a function $\varphi \in C(X)$, is an arbitrary subgradient of the functional $\lambda(\varphi)$ at the point φ (in other words it is a linear functional $\mu\colon C(X) \to \mathbb{R}$ such that $\lambda(\varphi + \psi) - \lambda(\varphi) \geq \mu(\psi)$ for all $\psi \in C(X)$). Evidently the set of all equilibrium measures corresponding to a certain function φ is convex and closed (with respect to the *-weak topology). This set is nonempty by the convex analysis theorem on the existence of a supporting hyperplane. It consists of a unique measure μ if and only if there exists the Gâteaux derivative $\lambda'(\varphi)$. In this case $\mu = \lambda'(\varphi)$.

The definitions of the dual entropy (5.7) and equilibrium measure imply

PROPOSITION 6.2. *For any function $\varphi \in C(X)$ and any measure $\mu \in C(X)^*$ the Young inequality holds true*

$$S(\mu) \le \lambda(\varphi) - \mu(\varphi). \tag{6.1}$$

This inequality turns into equality iff μ is an equilibrium measure corresponding to φ.

The following propositions are borrowed from [**21**], [**10**].

PROPOSITION 6.3. *The effective domain of S is contained in M_α, that is if $S(\mu) > -\infty$, then μ is a probability and α-invariant measure. In particular this is true for all equilibrium measures.*

This proposition shows that it suffices to define the dual functional (dual entropy) only on invariant probability measures. This will be done in the next section.

PROPOSITION 6.4. *If $S(\mu) > -\infty$, then μ belongs to the closure of the set of equilibrium measures* (*with respect to the norm of the space $C(X)^*$*).*

Proposition 6.3 implies that the supremum in (5.8) is attained on the set of invariant probability measures M_α. Therefore, every spectral potential has the form

$$\lambda(\varphi) = \sup_{\mu \in M_\alpha} \bigl\{\mu(\varphi) + S(\mu)\bigr\}, \tag{6.2}$$

where $S(\mu)$ is the corresponding dual entropy.

Proposition 6.2 implies in turn that supremum in (6.2) is in fact maximum that is

$$\lambda(\varphi) = \max_{\mu \in M_\alpha} \bigl\{\mu(\varphi) + S(\mu)\bigr\}, \tag{6.3}$$

and this maximum is attained precisely on equilibrium measures, corresponding to the function φ.

The forgoing observation implies in addition that the uniqueness of an extremal measure in (6.3) is equivalent to the existence of the Gâteaux derivative $\lambda'(\varphi)$.

7. t-entropy and variational principle for transfer operators

In this section we give a direct definition for the dual entropy not leaning on the Legendre duality (as in (5.7), (5.8)) but only on the properties of the initial dynamical system and the transfer operator chosen. As a result we obtain the variational principle for transfer operators for an arbitrary dynamical system (Theorem 7.3) which can be considered as a complete spectral counterpart of the variational principle for topological pressure (4.6).

As it was already noted Proposition 6.3 implies that it suffice to define the dual entropy only on invariant probability measures.

Let $A\colon C(X) \to C(X)$ be a transfer operator with the spectral potential $\lambda(\varphi)$, associated with a dynamical system (X, α), where α is a continuous mapping of a compact space X.

DEFINITION 7.1 ([**23**]). *t-entropy* is the functional τ on M_α such that its value at $\mu \in M_\alpha$ is defined by the following formula

$$\tau(\mu) = \inf_{n,D} \frac{1}{n} \sum_{g \in D} \mu(g) \ln \frac{\mu(A^n g)}{\mu(g)}. \tag{7.1}$$

The infimum in (7.1) is taken over all the partitions of unity D in the algebra $C(X)$. If we have $\mu(g) = 0$ for a certain function $g \in D$, then we set the corresponding summand in (7.1) to be zero independently of the value $\mu(A^n g)$. And if there exists a function $g \in D$ such that $A^n g \equiv 0$ and simultaneously $\mu(g) > 0$, then we set $\tau(\mu) = -\infty$.

The next two assertions can be extracted from [**10**] and [**23**].

PROPOSITION 7.2.

$$\tau(\mu) = \lim_{n\to\infty} \inf_D \frac{1}{n} \sum_{g\in D} \mu(g) \ln \frac{\mu(A^n g)}{\mu(g)}.$$

In the next theorem we consider transfer operators defined in Section 5 and call this situation the *model setting*. This is because of the fact that general operator algebraic picture (as abstract weighted shift operators in Section 2) can be given as well, and in essence, it can be reduced to the model setting (for details see [**10**]). This theorem means, in particular, that t-entropy $\tau(\mu)$ coincides with the dual entropy $S(\mu)$.

THEOREM 7.3 (Variational Principle in the model setting). *Let $A\colon C(X) \to C(X)$ be a transfer operator for a continuous mapping $\alpha\colon X \to X$ of a Hausdorff compact space X. Then its spectral potential $\lambda(\varphi)$ satisfies the variational principle*

$$\lambda(\varphi) = \max_{\mu\in M_\alpha} \bigl\{\mu(\varphi) + \tau(\mu)\bigr\}, \qquad \varphi \in C(X), \tag{7.2}$$

and t-entropy satisfies the equality

$$\tau(\mu) = \inf_{\varphi\in C(X)} \bigl\{\lambda(\varphi) - \mu(\varphi)\bigr\}, \qquad \mu \in M_\alpha. \tag{7.3}$$

REMARK 7.4. 1) In fact the equalities established in this theorem are much deeper than simply the explicit calculation of the Legendre dual objects arising in the procedure of spectral radius evaluation. Much more important (from our point of view) is the observation that formula (7.2) links the spectral characteristics of the transfer operator (the left-hand part) with the stochastic characteristics ($\tau(\mu)$ in the right-hand part) of the dynamical system. This ideology will be developed further in Section 8 where in particular the interrelation between $\tau(\mu)$ and the distribution of empirical measures is described.

2) Formula (7.2) reveals the partition of the process of calculation of the spectral radius into the *static component* (the first summand in the right-hand part depends only on the weight φ) and the *dynamical component* (the second summand depends only on the shift α and the transfer operator A).

3) The duality established in Theorem 7.3 and the thermodynamic formalism developed in Section 5 leads naturally to introduction of the thermodynamic 'ideology' into the spectral analysis of transfer operators. Having in mind this motivation it is reasonable to call the functionals (measures) μ at which the maximum in the right-hand part of (7.2) is attained the *equilibrium states*. We recall in this connection that in accordance with a common physical point of view the equilibrium states are the states at which the system 'exists in reality'. From this point of view the duality principle adds dialectics to the spectral analysis of transfer operators: since $\tau(\mu)$ describes the measure of the 'most typical' trajectories (see, in particular, Entropy Statistic Theorem in Section 8) and the value $\mu(\varphi)$ calculates the 'living conditions' (recall the corresponding discussion in [**8**]) then the duality principle

tells us that the process realizes at a state having the best combination of these components.

8. Properties of t-entropy. Entropy Statistic Theorem and Variational Principle for t-entropy

In this section we present a number of properties of t-entropy and in particular its upper semicontinuity. In addition we uncover the statistical nature of t-entropy and show that t-entropy itself satisfies certain variational principle.

The following Propositions 8.1 – 8.5 are borrowed from [**10**].

PROPOSITION 8.1. *The functional $\tau(\mu)$ satisfies the inequality $\tau(\mu) \le \lambda(0)$.*

REMARK 8.2. In particular, if A is a conditional expectation operator (i. e., $A\mathbf{1} = \mathbf{1}$) then $\|A^n\| = 1$ and therefore $\lambda(0) = 0$. Thus in this case $\tau(\mu) \le 0$.

PROPOSITION 8.3. *If $A\colon C(X) \to C(X)$ is an invertible conditional expectation operator then $\tau(\mu) = 0$ for any $\mu \in M$.*

PROPOSITION 8.4. *$\tau(\mu)$ depends concavely on $\mu \in M$.*

PROPOSITION 8.5. *$\tau(\mu)$ is upper semicontinuous on M.*

Now we present a certain result on the interrelation between t-entropy and the statistics of the distribution of empirical measures. This result is important in its own right and plays for $\tau(\mu)$ a role similar to that the Shannon–McMillan–Breiman theorem plays for $h(\mu)$ (see, for example [**27**], Chapter 4). The Entropy Statistic Theorem (Theorem 8.6) not only uncovers the statistical nature of $\tau(\mu)$ but also serves as the main technical instrument in the proof of the variational principle for transfer operators (Theorem 7.3).

Consider an arbitrary point $x \in X$. The *empirical measures* $\delta_{x,n} \in M$ are defined by the formula

$$\delta_{x,n}(f) = \frac{1}{n}\Big(f(x) + f(\alpha(x)) + \cdots + f\big(\alpha^{n-1}(x)\big)\Big), \qquad f \in C(X). \tag{8.1}$$

Evidently, the measure $\delta_{x,n}$ is concentrated on the trajectory of the point x of length n.

We endow the set M of probability measures (positive normalized functionals) with the *-weak topology of the dual space to $C(X)$. Given a measure $\mu \in M$ and a neighborhood $O(\mu)$ of it we define the sequence of sets $X_n(O(\mu))$ as follows:

$$X_n(O(\mu)) := \big\{x \in X \bigm| \delta_{x,n} \in O(\mu)\big\}. \tag{8.2}$$

The next theorem is due to Bakhtin [**22, 10**].

THEOREM 8.6 (Entropy Statistic Theorem). *Let $A\colon C(X) \to C(X)$ be a certain transfer operator for (X, α). Then for any measure $\mu \in M$ and any number $t > \tau(\mu)$ there exist a neighborhood $O(\mu)$ in the *-weak topology, a (large enough) number $C(t, \mu)$ and a sequence of functions $\chi_n \in C(X)$ majorizing the index functions of the sets $X_n(O(\mu))$ such that for all n the following estimate holds*

$$\|A^n\chi_n\| \le C(t, \mu)e^{nt}.$$

The following assertion (that can be extracted from [**10**] and [**23**]) shows that t-entropy itself satisfies a certain variational principle. We would like to stress that supremum here is taken over the set M of all probability measures, which are not necessarily α-invariant.

THEOREM 8.7 (Variational principle for t-entropy).

$$\tau(\mu) = \lim_{n\to\infty} \inf_{D} \frac{1}{n} \sup_{m\in M} \sum_{g\in D} \mu(g) \ln \frac{m(A^n g)}{\mu(g)}\,, \tag{8.3}$$

where M is the set of all probability measures.

We would like to emphasize that the main mathematical basis and t-entropy idealogy take their roots in the papers [**5, 6, 7, 8, 22**]. Namely, particular cases of t-entropy and the corresponding variational principles have been considered in [**5, 6, 7, 8**] for the situation when the initial transfer operators are the conditional expectation operators and the definition of t-entropy was introduced there in a *different* way, and in [**22**] the variational principle for the spectral radius of weighted shift operators with positive weights in $L^1(X,m)$ was proved. The interrelation between the definition of t-entropy given in this survey (Definition 7.1) and alternative definitions is analysed in [**10, 22**].

9. Variational principle for weighted shift operators

The example given in Section 4 (introduction of transfer operator when calculating the norm of weighted shift operator) shows the tight interrelation between transfer operators and weighted shift operators. Developing the idea of this example we go further and present in this section the variational principle for the spectral radius of weighted shift operators acting in $L^p(Y,m)$ spaces.

Model example. Just as in the Model Example in Section 5, let $(Y,\mathfrak{A},m)$ be a measurable space with a σ-finite measure m, and β be a measurable mapping of Y into itself satisfying the condition

$$m\big(\beta^{-1}(G)\big) \le Cm(G), \qquad G \in \mathfrak{A}, \tag{9.1}$$

where the constant C does not depend on G. Let us consider the space $L^p(Y,m)$, where $1 \le p \le \infty$. Set the *shift operator* T (generated by the mapping β) by the formula

$$[Tf](x) = f(\beta(x)), \qquad f \in L^p(Y,m). \tag{9.2}$$

Inequality (9.1) implies that the norm of this operator does not exceed $C^{1/p}$.

Note that in the case when $p=\infty$ this operator defines an *endomorphism* of the algebra $L^\infty(Y,m)$, that is

$$T(fg) = Tf \cdot Tg, \qquad f,g \in L^\infty(Y,m).$$

Just as in the Model Example in Section 5, we define the linear operator $A\colon L^\infty(Y,m) \to L^\infty(Y,m)$ by means of the identity

$$\int_Y f\cdot g\circ\beta\, dm \equiv \int_Y (Af) g\, dm, \qquad g\in L^1(Y,m), \quad f\in L^\infty(Y,m) \tag{9.3}$$

(in other words, A is adjoint to the shift operator T on $L^1(Y,m)$).

The definition implies that A is a positive operator and it satisfies the homological identity

$$A\big((Tf)g\big) = fAg, \qquad f,g \in L^\infty(X,m). \tag{9.4}$$

Therefore A is a transfer operator (for the C^*-dynamical system $(L^\infty(Y,m),T)$). And in the case of a β-invariant measure m it is a conditional expectation operator.

For any $\psi \in L^\infty(Y,m)$ the operator ψT acting on $L^p(Y,m)$ and given by

$$\psi T\colon f \mapsto \psi \cdot Tf$$

will be called a *weighted shift* operator (with the weight ψ). Note, in particular, that

$$T\psi = T(\psi)T, \qquad \psi \in L^\infty(Y,m). \tag{9.5}$$

REMARK 9.1. We would like to stress the usage of measurable functions in the foregoing Model Example and not continuous functions is not essential. Indeed, just as it was remarked in Section 5, by means of Gelfand transform one can identify the algebra $L^\infty(Y,m)$ with the corresponding algebra $C(X)$, where X is the maximal ideal space of $L^\infty(Y,m)$. Under this identification the shift mapping $f \mapsto f \circ \beta$ on $L^\infty(Y,m)$ is identified with the shift mapping $g \mapsto g \circ \alpha$ on $C(X)$. Finally, the set $M_\alpha(X)$ of all α-invariant probability measures on X is identified with the set of all β-invariant finitely additive probability measures on Y which are absolutely continuous with respect to m. We will denote the latter set by $M_\beta(Y,m)$.

Since the set $M_\beta(Y,m)$ consists of *finitely* additive measures one can come across certain difficulties when defining the integrals by these measures for unbounded functions, and namely such integrals are needed in the next theorem. Fortunately, we can introduce them in a rather natural way by using the corresponding measures on $C(X)$. Namely, let $\psi \in L^\infty(Y,m)$ and let $\hat{\psi} \in C(X)$ be its Gelfand transform, let also $\mu \in M_\beta(Y,m)$ and let $\hat{\mu} \in M_\alpha(X)$ be the corresponding measure mentioned above. Then we set

$$\int_Y \ln|\psi|\, d\mu := \int_X \ln|\hat{\psi}|\, d\hat{\mu}. \tag{9.6}$$

THEOREM 9.2 (Variational principle for weighted shift operators, [**10**]). *For the spectral radius of the operator $\psi T : L^p(Y,m) \to L^p(Y,m)$, $1 \le p < \infty$, the following variational principle holds:*

$$\ln r(\psi T) = \max_{\mu \in M_\beta(Y,m)} \left\{ \int_Y \ln|\psi|\, d\mu + \frac{\tau(\mu)}{p} \right\}, \tag{9.7}$$

where $\tau(\mu)$ is the t-entropy assigned to the transfer operator (9.3) *and the integral is understood in the sense of* (9.6) .

REMARK 9.3. 1) If $p \to \infty$ then formula (9.7) transforms into the formula

$$\ln r(\psi T) = \max_{\mu \in M_\beta(Y,m)} \left\{ \int_Y \ln|\psi|\, d\mu \right\}.$$

This restores the variational principle for the space $L^\infty(Y,m)$ (cf. Section 3).

2) Similarly to the method of passage from model examples to a general operator algebraic description of weighted shift operators and the corresponding variational principles in reversible dynamics (recall Section 2) one can implement the corresponding procedure in the irreversible situation under consideration (for details see [**10**]).

10. Multiterm functional operators. Variational principles etc.

Multiterm functional operators (generated by weighted shift operators) are the operators of the form

$$Wu(x) = \sum_{k=1}^{s} a_k(x)u(\alpha_k(x)), \tag{10.1}$$

where $\alpha_k\colon X \to X$ are given mappings and a_k are given functions. These operators have more complicated structure in comparison with a single weighted shift operator, and their spectral properties are studied only for a number of rather special cases (see, for example, [**3, 19, 18, 42, 43, 28**]). In this section we present this trend of investigation.

Just as in the analysis of weighted shift operators in the spaces $L^p(X,m)$ reversible and irreversible situations differ essentially, moreover in $L^p(X,m)$ the results also depend on the properties of the group or semigroup G generated by the given mappings. On the other hand in the space $C(X)$ the result is formulated in the same way for invertible and noninvertible mappings and does not depend on the group properties.

Spectral potential of a multiterm functional operator. In [**18, 19**] it is shown that variational principles for the spectral exponent of operators (10.1) possess properties similar to that discussed above in the case of a single term operator. Henceforth we present these results in brief.

Let us consider, for definiteness, the situation when X is a compact space with a Borel measure m, and the mappings α_k, $k \in \Delta = \{1, 2, \dots, s\}$ are continuous (and can be noninvertible). We assume that the measure m is quasiinvariant with respect to the mappings given, and for each α the formula

$$[T_\alpha u](x) := \left(\frac{dm_\alpha(x)}{dm(x)}\right)^{1/p} u(\alpha(x)), \tag{10.2}$$

where m_α is the measure given by $m_\alpha(E) := m(\alpha^{-1}E)$, and dm_α/dm is the Radon–Nikodim derivative, defines an isometry T_α in $L^p(X,m)$, $1 \le p < \infty$.

For each set of functions $\varphi_k \in C(X)$, $k = 1, 2, \dots, s$ we consider a multiterm functional operator

$$W = \sum_{k\in\Delta} e^{\varphi_k} T_{\alpha_k} = \sum_{k\in\Delta} a_k T_{\alpha_k}, \qquad a_k = e^{\varphi_k}, \tag{10.3}$$

where in the case of $L^p(X,m)$, $1 \le p < \infty$ the operator T_{α_k} is given by (10.2), while in $C(X)$ we have $[T_{\alpha_k}u](x) := u(\alpha_k(x))$.

Remark 10.1. One can consider operators with coefficients a_k, belonging to a certain subalgebra $\mathcal{A} \subset L^\infty(X,m)$, that is invariant with respect to all the mappings α_k. In this case all the results will be analogous to that described below but with replacement of the space X by the space $M(\mathcal{A})$ of the maximal ideals of the coefficient algebra. This approach allows one to obtain more effective results in a number of situations, but in contrast to single weighted operator even in the case of constant coefficients ($\mathcal{A} = \mathbb{C}$) the spectral exponent calculation for operators (10.3) for arbitrary mappings α_k is a profound problem.

The set $\varphi = (\varphi_1, \dots, \varphi_s)$ is an element of the space $C(X)^s$ of real-valued continuous vector-functions on X, which is naturally isomorphic to the space $C(Y)$, where $Y = X \times \Delta$. One has a natural partial order on $C(Y)$: if $\varphi = (\varphi_1, \dots, \varphi_s)$ and $\psi = (\psi_1, \dots, \psi_s)$, then $\varphi \le \psi$ means that $\varphi_i(x) \le \psi_i(x)$ for all i.

The *spectral potential* $\lambda\colon C(Y) \to \mathbb{R}$ of the multiterm operator operator W is given by the formula

$$\lambda(\varphi) = \ln r(W). \tag{10.4}$$

Elementary properties of the spectral potential (10.4) for operators in the classical spaces $C(X)$ and $L^p(X, \mu)$ are presented in the next

PROPOSITION 10.2. *The spectral potential $\lambda(\varphi)$ possesses the following properties*:

a) (monotonicity) *if $\varphi \le \psi$, then $\lambda(\varphi) \le \lambda(\psi)$;*

b) (additive homogeneity) $\lambda(\varphi + t\mathbf{1}) = \lambda(\varphi) + t$ *for any $t \in \mathbb{R}$, where* $\mathbf{1} = (1, 1, \dots, 1) \in C^m(X)$;

c) (convexity) $\lambda(t\varphi + (1-t)\psi) \le t\lambda(\varphi) + (1-t)\lambda(\psi)$ *for all $\varphi, \psi \in C(X)$ and $t \in [0, 1]$;*

d) (Lipschitz property) $|\lambda(\varphi) - \lambda(\psi)| \le \|\varphi - \psi\|$ *for all $\varphi, \psi \in C(X)$.*

Note that in contrast to the properties of spectral potentials of transfer operators (see Proposition 5.5) here we do not have the invariance (and do not have a singled out mapping).

As in Section 5 convexity and continuity of the spectral potential in question implies that the spectral potential can be represented by means of the Legendre dual functional λ^*, defined on the dual space (recall (5.5), (5.6)), thus here we also have the corresponding variational principle for the spectral potential. From monotonicity and additive homogeneity it follows that the effective domain implies that the effective domain of λ^* belongs to the set of positive normalized functionals (that is probability measure on Y). The authors of [**18, 19**] exploited the dual functional (not the dual entropy, as we did in the previous sections) and therefore when quoting their results we will also make use of this object.

THEOREM 10.3 (Variational principle for a multiterm functional operator, [**19**]). *Under the assumptions mentioned above for each of the spaces $C(X)$ and $L^p(X, m)$ there exists a convex lower semicontinuous functional $\mathcal{T}$ on $M(Y)$, such that for the spectral potential* (10.4) *in the space under consideration the following variational principle holds*

$$\lambda(\varphi) = \ln r(W) = \max_{\mu \in M(Y)} \left\{ \int_Y \varphi d\mu - \mathcal{T}(\mu) \right\}. \tag{10.5}$$

The effective domain $D(\mathcal{T})$ of the functional $\mathcal{T}$ belongs to the set $M(Y)$ of probability measures on Y.

The functionals $\mathcal{T}(\mu)$ on the dual space that stay in the right hand side of the corresponding variational principles define certain *new dynamical characteristics of the set of mappings α_k*. Thus one arrives at the problem of investigation of these characteristics, and in particular one has to analyse the following questions:

1) what measures belong to the effective domain of the functional $\mathcal{T}(\mu)$ (and, in particular, influence the operator spectral properties);

2) what is the direct procedure of calculation of these functionals on the given measures (by means of the mappings α_k);

3) what is the dynamical meaning of $\mathcal{T}(\mu)$?

REMARK 10.4. For t-entropy the corresponding questions have complete answers, recall in this connection Proposition 6.3, Definition 7.1 and Theorem 8.6.

Symbolic sequences. As the next step of investigation let us consider a reduction of the problem to the analysis of the so-called symbolic sequences.

To start with we introduce the necessary notation. Let G be a group or a semigroup, $\Omega = \{g_1, \dots, g_s\}$ be a finite subset of G and $|\Omega|$ is the number of elements of Ω.

We denote by Ω_n the set of elements that can be represented in the form of a product of n elements belonging to Ω:

$$\Omega_n = \{g_{i_1} g_{i_2} \dots g_{i_n} : i_j \in \Delta\}.$$

Some of these products may coincide and the number $|\Omega_n|$ depends on relations between the elements of Ω. For $n > 2$ as a rule one has $|\Omega_n| < |\Omega|^n$, and the equality $|\Omega_n| = |\Omega|^n$ takes place only in the case when the elements of Ω are generators of a free group.

As in the case of weighted shift and transfer operators the investigation is based on the usage of the known Gelfand–Beurling formula

$$r(W) = \lim_{n\to\infty} \|W^n\|^{1/n},$$

and the constructions presented below arise in the analysis of the expressions for operators W^n.

Since $T_\alpha a = (a \circ \alpha) T_\alpha$ and $T_{\alpha_k} T_{\alpha_l} = T_{\alpha_k \circ \alpha_l}$, it follows that

$$W^n = \sum_{\xi \in \Delta^n} a_\xi(x) T_{\alpha^\xi}, \tag{10.6}$$

where

$$\Delta^n = \{\xi = (\xi_1, \xi_2, \dots, \xi_n) : \xi_i \in \Delta\},$$

$$a_\xi(x) = a_{\xi_1}(x)\, a_{\xi_2}(\alpha_{\xi_1}(x)) \cdots a_{\xi_n}(\alpha_{\xi_1} \circ \cdots \circ \alpha_{\xi_{n-1}}(x)),$$

$$\alpha^\xi(x) = \alpha_{\xi_1} \circ \cdots \circ \alpha_{\xi_n}(x).$$

Note that for different $\xi \in \Delta^n$ the mappings α^ξ (and the corresponding operators T_{α^ξ}) may coincide. Let G be the semigroup of the mappings generated by the set $\Omega = \{\alpha_1, \alpha_2, \dots, \alpha_s\}$. Then the set of (different) operators T_{α^ξ} entering expression (10.6) is parametrized by the elements of Ω_n, and the number different elements here is equal to $|\Omega_n|$. Therefore the expression for W^n can be rewritten in the form

$$W^n = \sum_{d \in \Omega_n} a_d\, T_d, \quad \text{where } \; a_d(x) = \sum_{\alpha^\xi = d} a_\xi(x). \tag{10.7}$$

Note that the coefficients a_d here depend on n, and in the notation exploited this dependence appear in the condition $d \in \Omega_n$.

REMARK 10.5. Here we have the essential difference between the case of a single weighted shift operator and a multiterm functional operator. Namely, if $Wu(x) = a(x)u(\alpha(x))$, then the operator W^n has the expression containing the single summand:

$$W^n u(x) = \left[\prod_{i=0}^{n-1} a\big(\alpha^i(x)\big)\right] u(\alpha^n(x)),$$

where $\alpha^{i+1}(x) = \alpha(\alpha^i(x))$. Moreover, in each of the spaces under consideration one has the explicit expression for the norm $\|W^n\|$. In contrast to this in the case of a multiterm functional operator the investigation complicates since the number of summands in the expression for W^n grows with the increase of n, the coefficients for the corresponding summands have the form of the sum of products, and in addition in the spaces $L^p(X, m)$ there is no simple explicit expression for the norm of such operator.

By the *symbolic sequence* for the operator W we mean the sequence of functions

$$S_n(x) = \sum_{\xi \in \Delta^n} a_\xi(x),$$

where the functions $a_\xi(x)$ are the coefficients of the operator W^n given by formula (10.6).

THEOREM 10.6 ([**19**]). *Let $\alpha_k : X \to X$ be continuous mappings on a compact space X. Then for the operator* (10.6) *in the space $C(X)$ with nonnegative continuous coefficients a_k the following equality holds*

$$r(W) = \lim_{n\to\infty} \|S_n\|^{1/n}, \tag{10.8}$$

where

$$\|S_n\| = \max_x |S_n(x)|$$

is the norm of the function S_n in the space $C(X)$.

An analogous statement in $L^p(X, m)$ takes place only under additional assumptions.

The case of invertible mappings. Let us discuss in more detail the situation when G is a group of mappings, generated by given invertible mappings α_k. Recall that the set $\Omega \subset G$ is called *subexponential*, if $\lim |\Omega_n|^{1/n} = 1$.

For example, any mutually commuting family of mappings is subexponential, since in this case one has $|\Omega_n| \le n^{|\Omega|}$.

THEOREM 10.7 ([**19**]). *Let (X, m) be a space with a measure m and α_k, $k = 1, \dots, s$ be continuous invertible mappings on the space X, the measure μ be quasiinvariant with respect to each of these mappings and the functional operator in the space $L^p(X, m)$ be given in the form $W = \sum_{k=1}^{s} a_k T_{\alpha_k}$.*

If the set $\Omega = \{\alpha_1, \alpha_2, \dots, \alpha_s\}$ is subexponential, the coefficients a_k belong to the space $C(X)$ and are nonnegative, then

$$r(W) = \lim_n \max_{d\in\Omega_n} \|a_d\|^{1/n} = \lim_n \|S_n\|^{1/n} = \lim_n \left(\sum_{d\in\Omega_n} \|a_d\| \right)^{1/n}.$$

As a corollary one obtains

PROPOSITION 10.8. *If the coefficients a_k are constant then*

$$S_n = \left(\sum_{k=1}^{s} a_k \right)^n$$

and under conditions of Theorems 10.6 *and* 10.7 *one has*

$$r(W) = \sum_{k=1}^{s} a_k. \tag{10.9}$$

REMARK 10.9. A certain additional condition on the group G in Theorem 10.7 is necessary. In particular, equality (10.9) for the operators with constant coefficients in the case of arbitrary group G is not true, though it is not evident from the first sight. This follows from subtle results on symmetric wanderings on groups obtained by Kesten [**42, 43**] and Grigorchuk [**36**].

It is shown in [**42**] that equality (10.9) is true only for amenable groups, and in [**43**] the following assertion is proved.

THEOREM 10.10. *Let F_s be a discrete free group with generators $h_1, \dots, h_s$. Then in the space $l^2(F_s)$ for the functional operator* (*with equal constant coefficients*)

$$Wu(x) = \frac{1}{2s}\sum_{k=1}^{s}\big(u(h_k x) + u(h_k^{-1}x)\big), \qquad x \in G, \tag{10.10}$$

one has

$$r(W) = \frac{\sqrt{2s-1}}{s}.$$

In this example for $s > 1$ we have

$$r(W) = \frac{\sqrt{2s-1}}{s} < \sum a_k = 1.$$

In [**36**] there are considered discrete groups that are free products of cyclic groups, for example, the groups $G_{l,k}$ with two generators a and b satisfying the relations $a^l = b^k = e$. It is shown that if $W_{l,k}$ is the operator of the form (10.10) in $l^2(G_{l,k})$, then for $l, k \to \infty$ one has

$$r(W_{l,k}) \to \frac{\sqrt{3}}{2} = r(W) < 1$$

and therefore, for sufficiently large l and k, equality (10.9) is not true.

In connection with Theorem 10.3 and the results of [**42, 43, 36**] one naturally arrives at the next problem. Let G be a discrete group with generators $h_1, \dots, h_s$. Let us consider the functional operators in the space $l^2(G)$ of the form

$$Wu(x) = \sum_{k=1}^{s} e^{\varphi_k}\big(u(h_k x) + u(h_k^{-1}x)\big), \qquad x \in G$$

with (different) constant coefficients $a_k = e^{\varphi_k}$.

For these operators the spectral potential $\lambda(\varphi)$ is a continuous convex function on $\mathbb{R}^s$, and there exists a convex functional è $\mathcal{T}_G(t)$, defined on the simplex

$$\mathcal{S} := \Big\{t \in \mathbb{R}^s : t_k \geq 0,\ \sum t_k = 1\Big\},$$

such that

$$\ln r(W) = \lambda(\varphi) = \max_{t\in\mathcal{S}}\Big\{\sum \varphi_k t_k - \mathcal{T}_G(t)\Big\}.$$

For any amenable, and, in particular, subexponential group G we have

$$\lambda(\varphi) = \ln r(W) = 2\sum e^{\varphi_k}$$

and
$$\lambda^*(t) = \mathcal{T}_G(t) = \sum t_k \ln t_k - \ln 2.$$
But for an arbitrary group G the latter equalities are not true, thus one arrives at the problem of computation of the functional $\mathcal{T}_G(t)$ for arbitrary groups, and, in particular, for a free group.

Remark 10.11. Equality (10.8) in Theorem 10.6 is in fact a record of the Gelfand–Beurling formula since $\|W^n\| = \|S_n\|$. In the case of the spaces $L^p(X, m)$ there is no explicit formula for $\|W^n\|$ and Theorem 10.7 reduces the question on spectral radius to consideration of the explicitly given value $\|S_n\|$. For example, the operator of the form (10.10) is selfadjoint and therefore in this case $r(W) = \|W\|$. Therefore Kesten's result can be considered as an example of a (rather nonevident) norm computation for a multiterm functional operator (with constant coefficients).

The symbolic sequences for the first time were exploited in investigation of functional operators in the work by Brenner [**28**] (see also in this connection [**15**, Sections 22, 23], where the corresponding apparatus is developed in a general C^*-algebraic situation, that is for multiterm operators with operator-valued coefficients; in connection with the norm calculation for multiterm functional operators in the spaces $L^p(X, m)$ one should also mention [**58**], where in addition the structure of the Banach algebras generated by such operators is described).

The measure typicalness exponent. In [**19**] there is obtained a description of the dual functional independent on the spectral potential which can serve as a certain 'multimapping' analogue to t-entropy.

A point $x \in X$ and an element $\xi \in \Delta^n$ define the trajectory of the length n, namely, the sequence of points of Y: (x, ξ_1), (x_1, ξ_2), ..., (x_{n-1}, ξ_n), where $x_i = \alpha_{\xi_1} \circ \cdots \circ \alpha_{\xi_i}(x)$. This trajectory can be considered as a trajectory of a certain random wandering on the set Y. Let $\delta(x, k)$ be the Dirac measure of the point
$$(x, k) \in Y = X \times \Delta.$$
An element $\xi \in \Delta^n$ and a point $x \in X$ define the so-called *empiric measure* on Y concentrated on the set of points of the trajectory:
$$\mu_{\xi,x} := \frac{1}{n}\big(\delta(x, \xi_1) + \delta(x_1, \xi_2) + \cdots + \delta(x_{n-1}, \xi_n)\big). \tag{10.11}$$
This measure naturally appears in the analysis of the expression for coefficients of W^n. Namely, if all the coefficients of operator W are equal $a_1(x) = \cdots = a_s(x) = e^\varphi$, then:
$$\ln a_\xi(x) = n \int_Y \varphi \, d\mu_{\xi,x}.$$
Some of the points of the trajectory may coincide, thus the empiric measure of a given point of the trajectory is equal to k/n, where k is the number of points of the trajectory, coinciding with the given one. This means that the empiric measure of a point is the 'relative time of the trajectory staying' in the given point.

Let μ be an arbitrary measure on Y and W be a neighborhood of it in *-weak topology. Denote by $N(n, W, x)$ the number of elements $\xi \in \Delta^n$ such that $\mu_{\xi,x} \in W$. Let us define the functional $T(\mu)$ on the set of measures on Y by the expression
$$T(\mu) := \inf_{W \in O(\mu)} \limsup_{n\to\infty} \frac{1}{n} \max_x \ln N(n, W, x), \tag{10.12}$$

where $O(\mu)$ is the set of all the neighborhoods of the measure μ. The number $T(\mu)$ will be called the *measure μ typicalness exponent* with respect to the set of mappings $\alpha_k,\ \ k \in \Delta$. In particular, if μ possesses a neighborhood that does not contain empiric measures, then $T(\mu) = -\infty$ and this measure does not belong to the effective domain of the functional $T(\,\cdot\,)$.

The quantity $T(\mu)$ characterizes the number of elements $\xi \in \Delta^n$, generating empiric measures, that are close to the given one; by means of $T(\mu)$ one can describe the behavior of the values $S_n(x)$ of a symbolic sequence. The next two theorems can be naturally considered as 'general mappings' counterparts of Theorem 10.10.

THEOREM 10.12 ([**19**]). *Let X be a compact space, $\alpha_1, \ldots, \alpha_s$ be arbitrary continuous mappings on X, $\varphi = (\varphi, \ldots, \varphi)$ be a continuous vector-valued function on X, $Y = X \times \Delta$ and $T(\mu)$ be the functional on the set $M(Y)$ of probability measures on Y given by formula* (10.12).

Then the spectral potential of the operator

$$W = \sum_{k=1}^{s} e^{\varphi} T_{\alpha_k} \tag{10.13}$$

in the space $C(X)$ satisfies the following variational principle

$$\lambda(\varphi) = \ln r(W) = \max_{\mu \in M(Y)} \left\{ \int_Y \varphi d\mu + T(\mu) \right\}.$$

THEOREM 10.13 ([**19**]). *Let X be a compact space, $\alpha_1, \ldots, \alpha_s$ be continuous invertible mappings of X, m be a measure on X which is quasiinvariant with respect to each of these mappings and $\varphi = (\varphi, \ldots, \varphi)$ be continuous vector-valued function on X.*

If the set $\Omega = \{\alpha_1, \ldots, \alpha_s\}$ is subexponential, then for the spectral potential of operator (10.13) *in the space $L^p(X, m)$ the following variational principle holds true*

$$\lambda(\varphi) = \ln r(W) = \max_{\mu \in M(Y)} \left\{ \int_Y \varphi d\mu + T(\mu) \right\}.$$

An explicit form of the functional $T_M(\mu)$ can be obtained only for special cases (see [**19**]).

EXAMPLE 10.14. Let us consider in the space $L_2(\mathbb{R}^l)$ the operators of the form

$$Wu(x) = \sum_{k=1}^{s} a_k(x) u(x + h_k), \tag{10.14}$$

where $h_k \in \mathbb{R}^l$ and each of the coefficients a_k is a continuous nonnegative function, having limit at infinity $a_k(\infty)$. Here as X we consider the space $\mathbb{R}^l$, compactificated by a single point ∞.

THEOREM 10.15 ([**19**]). *If there exists a linear function f on $\mathbb{R}^l$, such that $f(h_k) > 0$ for all k, then the typicalness functional T takes finite values only on the set of probability measures comcentrated on the set $\infty \times \Delta \subset X \times \Delta$, and its value on the measure $\mu = \sum_{k=1}^{s} t_k \delta(\infty, k)$, where $t_k \geq 0$ and $\sum_{k=1}^{s} t_k = 1$, is given by the formula*

$$T(\mu) = \sum_{k=1}^{m} t_k \ln t_k,$$

and for the operator (10.14) *one has*

$$r(W) = \sum_{k=1}^{s} a_k(\infty). \tag{10.15}$$

In particular, on $\mathbb{R}$ ($l = 1$) the condition of the theorem means that the result takes place in the case when all h_k are positive, or all h_k are negative. Even in the case when a single number h_k is zero and the rest are positive formula (10.15) is not true, but in this case one can also derive the explicit formula. But if among the numbers h_k there are numbers of different signs, then the problem complicates and there is no way to obtain an explicit formula for $T(\mu)$.

Irreversible mappings. Starters. In the case of multiterm functional operators, generated by a set of irreversible mappings an independent construction of the dual functional is not obtained yet. This construction may turn out to be rather complicated, since it has to involve, as a particular case the t-entropy constructions for each of the mappings α_k and a combinatorial component, analogous to functional (10.12).

In this direction there are results only for a number of special cases. In [**64, 65, 66**] there is considered the set of operators of the form

$$W = w(e^{\varphi}T_\alpha), \qquad \varphi \in C(X),$$

where

$$w(z) = \sum_{k=0}^{s} a_k z^k, \qquad a_k > 0,$$

that is operators that are polynomials of weighted shift operators, generated by a given mapping.

The spectral radius of such operators can be found by means of a general reasoning related to the positivity of the operator in question:

$$r\bigl(w(e^{\varphi}T_\alpha)\bigr) = w\bigl(r(e^{\varphi}T_\alpha)\bigr).$$

By using the notation $\boldsymbol{c} := (c_0, \dots, c_s)$, $c_k = \ln a_k$ one can consider the spectral potential of the operator W as a functional on the space $C(X) \times \mathbb{R}^{s+1}$:

$$\lambda(\varphi, \boldsymbol{c}) := \ln r\bigl(w(e^{\varphi}T_\alpha)\bigr). \tag{10.16}$$

In view of the foregoing discussion it is reasonable to analyze the Legendre dual functional to functional (10.16). Namely, the explicit form of this functional is the principal result of [**64, 65, 66**] which is stated in the next

THEOREM 10.16. *For the functional* $\lambda(\varphi, \boldsymbol{c})$ *the following variational principle holds*

$$\lambda(\varphi, \boldsymbol{c}) = \max_{(\bar\nu, \boldsymbol{t}) \in \mathcal{M}} \left\{ \int_X \varphi \, d\bar\nu + \sum_{k=0}^{s} c_k t_k - \widetilde{\lambda}^*(\bar\nu, \boldsymbol{t}) \right\},$$

where

$$\widetilde{\lambda}^*(\bar\nu, \boldsymbol{t}) = \frac{1}{p}\, \bar\nu(X) \tau_\alpha\left(\frac{\bar\nu}{\bar\nu(X)}\right) + \sum_{k=0}^{s} t_k \ln t_k$$

and

$$\mathcal{M} = \bigl\{ (\bar\nu, \boldsymbol{t}) : \boldsymbol{t} = (t_k)_{k=0}^{s} \in \mathbb{R}^{s+1},\ t_k \ge 0,\ \textstyle\sum_{k=0}^{s} t_k = 1,$$
$$\text{and } \bar\nu \text{ is } \alpha\text{-invariant and such that } \bar\nu(X) = \textstyle\sum_{k=0}^{s} k t_k \bigr\}.$$

When $\bar{\nu}(X) = 0$ *then* $\widetilde{\lambda}^*$ *vanishes.*

In [**67**] there are considered the operators that are analytic functions of weighted shift operators and here the explicit form of the dual functional is obtained as well.

Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$ be an analytic function with nonnegative coefficients a_n for which the radius of convergence is greater than $r(e^{\varphi}T_\alpha)$. For this reason the operator

$$f(e^{\varphi}T_\alpha) := \sum_{n=0}^{\infty} a_n (e^{\varphi}T_\alpha)^n$$

is well defined and bounded.

THEOREM 10.17. *For the functional* $\widetilde{\lambda} = \ln r(f(e^{\varphi}T_\alpha))$ *the following variational principle holds*

$$\widetilde{\lambda}(\varphi) = \sup_{\mu \in M_\alpha^l} \bigl\{ \langle \mu, \varphi \rangle - \widetilde{\lambda}^*(\mu) \bigr\},$$

where $M_\alpha^l = \{\mu \in C(X)^* : \mu \in M_\alpha \text{ and } \mu(X) \in [l, +\infty)\}$,

$$\widetilde{\lambda}^*(\mu) = \frac{1}{p}\, \mu(X) \tau_\alpha\Bigl(\frac{\mu}{\mu(X)}\Bigr) + \min_{(t_k) \in S_{\mu(X)}} \liminf_{N \to \infty} \sum_{k \in I,\, k \le N} t_k \ln \frac{t_k}{a_k} \qquad \text{for } \mu(X) \neq 0$$

and

$$S_{\mu(X)} = \bigl\{ (t_k)_{k \in I} \in S : \textstyle\sum_{k \in I} k t_k < +\infty \text{ and } \mu(X) = \sum_{k \in I} k t_k \bigr\};$$

if $\mu(X) = 0$ *then* $\widetilde{\lambda}^*(0) = -\ln a_0$.

It is interesting to note that for the function $f(z) = e^z$ and the operator

$$W = e^{\varphi}T_\alpha$$

we obtain

$$\ln r(e^W) = r(W)$$

and the theorem brings us to the variational principle for the spectral radius (but not the spectral potential):

$$r(e^{\varphi}T_\alpha) = \max_{\mu \in \boldsymbol{M}_\alpha} \left\{ \int_X \varphi\, d\mu - \frac{\mu(X)}{p}\, \tau_\alpha\Bigl(\frac{\mu}{\mu(X)}\Bigr) - \mu(X) \ln \mu(X) + \mu(X) \right\},$$

where the maximum is taken over the set $\boldsymbol{M}_\alpha$ of invariant (not normalized) measures.

In particular, in the case of invertible mapping we have

$$r(W) = \max_{\mu \in \boldsymbol{M}_\alpha} \left\{ \int_X \varphi\, d\mu - \mu(X) \ln \mu(X) + \mu(X) \right\}.$$

References

1. A. B. Antonevich, *On a class of pseudodifferential opearators with deviating argument on the torus.* Differentsial'nye Uravneniya **11**(9) (1975), 1550–1557 (in Russian). MR0387851 (52:8689)
2. A. B. Antonevich, *Operators with a shift generated by the action of a compact Lie group.* Sibirsk. Mat. Zh. **20**(3) (1979), 467–478 (in Russian). MR537353 (81h:47011)
3. A. B. Antonevich, *Invertibility conditions for operators with convex rationally independent system of shifts.* Doklady AN SSSR **256**(1) (1981), 11 - 14 (in Russian). MR599090 (82h:47030)
4. A. B. Antonevich, *Linear functional equations. Operator approach.* Universitetskoe, Minsk, 1988 (in Russian); English transl. Birkhauser Verlag, 1996. MR1382652 (97i:47001)

5. A. B. Antonevich, V. I. Bakhtin, A. V. Lebedev, *Variational principle for the spectral radius of weighted composition and Perron–Frobenius operators.* Trudy Instituta Matematiki NAN Belarusi **5** (2000), 13–17 (in Russian).
6. A. B. Antonevich, V. I. Bakhtin, A. V. Lebedev, *Variational principle for the spectral radius of weighted composition and weighted mathematical expectation operators.* Doklady NAN Belarusi **44**(6) (2000), 7–10 (in Russian). MR1851869 (2002f:47067)
7. A. B. Antonevich, V. I. Bakhtin, A. V. Lebedev, *Thermodynamics and Spectral Radius.* Nonlinear Phenomena in Complex Systems **4**(4) (2001), 318–321. MR1950124 (2003m:47040)
8. A. B. Antonevich, V. I. Bakhtin, A. V. Lebedev, *Spectra of Operators Associated with Dynamical Systems: From Ergodicity to the Duality Principle.* Twenty Years of Bialowieza: A mathematical Anthology Aspects of Differential Geometric Methods in Physiscs. World Scientific Monograph Series in Mathematics. V. 8, Chapter 7, 129–161. MR2181550 (2006f:47034)
9. A. B. Antonevich, V. I. Bakhtin, A. V. Lebedev, *Crossed product of a C^*-algebra by an endomorphism, coefficient algebras and transfer operators.* arXiv:math. OA/0502415 33pp.
10. A. B. Antonevich, V. I. Bakhtin, A. V. Lebedev, *On t-entropy and variational principle for the spectral radii of transfer and weighted shift operators.* Ergodic Theory Dynam. Systems **31** (2011), 995–1045. MR2818684
11. A. B. Antonevich, V. I. Bakhtin, A. V. Lebedev, D. S. Sarzhinsky, *Legendre analysis, thermodynamic formalizm and spectrums of Perron–Frobenius operators.* Doklady RAN **390**(3) (2003) 353–356 (in Russian).
12. A. Antonevich, M. Belousov, A. Lebedev, *Functional differential equations: II. C^*-applications: Part 1: Equations with continuous coefficients.* Addison Wesley Longman, Pitman momographs and surveys in pure and applied mathematics, **94**, 1998. MR1680633 (2000d:47073a)
13. A. Antonevich, M. Belousov, A. Lebedev, *Functional differential equations: II. C^*-applications: Part 2: equations with discontinuous coefficients and boundary value problems.* Addison Wesley Longman, Pitman momographs and surveys in pure and applied mathematics, **95**, 1998. MR1680637 (2000d:47073b)
14. A. B. Antonevich, A. V. Lebedev, *On spectral properties of operators with a shift.* Izv. AN SSSR., Ser. Mat. **47**(5) (1983), 915–941 (in Russian). MR718411 (85c:47028)
15. A. Antonevich, A. Lebedev, *Functional differential equations: I. C^* - theory.* Longman Scientific & Technical, 1994. MR1413495 (97e:46072)
16. A. B. Antonevich, V. B. Ryvkin, *The normal solvability of the problem of the periodic solution of a linear differential equation with deviating argument.* Differentsial'nye Uravneniya **10**(3) (1974), 1347–1353 (in Russian). MR0361362 (50:13807)
17. A. B. Antonevich, Yu. V. Vol'pert, *On the spectrum of a weighted diagonal shift operator on the torus.* Vestsi Akad. Navuk BSSR Ser. Fiz.-Mat. navuk (3) (1976), 48–53 (in Russian).
18. A. B. Antonevich, K. Zajkowski, *Variational principle for the spectral radius of a model functional operator.* Trudy Inst. Mat. NAN Belarusi Minsk, **12**(2) (2004), 18–25 (in Russian).
19. A. B. Antonevich, K. Zajkowski, *Variational principles for the spectral radius of functional operators.* Sbornik Mathematics **197**(5) (2006), 633–680. MR2264327 (2007j:47062)
20. Ch. Babbage, *An essay towards the calculus of functions.* Philos. Trans. London. **106**(2) (1816), 179–256.
21. V. I. Bakhtin, *Positive processes.* Ergodic Theory Dynam. Systems **27** (2007), 639–670. MR2322172 (2010c:37013)
22. V. I. Bakhtin, *On t-entropy and variational principle for the spectral radius of weighted shift operators.* Ergodic Theory Dynam. Systems **30** (2010), 1331–1342. MR2718896 (2012a:37058)
23. V. I. Bakhtin, A. V. Lebedev, *A new definition of t-entropy for transfer operators.* arXiv: 0906.4969v1 [math.DS] 26 Jun 2009.
24. V. Baladi, *Positive transfer operators and decay of correlations.* World Scientific Publ., River edge, NJ, 2000. MR1793194 (2001k:37035)
25. V. Baladi, M. Holschneider, *Approximation of nonessential spectrum of transfer operators.* Nonlinearity **12** (1999), 525–538. MR1690191 (2000h:47008)
26. V. Baladi, L.-S. Young, *On the spectra of randomly pertrubated expanding mapps.* Comm. Math. phys. **156** (1993), 355–385. MR1233850 (94g:58172)
27. P. Billingsley, *Ergodic theory and information.* John Wiley and Sons, New York–London–Sydney, 1965. MR0192027 (33:254)

28. V. V. Brenner, *Operators with locally independent shifts. Diss. kand. fiz.-mat. nauk.* Minsk, 1981 (in Russian).
29. J. Campbell, Yu. Latushkin, *Sharp estimates in Ruelle theorems for matrix transfer operators.* Comm. Math. phys. **185** (1996), 379–396. MR1463047 (98j:58088)
30. P. Collet, S. Isola, *On the essential spectrum of the transfer operator for expanding Markov maps.* Comm. Math. phys. **139** (1991), 551–557. MR1121133 (92h:58157)
31. V. Didenko, *Estimates of the spectral radius of refinement and subdivision operators with isotropic dilations.* J. of Oper. Theory **55** (2007). MR2385440 (2009m:47007)
32. I. Ekland, R. Temam, *Convex analysis and variational problems.* North-Holland, Amsterdam, 1976.
33. R. Exel, *A new look at the crossed-product of a C^*-algebra by an endomorphism.* Ergodic Theory Dynam. Systems **23**(6) (2003), 1733–1750. MR2032486 (2004k:46119)
34. R. Exel, A. Vershik, *C^*-algebras of irreversible dynamical systems.* Canad. J. Math. **58**(1) (2006), 39–63. MR2195591 (2007d:46058)
35. D. Fried, *The flat-trace asymptotics of a uniform system of contructions.* Ergodic Theory Dynam. Systems **15** (1995), 1061–1073. MR1366308 (97c:58124)
36. R. I. Grigorchuk, *Symmetric random wondering on discrete groups.* Multicomponent random systems. M.: Nauka (1978), 132–152 (in Russian).
37. V. M. Gundlach, Yu. Latushkin, *Essential spectral radius of Ruelle's operator on smooth Hölder spaces.* Comp. rend. Acad. sci. Paris, Serie I **325**, 197 889-894. MR1485911 (98k:58178)
38. V. M. Gundlach, Yu. Latushkin, *A sharp formula for the essential spectral radius of the Ruelle transfer operator on smooth and holder spaces.* Ergodic Theory Dynam. Systems **23** (2003), 175–191. MR1971201 (2004e:37035)
39. N. K. Karapetyants, *On a class of discrete convolution operators with oscillating coefficients.* Doklady AN SSSR **216**(1) (1974) 28–31 (in Russian). MR0344936 (49:9675)
40. N. K. Karapetyants, S. G. Samko, *On functional equation $\psi(x + \alpha) - b(x)\psi(x) = g(x)$.* Izv. AN Arm. SSR **5**(5) (1971), 441–445 (in Russian). MR0284877 (44:2101)
41. Yu. I. Karlovich, V. G. Kravchenko, *The Noether theory of a singular integral operators with a shift having periodic points.* Doklady AN SSSR **231** (1976) 277–280 (in Russian). MR0428114 (55:1143)
42. H. Kesten, *Fill Banach mean values on countable groups.* Math. scand. **7**(1), (1959). MR0112053 (22:2911)
43. H. Kesten, *Symmetric random walks on groups.* Trans. Amer. Math. Soc. **92**, (1959), 336–354. MR0109367 (22:253)
44. A. Kitaev, *Fredholm determinants for hyperbolic diffeomorphisms of finite smoothness.* Nonlinearity **12** (1999), 141–179. MR1668543 (2000a:37015)
45. A. Kitover, *On the spectrum of an automorphism with a weight and Kamowitz-Sheinberg theorem.* Funktsion. Anal. i Prilozhen. **13**(1) (1979), 70–71 (in Russian). MR527528 (80f:46053)
46. A. Knauf and Ya. Sinai, *Classical Noninegrability, Quantum chaos.* Birkhauser-Verlag, Basel, 1997. MR1461015 (98k:58181)
47. V. G. Kravchenko, *On singular integral operator with a shift.* Doklady AN SSSR **215**(6) (1974) 1301–1304 (in Russian). MR0348559 (50:1057)
48. B. K. Kwaśniewski, *C^*-algebras associated with reversible extensions of logistic maps.* arXiv: 1007.5395v1 [math.DS] 30 Jul 2010.
49. B. K. Kwaśniewski, *Spectral analysis of operators generating irreversible dynamical systems. Ph. D. thesis.* Warsaw, PAN, 2009 (in Polish).
50. B. K. Kwaśniewski, A. V. Lebedev, *Reversible extensions of irreversible dynamical systems: C^*-method.* Sbornik Mathematics **199**(11) (2008), 45–74. MR2485375 (2009k:46127)
51. Ju. D. Latushkin, *On the integral functional operators with a nonbijective shift.* Izv. AN SSSR. Ser. Mat. **45**(6) (1981), 1241–1257 (in Russian). MR641801 (83d:47051)
52. Ju. D. Latushkin, A. M. Stepin, *Weighted shift operators on a topological Markov chain.* Funktsion. Anal. i Prilozhen. **22**(4) (1988), 86–87 (in Russian). MR977005 (90d:47034)
53. Ju. D. Latushkin, A. M. Stepin, *Weighted shift operators, spectral theory of linear extentions and multiplicative ergodic theorem.* Matem. Sbornik **181** (6) (1990), 724–742 (in Russian).
54. Ju. D. Latushkin, A. M. Stepin, *Weighted shift operators and linear extensions of dynamical systems.* Uspekhi. Mat. Nauk **46**(2) (1991), 85–143 (in Russian). MR1125273 (92k:47062)

55. A. V. Lebedev, *On the invertibility of elements in C^*-algebras generated by dynamical systems.* Uspekhi Mat. Nauk **34**(4) (1979), 199–200 (in Russian). MR548434 (81g:47041)
56. A. V. Lebedev, *On the spectrum of weighted shift operators generated by Morse–Smale diffeomorphism.* Vestsi Akad. Navuk BSSR Ser. Fiz.-Mat. navuk (3) (1979), 24–28 (in Russian). MR544294 (81i:47033)
57. A. V. Lebedev, *On the spectrum of intransitive weighted shift operators on the circle.* Voprosy funkts. analiza., Petrozavodsk (1980), 8–18 (in Russian).
58. A. V. Lebedev. *On the structure of Banach algebras associated with automorphisms.* Trans. MMS. **66** (2005), 114–155. MR2193431 (2007e:46043)
59. A. Lebedev, O. Maslak, *The variational principles for the spectral characteristics of the operators generated by dynamical systems.* Proceedings of the Fifth Annual Seminar NPCS'96. Minsk, 1997, 165–170.
60. A. Lebedev, O. Maslak, *The spectral radius of a weighted shift operator, variational principles, entropy and topological pressure. Spectral and evolutionary problems.* Proceedings of the Eight Crimean Autumn Mathematical School-Symposium. Simferopol, 1998, 26–34. MR1725042 (2000h:47046)
61. G. S. Litvinchuk, *Boundary value problems and singular integral equations with a shift.* M. Nauka, 1977 (in Russian). MR0587791 (58:28532)
62. S. Lo, *Weighted shift operators in certain Banach spaces of functions. Diss. kand. fiz.-mat. nauk.* Minsk, 1981 (in Russian).
63. Ae. Mukhamadiev, B. N. Sadovskij, *On a spectral radius estimate for a certain operator related to neutral type equations.* Mat. zametki **13**(1) (1973), 61–78 (in Russian).
64. U. Ostaszewska, K. Zajkowski, *Spectral exponent of finite sums of weighted positive operators in Lp-spaces.* Positivity **11** (4) (2007), 549–562. MR2346441 (2008h:47068)
65. U. Ostaszewska, K. Zajkowski, *Legendre-Fenchel transform of the spectral exponent of polynomials of weighted composition operators.* Positivity **14**(3) (2010), 373–381. MR2680501
66. U. Ostaszewska, K. Zajkowski, *Variational principle for the spectral exponent of polynomials of weighted composition operators.* J. Math. Anal. Appl. **361** (2010), 246–251. MR2567298 (2010i:47041)
67. U. Ostaszewska, K. Zajkowski, *Legendre-Fenchel transform of the spectral exponent of analytic functions of weighted composition operators.* Journal of Convex Analysis **18**(2) (2011), 367–377. MR2828492
68. S. Parrot, *Weighted translation operators. Ph. D. diss.* Univ. of Michigan, Ann Arbor. Mich., 1965. MR2614859
69. W. Parry, M. Pollicot, *Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics.* Societe Methematique de france, Asterisque **187-188**. Paris, 1990. MR1085356 (92f:58141)
70. H. Rough, *Generalized fredholm determinants and Selberg zeta functions for Axiom A dynamical systems.* Ergodic Theory Dynam. Systems **16** (1996), 805–819. MR1406435 (97j:58125)
71. D. Ruelle, *Statistical mechanics on a compact set with Z^ν action satisfying expansiveness and specification.* Trans. Amer. Math. Soc. **185** (1973), 237–252. MR0417391 (54:5441)
72. D. Ruelle, *Thermodynamic Formalism.* Addison-Wesley, Reading. Mass., 1978. MR511655 (80g:82017)
73. D. Ruelle, *Characteristics exponents and invariant manifolds in Hilbert space.* Ann. Math. **115** (1982), 243-290. MR647807 (83j:58097)
74. D. Ruelle, *Dynamical zeta Functions for piecewise Monotone maps of the Interval.* CRM Monograph Series **4,** AMS, 1991. MR1274046 (95m:58101)
75. P. Semrl, *Nonlinear pertrubatins of homomorphisms on $C(X)$.* Quart. J. Math. Oxford Ser. 2 **50**(197) (1999), 87–109. MR1673258 (2000a:46074)
76. A. M. Stepin, *Homological equation in dynamical system theory.* Studies in the theory of functions of several real variables, Yaroslavl's state University, Yaroslavl' (1982), 106-117 (in Russian)
77. P. Walters, *A variational principle for the pressure on continuous transformations.* Amer. J. Math. **97**(4) (1975), 937–971. MR0390180 (52:11006)

University of Bialystok, Poland
Current address: Belarus State University, Belarus
E-mail address: `antonevich@bsu.by`

Belarus State University, Belarus
E-mail address: `bakhtin@tut.by`

University of Bialystok, Poland
Current address: Belarus State University, Belarus
E-mail address: `lebedev@bsu.by`

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11251

A condition for weak mixing of induced IETs

Michael Boshernitzan

Dedicated to A. M. Stepin on the occasion of his 70th birthday.

Abstract. Let $f\colon X \to X$, $X = [0, 1)$, be an ergodic IET (interval exchange transformation) relative to the Lebesgue measure on X. Denote by $f_t\colon X_t \to X_t$ the IET obtained by inducing f to the subinterval $X_t = [0, t)$, $0 < t < 1$. We show that

$$X_{wm} = \{0 < t < 1 \mid f_t \text{ is weakly mixing}\}$$

is a residual subset of X of full Lebesgue measure. The result is proved by establishing a generic Diophantine sufficient condition on t for f_t to be weakly mixing.

1. IETs: Minimality, ergodicity and mixing

Denote by λ the Lebesgue measure on the real line $\mathbb{R}$. Denote by $\mathbb{Z}$, $\mathbb{N} = \{k \in \mathbb{Z} \mid k \geq 1\}$ the sets of integers and of natural numbers, respectively. We write $\sharp(S)$ for the cardinality of a set S.

An IET (interval exchange transformation) is a pair (X, f) where $X = [0, b) \subset \mathbb{R}$ is a bounded interval, and f is a right continuous bijection $f\colon X \to X$ of it with a finite set D of discontinuities and such that $f'(x) = 1$ for all $x \in X\backslash D$. (The last conditions means that f is a translation in a neighborhood of every its continuity point). We often refer to the map f itself as an IET.

Let $\sharp(D) = r - 1$, $r \geq 2$. Then without loss of generality

$$D = \{d_k\}_{k=1}^{r-1}; \quad d_0 = 0 < d_1 < \ldots < d_{r-1} < b = d_r. \tag{1.1}$$

(The conventions $d_0 = 0$, $d_r = b$ are used for convenience. Note that $0, b \notin D$).

An IET (X, f) (or f) with $\sharp(D) = r - 1$ is also called (more specifically) an r-IET referring to the fact that f exchanges the r intervals $X_k = [d_{k-1}, d_k) \subset X$, $1 \leq k \leq r$, according to some permutation $\rho \in S_r$.

An r-IET is completely determined by this permutation $\rho \in S_r$ and the lengths $\lambda_k = d_k - d_{k-1} > 0$, $1 \leq k \leq r$, of exchanged subintervals X_k. Thus r-IETs can be identified with pairs $(\vec{\lambda}, \rho)$ where $\vec{\lambda} \in (\mathbb{R}^+)^r$ and $\rho \in S_r$: $(X, f) = (\vec{\lambda}, \rho)$.

2010 *Mathematics Subject Classification.* Primary 37A25, 37E05.

Key words and phrases. Interval exchange transformation, weak mixing, induced transformations.

The author was supported in part by NSF Grant: DMS-1102298.

A permutation $\rho \in S_r$ is called *irreducible* if $\rho(\{1, 2, \ldots, k\}) \neq (\{1, 2, \ldots, k\})$, for all $k < r$.

An IET (X, f) is called minimal if all f-orbits are dense in X. Note that the irreducibility of ρ is a necessary condition for f to be minimal because otherwise X splits into two f-invariant subintervals.

Keane [**16**] proved that if $r \geq 2$ and if $\rho \in S_r$ is irreducible then the IET $(\vec{\lambda}, \rho)$ is minimal provided that the lengths λ_k of exchanged intervals are linearly independent over the rationals (a generic assumption on $\vec{\lambda}$).

Masur [**21**] and Veech [**26**] independently proved the following result (conjectured by Keane in [**16**]): If $r \geq 2$ and $\rho \in S_r$ is irreducible, then for Lebesgue almost all $\vec{\lambda} \in (\mathbb{R}^+)^r$ the IET $(\vec{\lambda}, \rho)$ is uniquely ergodic (all its orbits are uniformly distributed). Alternative approaches to Keane's conjecture were later given by Rees [**22**], Kerckhoff [**20**] and Boshernitzan [**3**]. (Boshernitzan exhibited some Diophantine conditions, including a generic one, Property P, for unique ergodicity of IETs, see [**3**], [**5**] and [**28**], an improvement by Veech. The idea to use a suitable generic Diophantine condition to establish a metric result is also central in the present paper).

Avila and Forni [**1**] proved that Lebesgue almost all IETs are weakly mixing (assuming that ρ is irreducible and not a rotation). Partial results in this directions were obtained earlier by Katok and Stepin [**15**] who established weak mixing of generic 3-IETs and Veech [**27**] who proved generic weak mixing for a large class of permutations (Veech permutations). It was later shown by Boshernitzan and Nogueira [**8**] that if an IET $f = (\vec{\lambda}, \rho)$ satisfies property P (a generic condition used in [**3**]), or even a weaker condition [**28**], and if ρ is a Veech permutation, then f is weakly mixing.

Note that Katok proved that IETs are never (strongly) mixing [**13**]. On the other hand, Chaika [**9**] constructed a 4-IET which is topologically mixing. Boshernitzan and Chaika [**7**] showed that 3-IETs are never topologically mixing.

We refer the reader to a very nice book [**29**] (available on the web) by Marcelo Viana which may serve a nice introduction and survey reference to the subject of IETs.

2. The results

In what follows let (X, f) be a fixed aperiodic r-IET, $X = [0, b)$, $r \geq 2$. (Aperidicity of f means absence of f-periodic points).

Denote by X_t, $0 < t < b$, the subinterval $[0, t) \subset X$ and by f_t the IET obtained by inducing f to X_t. It is well known that each (X_t, f_t) is an s-IET with $s = s(t) \leq n + 1$ (see [**10**] or [**16**]). (In fact, $s(t) \geq 2$ due to the aperiodicity assumption).

The central result of the paper is the following. Recall that λ stands for the Lebesgue measure on X.

THEOREM 1. *Let (X, f) be a λ-ergodic IET, $X = [0, b)$. Then the set*

$$X_{\rm wm} = X_{\rm wm}(\lambda) = \{0 < t < b \mid f_t \text{ is weakly mixing (relative } \lambda)\} \tag{2.1a}$$

is a residual set of full measure: $\lambda(X_{\rm wm}) = 1$.

Because of the above result, we refer to the complement set

$$X_{\rm nwm} = (0, b) \backslash X_{\rm wm} \tag{2.1b}$$

as “the exceptional set for (weakly mixing induction of) (X, f)”.

Remark 1. Two additional versions of Theorem 1 (for minimal but not uniquely ergodic IETs) are given in Section 7.

We need some notation. Recall that $D = \{d_k\}_{k=1}^{r-1}$ stands for the set of discontinuities of f. Denote by

$$D_0 = D \cup \{0, b\} = \{d_k\}_{k=0}^{r}$$

the set of $(r + 1)$ points in (1.1), and denote by

$$D' = \bigcup_{k=-\infty}^{\infty} f^k(D) \tag{2.2}$$

the set of points whose orbits hit D.

The aperiodicity of f implies that D' is a dense countable subset of $X = [0, b)$ containing 0 (see e.g. [**5**, Section 2]).

For $x \in X$ and $n \geq 1$, we set

$$\rho(x) = \mathbf{dist}(x, D_0) = \min_{0\leq k\leq r} |x - d_k| \tag{2.3a}$$

$$\rho_n(x) = \min_{-n\leq k\leq n-1} \rho(f^k(x)) \tag{2.3b}$$

$$\Delta_n(x) = \min_{\substack{|p|,|q|\leq n\\ p\neq q}} |f^p(x) - f^q(x)| \tag{2.3c}$$

and

$$\rho'_n(x) = \min(\rho_n(x), \tfrac{1}{2}\Delta_n(x)). \tag{2.3d}$$

Note that for all $x\in D'$ both $\rho_n(x)$ and $\rho'_n(x)$ vanish for large n, while for $x\in X\backslash D'$ we have $\rho_n(x)\geq\rho'_n(x)>0$ due to the aperiodicity of f.

Important interpretations of the values $\rho_n(x)$ and $\rho'_n(x)$ are given by Propositions 1 and 2 in the next section. The following two functions

$$\phi\colon X \to \mathbb{R}; \qquad \phi(x) = \phi_f(x) = \limsup_{n\to\infty} n\rho_n(x); \tag{2.4a}$$

and

$$\psi\colon X \to \mathbb{R}; \qquad \psi(x) = \psi_f(x) = \limsup_{n\to\infty} n\rho'_n(x); \tag{2.4b}$$

will play central role in the paper. (We usually suppress the dependence on f by writing ϕ, ψ rather than ϕ_f, ψ_f due to the standing assumption that (X, f) is a fixed aperiodic IET).

Theorem 2 (A sufficient condition for weak mixing of f_t). *Let $f\colon X \to X$ be a λ-ergodic IET, $X = [0, b)$. If $\psi(t) > 0$ for some $t \in (0, b)$, then the induced IET $f_t\colon X_t \to X_t$, $X_t = [0, t)$, is weakly mixing.*

For an aperiodic IET $f\colon X \to X$, $X = [0, b)$, the set

$$X_{\psi=0} = \big\{t \in (0, b) \mid \psi(t) = 0\big\} \tag{2.5a}$$

will be referred as “the critical set for (X, f)”. We also adopt similar notation

$$X_{\psi>0} = (0, b)\backslash X_{\psi=0} = \big\{t \in (0, b) \mid \psi(t) > 0\big\} \tag{2.5b}$$

for the complement of this set.

Theorem 2 claims that for an λ-ergodic IET (X, f) the inclusion $X_{\mathrm{nwm}} \subset X_{\psi=0}$ takes place. (In other words, every exceptional point must be critical, see (2.1b) and (2.5b)).

Equivalently, $X_{\psi>0} \subset X_{\mathrm{wm}}$ (see (2.1a) and (2.5a)).

The importance of Theorem 2 is twofold. First, it provides a *generic* condition (Theorem 3) sufficient to establish Theorem 1 claiming that the exceptional set X_{nwm} is small in both measure and topology categories.

Secondly, it allows (under certain Diophantine conditions on t and f) to establish more delicate information on the "smallness" of the exceptional set X_{nwm}. In particular, there are examples of ergodic IETs (X, f) for which one can show that the critical set $X_{\psi=0}$ coincide with D', and hence $X_{\mathrm{nwm}} \subset D'$ is at most countable (see Section 8).

An IET (X, f) is called *persistently weakly mixing* if $X_{\mathrm{nwm}} = \emptyset$. It is easy to see that there are no persistently weakly mixing r-IETs with $r < 4$ (because then f_t become rotations for a countable set of t).

We believe that there exist persistently weakly mixing IETs.

Question. What is the "size" (in the sense of measure, category and cardinality) of X_{nwm} for "most" 4-IETs with permutation $\rho = (4321)$?

The answers for the same questions for r-IETs with $r = 2$ or 3 are known (see Section 8 for the answers without proofs).

Theorem 3. *Let $f\colon X \to X$ be an aperiodic IET, $X = [0, b)$. Then the critical set $X_{\psi=0}$ is meager and has Lebesgue measure* 0.

Theorem 1 follows immediately from Theorems 2 and 3, the proofs of which are presented in Sections 5 and 4, respectively.

3. Some notation, terminology and lemmas

The discussion in this section continues under the assumption that (X, f) is a fixed aperiodic r-IET, $X = [0, b)$, $r \geq 2$. Note that f^{-1} (the compositional inverse of f) is also an aperiodic r-IET on X.

Definition 1. An open subinterval $Y \subset X$ is called f-basic if the following equivalent conditions are met:

(b1) $f|_Y$ is a translation;

(b2) $f|_Y$ is continuous;

(b3) $Y \cap D = \emptyset$.

Given an f-basic interval Y, we write $\overset{+}{Y}(f)$ for the translation constant $f|_Y(y) - y$.

Observe that if an interval $Y \subset X$ is f-basic then $f(Y)$ is an f^{-1}-basic interval.

Definition 2. A sequence $\vec{Y} = (Y_k)_{k=1}^n$ of subsets of X is called an f-stack if the following conditions are met:

(s1) Each of the sets Y_k, $1 \leq k \leq n-1$, is an f-basic interval;

(s2) $f(Y_k) = Y_{k+1}$, for $1 \leq k \leq n-1$.

An f-stack $\vec{Y} = (Y_k)_{k=1}^n$ is called *distinct* if the sets Y_k are pairwise disjoint. Given an f-stack $\vec{Y} = (Y_k)_{k=1}^n$, we use the following terminology:

- **The width of $\vec{Y}$:** $\omega(\vec{Y}) = \lambda(Y_1)$ (in fact, all $\lambda(Y_k)$ are equal);
- **The support of $\vec{Y}$:** $\mathbf{supp}(\vec{Y}) = \bigcup_{k=1}^{n} Y_k \subset X$;
- **The length of $\vec{Y}$:** $h(\vec{Y}) = n$;
- **The measure of $\vec{Y}$:** $\lambda(\vec{Y}) = \lambda(\mathbf{supp}(\vec{Y}))$.

Note that if $\vec{Y} = (Y_k)_{k=1}^n$ is a distinct f-stack, then $\lambda(\vec{Y}) = \omega(\vec{Y})h(\vec{Y})$.

Observe that if $(Y_k)_{k=1}^n$ is an f-stack then the inverted sequence $(Y_{n+1-k})_{k=1}^n$ forms an f^{-1}-stack; in particular, the last set Y_n must also be an open subinterval of X (but not necessarily an f-basic one).

For $x \in \mathbb{R}$ and $\varepsilon > 0$, denote by $B_\varepsilon(x) = (x - \varepsilon, x + \varepsilon)$ the ε-neighborhood of $x \in \mathbb{R}$.

Proposition 1. *Let (X, T) be an aperiodic IET. For $\varepsilon > 0$, $x \in X \setminus D'$ and $n \geq 1$, the following three conditions are equivalent:*

(1a) *The sequence of intervals $\Big(B_\varepsilon\big(T^k(x)\big)\Big)_{k=-n}^{n}$ forms an f-stack (of length $(2n+1)$).*

(1b) $\varepsilon \leq \rho_n(x)$.

(1c) *There exists an f-stack $(Z_k)_{k=-n}^n$ with $Z_0 = B_\varepsilon(x)$.*

Proof. Follows from the definition of $\rho_n(x)$ (see (2.3a)). □

Proposition 2. *Under the assumptions and notations as in Proposition* 1, *assume that the equivalent conditions* (1a), (1b) *and* (1c) *hold. Then the following three conditions are equivalent:*

(2a) *The f-stack $\Big(B_\varepsilon\big(T^k(x)\big)\Big)_{k=-n}^{n}$ is distinct;*

(2b) $\varepsilon \leq \rho'_n(x)$.

(2c) *There exists a distinct f-stack $(Z_k)_{k=-n}^n$ with $Z_0 = B_\varepsilon(x)$.*

Proof. Follows from the definition of $\rho'_n(x)$ (see (2.3d)). □

The following (well known) lemma will be used in the proof of Theorem 3.

Lemma 1. *Let (X, f) be a minimal r-IET, $X = [0, b)$. Then for every N, there exists a distinct f-stack $\vec{Y}$ of length at least N and of measure at least $\frac{b}{r}$.*

The short proof is included for completeness.

Proof of Lemma 1. Pick a small subinterval $Y \subset X$, $0 < \lambda(Y) < \frac{b}{rN}$, so that the induced map g on Y is an s-IET, with some $2 \leq s \leq r$. Let $Y_k \subset Y$, $1 \leq k \leq s$, be the subintervals of Y exchanged by g: $g(Y_k) = f^{n_k}(Y_k)$. By minimality, the images $f^n(Y_k)$ cover $X = [0, b)$ before returning to Y.

More precisely, the family of subintervals $\big\{f^n(Y_k)\,\big|\,1 \leq k \leq s,\ 0 \leq n \leq n_k\big\}$ partitions the interval $X = [0, b)$. Thus $\lambda\Big(\bigcup_{n=0}^{n_k} f^n(Y_k)\Big) \geq \frac{b}{s} \geq \frac{b}{r}$, for some $k \in [1, s]$. To satisfy the conditions of the Lemma 1, one takes $\vec{Y} = \big(f^n(Y_k)\big)_{n=0}^{n_k}$. □

Remark. Lemma 1 holds under the weaker assumption that (X, f) is aperiodic (rather than minimal). We do not use the stronger version, and its proof is not included.

We would need the following notation. For an open finite interval $Y \subset \mathbb{R}$, denote by $\Theta(Y)$ the middle third subinterval of Y defined as the interval with the same center but 3 times shorter:

$$\Theta\big(B_\varepsilon(x)\big) = B_{\varepsilon/3}(x); \qquad \Theta\big((a,b)\big) = \big(\tfrac{2a+b}{3}, \tfrac{a+2b}{3}\big); \qquad \lambda(\Theta(Y)) = \tfrac{1}{3}\lambda(Y). \tag{3.1}$$

4. Proof of Theorem 3

One easily validates the f-invariance of ψ: $\psi(x) = \psi(f(x))$. So the Borel f-invariant set $X_{\psi>0}$ must have Lebesgue measure either 0 or 1, in view of the ergodicity of f.

Since (X, f) is an ergodic IET, it is minimal, so Lemma 1 applies. It follows that there exists a sequence $(\vec{Y}_n)_{n\geq 1}$ of distinct f-stacks with lengths $h(n) := h(\vec{Y}_n)$ approaching infinity and measures $\lambda(\vec{Y}_n) \geq \frac{b}{r}$ (see Definition 2). Let

$$\vec{Y}_n = \big(Y_{n,k}\big)_{k=1}^{h(n)} = (Y_{n,1}, Y_{n,2}, \ldots, Y_{n,h(n)}), \qquad n \in \mathbb{N}.$$

We may assume that all $h(n) \geq 6$. Let $p(n) = \left[\frac{h(n)}{3}\right]$ and $q(n) = h(n) - p(n) + 1$.

Consider the following sequence of distinct f-stacks $(\vec{Z}_n)_{n\geq 1}$:

$$\vec{Z}_n = \Big(\Theta\big(Y_{n,k}\big)\Big)_{k=p(n)+1}^{q(n)-1} = \Big(\Theta\big(Y_{n,p(n)+1}\big), \Theta\big(Y_{n,p(n)+2}\big), \ldots, \Theta\big(Y_{n,q(n)-1}\big)\Big), \; n \in \mathbb{N},$$

and set

$$Z_n = \mathbf{supp}(\vec{Z}_n) \subset X, \qquad n \in \mathbb{N},$$

(see Definition 2 for notation) and

$$Z = \bigcap_{n=1}^{\infty} \left(\bigcup_{k=n}^{\infty} Z_k \right) = \{z \in X \mid z \in Z_n, \text{ for infinitely many } n \geq 1\}.$$

The set Z is a dense G_δ subset of X, hence a residual subset of X.

Observe the following inequalities for the lengths and the widths of $\vec{Z}_n$:

$$h(\vec{Z}_n) = q(n) - p(n) - 1 \geq \tfrac{h(n)}{3}, \qquad \omega(\vec{Z}_n) = \tfrac{1}{3}\omega(\vec{Y}_n) \qquad (n \in \mathbb{N}).$$

It follows that

$$\lambda(Z_n) = \lambda(\vec{Z}_n) \geq \tfrac{1}{9}\lambda(\vec{Y}_n) \geq \tfrac{b}{9r} \qquad (n \in \mathbb{N}),$$

and therefore

$$\lambda(Z) \geq \limsup_{n\to\infty} \lambda(\vec{Z}_n) \geq \tfrac{b}{9r}.$$

LEMMA 2. *$\psi(z) > 0$ for all $z \in Z$.*

Since the function ψ is easily seen to be Borel measurable and f-invariant ($\psi(x) = \psi(f(x))$, for all $x \in X$), the ergodicity of f implies that $\lambda(X_{\psi>0}) \in \{0, 1\}$. The proof of Lemma 2 would imply that $\lambda(X_{\psi>0}) = 1$, completing the proof of Theorem 3.

PROOF OF LEMMA 2. Let $z \in Z$. Then there exists an increasing sequence of positive integers $(n_i)_{i=1}^{\infty}$ such that $z \in Z_{n_i}$ where

$$Z_{n_i} = \mathbf{supp}(\vec{Z}_{n_i}) = \bigcup_{k=p(n_i)+1}^{q(n_i)-1} \Theta(Y_{n_i,k}) \subset \mathbf{supp}(\vec{Y}_{n_i}) = \bigcup_{k=1}^{h(n_i)} Y_{n_i,k}, \qquad \text{for all } i \geq 1.$$

Set $\varepsilon_i :=\frac{1}{2}\omega(\vec{Z}_{n_i}) = \frac{1}{6}\omega(\vec{Y}_{n_i})$. Note that for every $i \geq 1$ the sequence

$$\vec{B}_i := \Big(f^k\big(B_{\varepsilon_i}(z)\big)\Big)_{k=-p(n_i)}^{p(n_i)}$$

forms a distinct f-stack due to the fact that $\vec{Y}_{n_i}$ does. By Proposition 2, $\varepsilon_i \leq \rho'_{p(n_i)}(z)$.

Since $\lim\limits_{i\to\infty} n_i = \infty$, we get both $\lim\limits_{i\to\infty} h(n_i) = \infty$ and $\lim\limits_{i\to\infty} p(n_i) = \infty$. One concludes that for all large i:

$$\rho'_{p(n_i)}(z) \cdot p(n_i) \geq \varepsilon_i \cdot h(n_i) \cdot \tfrac{p(n_i)}{h(n_i)} = \tfrac{1}{6} \cdot \lambda(\vec{Y}_{n_i}) \cdot \tfrac{p(n_i)}{h(n_i)} > \tfrac{1}{6} \cdot \tfrac{b}{r} \cdot \tfrac{1}{4} = \tfrac{b}{24r}.$$

The proof of Lemma 2 (and hence of Theorem 3) is completed by direct estimation (see (2.4b)):

$$\psi(z) = \limsup_{n\to\infty} n\rho'(n) \geq \tfrac{b}{24r} > 0.$$

□

5. Proof of Theorem 2

Denote by $S^1 = \{z \in \mathbb{C} \mid |z| = 1\} = \{e^{it} \mid t \in [0, 2\pi)\}$ the unit circle in the complex plane.

Theorem 2 is derived from the following proposition.

Proposition 3. *Let (X, f) be an aperiodic IET, $X = [0, b)$. Let $1 \neq \theta \in S^1$. Assume that $\psi(t) > 0$, for some $t \in (0, b)$. Then the equation*

$$F(f(x)) = \begin{cases} \theta \cdot F(x) & \text{if } x < t \\ F(x) & \text{if } x \geq t \end{cases} \tag{5.1}$$

has no (Lebesgue) measurable solutions $F : X \to S^1$.

Proof of Theorem 2. Since (X, f) is ergodic, all induced maps (X_t, f_t) also are. If some IET (X_t, f_t) fails to be weakly mixing, it has a nontrivial eigenvalue $\theta \in S^1$, $\theta \neq 1$.

Select an eigenfunction $G : X_t \to \mathbb{C}$ corresponding to θ so that $G(f_t(x)) = \theta \cdot G(x)$, for $x \in X_t$. The ergodicity of f_t implies that $|G|$ must be a constant which (without loss of generality) is assumed to be 1. Thus $G(X_t) \subset S^1$.

Define $F(x) = G(f^{k(x)}(x))$ with $k(x) = \min\big(\mathbb{N}_t(x)\big)$ where $\mathbb{N}_t(x) := \{k \geq 0 \mid f^k(x) \in X_t\}$. (Note that the set $\mathbb{N}_t(x)$ is not empty because f is ergodic and hence minimal).

The constructed function $F : X \to S^1$ is measurable (because G is) and is easily seen to satisfy (5.1). This contradicts the conclusion of Proposition 3, completing the proof of Theorem 2. □

6. Proof of Proposition 3

The proof goes by contradiction. Assume to the contrary that there are θ and F satisfying the conditions of Proposition 3 and that, in particular, (5.1) holds for some $t \in (0, b)$ such that

$$\psi(t) = \limsup_{n\to\infty} n\rho'_n(t) > 0.$$

Set $\varepsilon_k = \frac{\psi(t)}{2k}$, for $k \geq 1$. Then there exists an infinite subset $\mathbb{M} \subset \mathbb{N}$ of natural numbers such that $\rho'_n(t) > \varepsilon_n > 0$ for all $n \in \mathbb{M}$.

The statement of the following lemma follows from Proposition 2.

LEMMA 3. *The sequence* $\big(B_{\varepsilon_n}(f^k(t))\big)_{k=-n}^{n}$ *forms a distinct f-stack for every integer $n \in \mathbb{M}$.*

Some notation. For $n \in \mathbb{M}$ and $|k| \le n$, set the following open subintervals of X:

$$A_n^k = (f^k(t) - \varepsilon_n, f^k(t) + \varepsilon_n) = B_{\varepsilon_n}(f^k(t));$$
$$B_n^k = (f^k(t), f^k(t) + \varepsilon_n);$$
$$C_n^k = (f^k(t) - \varepsilon_n, f^k(t)).$$

Set the constants (all lying in $D_1 = \big\{z \in \mathbb{C} \bigm| |z| \le 1\big\}$):

$$\alpha_n^k = \mathcal{A}(A_n^k), \qquad \beta_n^k = \mathcal{A}(B_n^k), \qquad \gamma_n^k = \mathcal{A}(C_n^k), \tag{6.2}$$

where

$$\mathcal{A}(Y) = \tfrac{1}{\lambda(Y)} \int_Y F(x)\,dx \tag{6.3}$$

stands for the average of the function F over a subinterval $Y \subset X$.

Since F is S^1-valued, the constants in (6.2) lie in the unit disc $D_1 = \big\{z \in \mathbb{C} \bigm| |z| \le 1\big\}$.

By a passing to an infinite subset of $\mathbb{M}$, we may assume that the following six limits

$$\alpha^i = \lim_{\substack{n\in\mathbb{M}\\ n\to\infty}} \alpha_n^i, \qquad \beta^i = \lim_{\substack{n\in\mathbb{M}\\ n\to\infty}} \beta_n^i, \qquad \gamma^i = \lim_{\substack{n\in\mathbb{M}\\ n\to\infty}} \gamma_n^i, \qquad i \in \{0,1\} \tag{6.4}$$

exist and lie in D_1. We show that in fact all six constants $\alpha^1, \alpha^0, \beta^1, \beta^0, \gamma^1$ and γ^0 must lie in the unit circle $S^1 = \partial D_1$ (see Lemma 5 below). This will follow from Lemma 4 below.

Recall that, given an f-stack $\vec{Y}$, we write $h(\vec{Y})$ and $\lambda(\vec{Y})$ for the length and the measure of $\vec{Y}$ (see Definition 2).

LEMMA 4. *Let (X, f) be an aperiodic IET, $X = [0, b)$. Let $\theta \in S^1$ and $t \in (a, b)$. Assume that a measurable function $F\colon X \to S^1$ satisfies the equation (5.1). Let $(\vec{Y}_n)_{n=1}^{\infty}$ be a sequence of distinct f-stacks*

$$\vec{Y}_n = \big(Y_{n,k}\big)_{k=1}^{h_n} = (Y_{n,1}, Y_{n,2}, \ldots, Y_{n,h_n}),$$

satisfying the following three conditions:

(4A) $\lim\limits_{n\to\infty} h_n = \infty$ *(the lengths of stacks $h_n = h(\vec{Y}_n)$ approach infinity);*

(4B) $\liminf\limits_{n\to\infty} \lambda(\vec{Y}_n) > 0$ *(the measures of stacks $\lambda(\vec{Y}_n) = \sum\limits_{k=1}^{h(n)} \lambda(Y_{n,k})$ stay away from* 0*);*

(4C) $t \notin Y_{n,k}$, *for all n and $k \in [1, h_n]$.*

Then $\lim\limits_{n\to\infty} |\mathcal{A}(Y_{n,1})| = 1$ *(for notation see (6.3)).*

REMARK. In the above lemma, under the condition (4C) alone (i.e., without assuming (4A) and (4B)) we obviously have

$$|\mathcal{A}(Y_{n,1)}| = |\mathcal{A}(Y_{n,2)}| = \ldots = |\mathcal{A}(Y_{n,h_n)}|, \quad \text{for all } n \ge 1,$$

in view of the equation (5.1) the function F satisfies. In particular, the conclusion $\lim_{n\to\infty} |\mathcal{A}(Y_{n,1})| = 1$ in Lemma 4 is equivalent to the relation $\lim_{n\to\infty} |\mathcal{A}(Y_{n,h_n})| = 1$.

Proof of Lemma 4. In view of the assumption (4A),

$$\lambda(Y_{n,1}) = \lambda(Y_{n,2}) = \ldots = \lambda(Y_{n,h_n}) \le \tfrac{1}{h_n} \to 0 \qquad (\text{as } n \to \infty).$$

For subintervals $Y \subset X$, set $\mathcal{B}(Y) = \int_Y |F(x) - \mathcal{A}(Y)|\, dx$ where $\mathcal{A}(Y)$ stands for the average of F over Y (see (6.3)). We claim that

$$\lim_{n\to\infty} \Big(\sum_{k=1}^{h_n} \mathcal{B}(Y_{n,k}) \Big) = 0.$$

Indeed, let P_n be a partition of X with mesh $\le \frac{1}{h_n}$ containing all the intervals $Y_{n,k}$, $1 \le k \le h_n$. Then

$$\begin{aligned} 0 < \sum_{k=1}^{h_n} \mathcal{B}(Y_{n,k}) \le \sum_{Y\in P_n} \mathcal{B}(Y) &= \\ &= \int_X \Big| F(x) - (F \,|\, P_n)(x) \Big|\, dx \to 0 \qquad (\text{as } n \to \infty) \end{aligned}$$

where $(F \,|\, P_n)$ stands for the conditional expectation of F over the partition P_n.

Since F satisfies (5.1), we have

$$\mathcal{B}(Y_{n,1}) = \mathcal{B}(Y_{n,2}) = \ldots = \mathcal{B}(Y_{n,h_n}),$$

and hence $\lim_{n\to\infty} \mathcal{B}(Y_{n,1}) \cdot h_n = 0$. (Here we use the assumption (4C)). It follows that

$$\begin{aligned} \lim_{n\to\infty} \Big(\Big(\tfrac{1}{\lambda(Y_{n,1})} \cdot \mathcal{B}(Y_{n,1}) \Big) \cdot \big(h_n \cdot \lambda(Y_{n,1}) \big) \Big) &= \\ = \lim_{n\to\infty} \Big(\Big(\tfrac{1}{\lambda(Y_{n,1})} \cdot \mathcal{B}(Y_{n,1}) \Big) \cdot \lambda(\vec{Y}_n) \Big) &= 0. \end{aligned}$$

In view of the assumption (4B), we conclude that

$$\lim_{n\to\infty} \tfrac{1}{\lambda(Y_{n,1})} \cdot \mathcal{B}(Y_{n,1}) = \lim_{n\to\infty} \tfrac{1}{\lambda(Y_{n,1})} \int_{Y_{n,1}} |F(x) - \mathcal{A}(Y_{n,1})|\, dx = 0.$$

Let $\varepsilon > 0$ be given. Then for all sufficiently large n, one can select $x_n \in Y_{n,1}$ so that $F(x_n) \in S^1$ and $|F(x_n) - \mathcal{A}(Y_{n,1})| < \varepsilon$. This implies that $\big| |\mathcal{A}(Y_{n,1})| - 1 \big| < \varepsilon$, and, since $\varepsilon > 0$ is arbitrary, $\lim_{n\to\infty} |\mathcal{A}(Y_{n,1})| = 1$, completing the proof of Lemma 4. □

Lemma 5. *The six constants* $\alpha^1, \alpha^0, \beta^1, \beta^0, \gamma^1$ *and* γ^0 *(see* (6.2)*) lie in* S^1*.*

Proof. **Case of $\boldsymbol{\alpha^1}$** Recall that $\alpha^1 = \lim_{\substack{n\in\mathbb{M}\\ n\to\infty}} \alpha^1_n$ where $\alpha^1_n = \mathcal{A}(A^1_n)$.

We claim that the conditions of Lemma 4 are fulfilled with $\big(\vec{Y}_n\big)_{n\ge1} = \Big(\big(A^k_n\big)_{k=1}^n \Big)_{n\in\mathbb{M}}$. Indeed, in this case $\vec{Y}$ is a distinct f-stack in view of Lemma 3. We also have

$$h(\vec{Y}_n) = n \qquad \text{and} \qquad \lambda(\vec{Y}_n) = 2\varepsilon_n h(\vec{Y}_n) = 2\varepsilon_n n = \psi(t) = 2\varepsilon_1$$

for $n \in \mathbb{M}$. It follows from Lemma 4 that $|\alpha^1| = 1$.

Case of $\boldsymbol{\alpha^0}$. Similar argument. We set $\big(\vec{Y}_n\big)_{n\geq 1} = \Big(\big(A_n^k\big)_{k=-n+1}^{0}\Big)_{n\in\mathbb{M}}$ and take in account remark following Lemma 4 to get $|\alpha^0| = 1$.
Case of $\boldsymbol{\beta^1}$. We set $\big(\vec{Y}_n\big)_{n\geq 1} = \Big(\big(B_n^k\big)_{k=1}^{n}\Big)_{n\in\mathbb{M}}$ and in the same way apply Lemma 3 to get $|\beta^1| = 1$.
Case of $\boldsymbol{\beta^0}$. We set $\big(\vec{Y}_n\big)_{n\geq 1} = \Big(\big(B_n^k\big)_{k=-n+1}^{0}\Big)_{n\in\mathbb{M}}$ and in the same way apply Lemma 3 to get $|\beta^0| = 1$.
Cases of $\boldsymbol{\gamma^1}$ and $\boldsymbol{\gamma^0}$. Similar to the preceding two cases.

This completes the proof of Lemma 5.

□

Since $\mathcal{A}(A_n^k) = \frac{\mathcal{A}(B_n^k)+\mathcal{A}(C_n^k)}{2}$, it follows that, for both $i = 0, 1$, we have $\frac{|\beta^i+\gamma^i|}{2} = |\alpha^i| = 1$, and hence

$$\beta^i = \gamma^i = \alpha^i \in S^1.$$

On the other hand, the fact that F satisfies (5.1) implies that

$$\beta^1 = \beta^0; \qquad \gamma^1 = \theta \cdot \gamma^0.$$

We conclude that $\beta^1 = \gamma^1 = \theta \cdot \gamma^0 = \theta \cdot \beta^0 = \theta \cdot \beta^1$, whence $\theta = 1$, in contradiction with the initial assumption that $\theta \neq 1$.

The proof of Proposition 3 is complete.

7. Two extensions of Theorem 1

The following two theorems (Theorem 4 and 5) extends Theorem 1 to arbitrary f-invariant ergodic measures μ (rather than Lebesgue measure λ). These results are of interest in the case when IETs (X, f) are minimal but not uniquely ergodic. (Such IETs exist, see [**17**], [**18**]).

THEOREM 4. *Let (X, f) be a minimal μ-ergodic IET, $X = [0, b)$, where μ is an f-invariant Borel probability measure. Then the set*

$$X_{\mathrm{wm}}(\mu) = \{0 < t < b \mid f_t \text{ is weakly mixing (relative } \mu)\} \tag{7.1}$$

is a residual set of full μ-measure: $\mu(X_{\mathrm{wm}}(\mu)) = 1$.

PROOF. Let $\beta\colon X \to X$ be an increasing continuous bijection taking measure μ to λ. Then the composition $g = \beta^{-1}\circ f \circ \beta$ becomes a λ-ergodic IET which topologically and combinatorially is β-isomorphic to f. (This is the essense of the normalization procedure discussed e. g. in [**25**, Section 1]).

Theorem 1 applies to g to deduce Theorem 4. □

It is known that for any minimal IET (X, f) the set $\mathcal{P}_{\mathrm{erg}}(f)$ of ergodic f-invariant Borel probability measures on X is finite (see [**14**] and [**25**]).

THEOREM 5. *Let $f\colon X \to X$ be a minimal IET, $X = [0, b)$. Then*

$$X_{\mathrm{wm}}(\mathcal{P}_{\mathrm{erg}}) = \{0 < t < b \mid f_t \text{ is weakly mixing relative to every } \mu \in \mathcal{P}_{\mathrm{erg}}(f)\,\} \tag{7.2}$$

is a residual subset of X.

PROOF. Follows from Theorem 4 because

$$X_{\mathrm{wm}}(\mathcal{P}_{\mathrm{erg}}(f)) = \bigcap_{\mu\in\mathcal{P}_{\mathrm{erg}}(f)} X_{\mathrm{wm}}(\mu)$$

is a finite intersection of residual sets. □

8. Final comments

The discussion in this section is conducted under the assumption that (X, f) is a fixed λ-ergodic IET, $X = [0, 1]$, so that both X_{nwm} and $X_{\psi=0}$ (the exceptional and the critical sets, respectively) are defined (by (2.1b) and (2.5a)).

By Theorem 2, every exceptional point is critical, i.e., the inclusion

$$X_{\mathrm{nwm}}(f) = X_{\mathrm{nwm}} \subset X_{\psi=0} = X_{\psi=0}(f) \tag{8.1}$$

holds, and, by Theorem 3, the critical set $X_{\psi=0}$ is meager and has Lebesgue measure 0.

In this section we sketch some results on the size of the sets X_{nwm} and $X_{\psi=0}$, concerning their Hausdorff dimensions and cardinalities. The proofs and more detailed description of the results will appear elsewhere.

8.1. Case: Irrational rotation. Let $a \in \mathbb{R}$ and set $f(x) = R_a(x) = x + a \pmod 1$. (Such f can be viewed as a 2-IET). One verifies that in this case f_t is a 3-IET provided that $t \neq D' = \{na \mid n \in \mathbb{Z}\}$ (see (2.2)). Theorem 3 applies to deduce that $X_{\psi=0}$ is meager and has Lebesgue measure 0.

It follows from [**8**] that

$$\{\mathbb{Q} + \mathbb{Q}\,a\} \cap (X \setminus D') \cap X_{\psi=0} = \emptyset,$$

(i.e., no rational linear combination of 1 and a lying in $X \setminus D'$ is exceptional).

Both sets $X_{\mathrm{nwm}}(R_a)$ and $X_{\psi=0}(R_a)$ are countable if and only if a is badly approximable (the sequences of partial quotients in its continued fraction expansion is bounded). This fact can be deduced from [**24**]. In fact, if a is badly approximable then the three sets, X_{nwm}, $X_{\psi=0}$ and D', coincide.

We conjecture that for Lebesgue almost all a, the Hausdorff dimensions of the sets $X_{\mathrm{nwm}}(R_a)$ and $X_{\psi=0}(R_a)$ vanish. On the other hand, for some irrational a, both Hausdorff dimensions can be 1. (This last claim is based on an unpublished result by Yitwah Cheung).

8.2. Case: Linearly recurrent IETs. Every minimal r-IET f can be naturally coded by a minimal subshift Ω_f over the alphabet $\mathcal{A}_r = \{1, 2, \dots, r\}$ the following natural way described in [**16**]. (Every point $x \in X$ corresponds to the infinite word $\big(W_x(k)\big)_{k\in\mathbb{Z}} \in (\mathcal{A}_r)^{\mathbb{Z}}$ determined by the rule: $W_x(k) = s$ if $f^k(x) \in X_s = [d_{s-1}, d_s]$, see (1.1)).

A minimal r-IET f is said to be linearly recurrent if the corresponding subshift Ω_f is linearly recurrent in the sense of [**12**] (see also [**11**]). Linearly recurrent IETs also appear in the literature as "IETs of constant type" (a characterization in terms of Rauzy-Veech induction).

Linearly recurrent IETs include IETs of periodic type (also known as pseudo-Anosov IETs) and, in particular, minimal IETs over quadratic number fields (the IETs with the lengths of exchanged intervals lying in one and the same real quadratic number field). IETs over quadratic number fields always reduce to Pseudo-Anosov IETs (see [**6**]).

One can prove that for linear recurrent IETs (X, f) each of the sets $X_{\text{nwm}}(f)$ and $X_{\psi=0}(f)$ is at most countable.

We believe that for $r \geq 4$ there are pseudo-Anosov r-IETs which are persistently weakly mixing (i.e., for which $X_{\text{nwm}}(f) = \emptyset$).

References

[1] A. Avila, G. Forni, Weak mixing for interval exchange transformations and translation flows, *Ann. of Math.* **165** (2007), 637–664 MR2299743 (2010c:37002)

[2] V. Benesnevich, S. Velani, A note on zero-one laws in metrical Diophantine approximation. *Acta Arithmetica* **133** (2008), 363–474 MR2457266 (2009i:11092)

[3] M. Boshernitzan, A condition for minimal interval exchange transformation to be uniquely ergodic, *Duke J. of Math.* **52 (3)** (1985), 723–752 MR808101 (87i:28012)

[4] M. Boshernitzan, A condition for unique ergodicity of minimal symbolic flows, *Erg. Theory and Dynam. Sys.* **12** (1992), 425–428 MR1182655 (93j:58043)

[5] M. Boshernitzan, Rank two interval exchange transformations, *Erg. Theory Dynam. Systems* **8** (1988), no. 3, 379–394. MR961737 (90c:28024)

[6] M. Boshernitzan, C. Carroll, An extension of Lagrange's theorem to interval exchange transformations over quadratic fields. *J. Anal. Math.* **72** (1997), 21–44. MR1482988 (99a:58101)

[7] M. Boshernitzan, J. Chaika, Diophantine properties of IETs and general systems. Quantitative proximality and connectivity, arXiv:0910.5412 (2010)

[8] M. Boshernitzan, A. Nogueira, Generalized eigenfunctions of interval exchange maps, *Erg. Theory Dynam. Systems* **24** (2004), 697–705 MR2060994 (2005e:37007)

[9] J. Chaika, There exists a topologically mixing IET, arXiv: 0910.3986, to appear in *Erg. Theory and Dynam. Systems*

[10] I. P. Cornfeld, S. Fomin, Ya. G. Sinai, **Ergodic Theory**, Grundlehren der Mathematisches Wissenschaften [Fundamental Principles of Mathematical Sciences], 245. Springer-Verlag, New York, 1982. MR832433 (87f:28019)

[11] F. Durand, B. Host, C. Skau, Substitutive dynamical systems, Bratteli diagrams and dimension groups, *Erg. Theory and Dynam. Sys.* **19** (1999), 953-993. MR1709427 (2000i:46062)

[12] F. Durand, Linear recurrent subshifts have a finite number of non-periodic subshift factors, *Erg. Theory and Dynam. Sys.* **20** (2000), 1061-1078. MR1779393 (2001m:37022)

[13] A. Katok, Interval exchange transformations and some special flows are not mixing, *Israel Journal of Mathematics* **35 (4)** (1980), 301–310. MR594335 (82e:58060)

[14] A. Katok, Invariant measures of flows on oriented surfaces, Dokl. Nauk. SSR **211**, (1973), Sov. Math. Dokl. **14** (1973), 1104-1108. MR0331438 (48:9771)

[15] A. Katok, A. M. Stepin. Approximations in Ergodic Theory, *Russ. Math. Surv.* **22** (1967), 77–102. MR0219697 (36:2776)

[16] M. Keane, Interval exchange transformations, *Math. Z.* **141**, (1975), 25–31. MR0357739 (50:10207)

[17] M. Keane, Non-ergodic intervak exchange transformation, Isr. J. Math. **26** (1977), 188-196. MR0435353 (55:8313)

[18] H. B. Keynes, D. Newton, A "minimal", non-uniquely ergodic interval exchange transformation, *Math. Z.* **148** (1976), 101–105. MR0409766 (53:13518)

[19] H. B. Keynes, D. Newton, A "minimal", non-uniquely ergodic interval exchange transformation, *Math. Z.* **148** (1976), 101–105. MR0409766 (53:13518)

[20] S. Kerckhoff, Simplicial systems for interval exchange maps and measured foliations, *Erg. Theory and Dynam. Sys.* **(5)** (1982), 257–271. MR796753 (87g:58075)

[21] H. Masur, Interval exchange transformations and measured foliations, *Ann. of Math.* **(2) 115** (1982), 168–200. MR644018 (83e:28012)

[22] M. Rees, An alternative approach to the ergodic theory of measured foliations on surfaces. *Erg. Theory and Dynam. Sys.* **1** (1981), 461–488 MR662738 (84c:58065)

[23] Ya. G. Sinai, C. Ulcigrai, Weak mixing in interval exchange transformations of periodic type. *Lett. Math. Phys.* 74 (2005), no. 2, 111–133. MR2191950 (2009k:37004)

[24] W. Veech, Strict Ergodicity in Zero Dimensional Dynamical Systems and the Kronecker-Weyl Theorem Mod 2. *Trans. of Amer. Math. Soc.* **140** (1969), 1–33 MR0240056 (39:1410)

[25] W. Veech, Interval exchange transformations. *J. D'Analyse Math.* **33** (1978), 222–272 MR516048 (80e:28034)

[26] W. Veech, Gauss measures for transformations on the space of interval exchange maps. *Ann. of Math.* (2) **115** (1982), 201-242 MR644019 (83g:28036b)

[27] W. Veech, The metric theory of interval exchange transformations, I. Generic spectral properties. *Amer. Journal of Math.* **106** (6) (1984), 1331–1422 MR765582 (87j:28024a)

[28] W. Veech, Boshernitzan's criterion for unique ergodicity of an interval exchange transformation, *Erg. Theory and Dynam. Sys.* **7** (1987), 149–153 MR886376 (88d:28022)

[29] M. Viana, Ergodic theory of interval exchange maps, *Rev. Mat. Complut.* **19** (2006), no.1, 7–100. MR2219821 (2007f:37002)

Department of Mathematics, Rice University, Houston, Texas 77005
E-mail address: michael@rice.edu

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11234

Every countably infinite group is almost Ornstein

Lewis Bowen

This paper is dedicated to Anatoly Stepin on the occasion of his 70th birthday.

Abstract. We say that a countable discrete group G is *almost Ornstein* if for every pair of standard non-two-atom probability spaces $(K,\kappa),(L,\lambda)$ with the same Shannon entropy, the Bernoulli shifts $G\curvearrowright(K^G,\kappa^G)$ and $G\curvearrowright(L^G,\lambda^G)$ are isomorphic.

This paper proves every countably infinite group is almost Ornstein.

1. Introduction

All probability spaces in this paper are standard. Let G be a countable discrete group and (K,κ) a probability space. Let K^G be the space of maps $x : G \to K$ with the product measure κ^G. The group acts on this space by $gx(f) = x(g^{-1}f)$ for every $f,g \in G$ and $x \in K^G$. The dynamical system $G\curvearrowright(K^G,\kappa^G)$ is called the *Bernoulli shift* over G with *base space* (K,κ). If (L,λ) is another probability space then a measurable map $\phi : K^G \to L^G$ is *shift-equivariant* if $\phi(gx) = g\phi(x)$ for every $g \in G$ and a.e. $x \in K^G$. It is an *isomorphism* if in addition it has a measurable inverse and $\phi_*\kappa^G = \lambda^G$.

In the early years of measurable dynamics, von Neumann asked whether $\mathbb{Z}\curvearrowright(\{0,1\}^{\mathbb{Z}},u_2^{\mathbb{Z}})$ is isomorphic to $\mathbb{Z}\curvearrowright(\{0,1,2\}^{\mathbb{Z}},u_3^{\mathbb{Z}})$ where u_i is the uniform probability measure on $\{0,1,\dots,i-1\}$. Kolmogorov [**Ko58, Ko59**] answered this question by introducing dynamical entropy for general probability measure-preserving transformations and computing this invariant for Bernoulli shifts over $\mathbb{Z}$. To be precise, the *Shannon entropy* of a probability space (K,κ) is defined as follows. If there exists a countable set $K' \subset K$ such that $\kappa(K') = 1$ then

$$H(K,\kappa) = -\sum_{k\in K'} \kappa(\{k\})\log(\kappa(\{k\}))$$

where, by convention, $0\log(0) = 0$. Otherwise, $H(K,\kappa) = +\infty$. Kolmogorov proved that *if* $\mathbb{Z}\curvearrowright(K^{\mathbb{Z}},\kappa^{\mathbb{Z}})$ is isomorphic to $\mathbb{Z}\curvearrowright(L^{\mathbb{Z}},\lambda^{\mathbb{Z}})$ then $H(K,\kappa) = H(L,\lambda)$, thereby answering von Neumann's question in the negative.

In [**Or70a, Or70b**], Ornstein famously proved the converse:

Theorem 1.1 (Ornstein). *Let $(K,\kappa),(L,\lambda)$ be two probability spaces with $H(K,\kappa) = H(L,\lambda)$. Then $\mathbb{Z}\curvearrowright(K^{\mathbb{Z}},\kappa^{\mathbb{Z}})$ is isomorphic to the $\mathbb{Z}\curvearrowright(L^{\mathbb{Z}},\lambda^{\mathbb{Z}})$.*

2010 *Mathematics Subject Classification.* Primary 37A35, Secondary 37A20.
Key words and phrases. Bernoulli shifts, entropy, Ornstein.
The author was supported in part by NSF Grants #0968762 and #0954606.

Motivated by this result, Stepin [**St76**] made the following definition: a countable group G is *Ornstein* if for every pair of probability spaces (K,κ), (L,λ) with the same Shannon entropy, the Bernoulli shift $G\curvearrowright(K^G,\kappa^G)$ is isomorphic to $G\curvearrowright(L^G,\lambda^G)$. Note that no finite group is Ornstein. However, Ornstein and Weiss [**OW87**] proved that every countably infinite amenable group is Ornstein. Using a co-induction argument, Stepin [**St75**] observed that if G has an Ornstein subgroup H then G is itself Ornstein. Therefore, every group which contains an infinite amenable subgroup (for example, an infinite cyclic subgroup) is Ornstein. However, there are countable groups which do not satisfy this property (for example, Ol'shanskii's monster [**Ol91**]).

A probability space (K,κ) is a *two-atom space* if κ is supported on two atoms: i.e., there exist elements $k_0,k_1\in K$ such that $\kappa(\{k_0,k_1\})=1$. We say that a group G is *almost Ornstein* if whenever $(K,\kappa),(L,\lambda)$ are standard probability spaces, neither of which is a two-atom space and $H(K,\kappa)=H(L,\lambda)$ then $G\curvearrowright(K^G,\kappa^G)$ is isomorphic to $G\curvearrowright(L^G,\lambda^G)$. So G is almost Ornstein if it is Ornstein with the possible exception of the 1-parameter family of two-atom probability spaces. Our main results are:

THEOREM 1.2. *Every countably infinite group is almost Ornstein.*

THEOREM 1.3. *If G is a countable group, $C<G$ is a cyclic subgroup of prime order and the normalizer of C in G has infinite index, then G is Ornstein.*

Unfortunately, Theorem 1.3 is not sufficient to conclude that every infinite countable group is Ornstein. There is a group G, constructed in Theorem 31.8 of [**Ol91**] which is generated by 2 elements and splits as a central extension

$$1\to C\to G\to T\to 1$$

where $C=\mathbb{Z}/p\mathbb{Z}$ (p is a prime $>10^{70}$) and G and T are Tarski Monsters (meaning that all proper subgroups of G and T are finite and cyclic). In addition, every nontrivial subgroup of G contains C. Therefore, G does not contain any cyclic subgroup of prime order with infinite-index normalizer. According to D. Osin [**Os11**] it is possible to modify the construction (in the spirit of section 37 of [**Ol91**]) to ensure that G is also non-amenable. Therefore, it does not contain any infinite amenable subgroups and we cannot say whether or not it is Ornstein. However we can say:

COROLLARY 1.4. *If every countable infinite group G with a nontrivial center is Ornstein, then every countably infinite group is Ornstein.*

PROOF. Assuming the hypothesis, let G be an arbitrary countably infinite group. We need to show it is Ornstein. By Stepin's Theorem [**St75**], we may assume G does not contain any infinite cyclic subgroups. Therefore, it contains a cyclic subgroup C of prime order. By Theorem 1.3, we may assume the normalizer of C has finite index in G. Because C is finite, its centralizer has finite index in its normalizer. Therefore, the centralizer H of C has finite index in G and so H is infinite. By hypothesis, H is Ornstein. By Stepin's Theorem [**St75**], this implies G is Ornstein. $\square$

In [**Bo10**] (see also [**KL1, KL2**] for an alternative approach) it is shown that if G is any sofic group, then the entropy of the base space is an invariant. Therefore:

COROLLARY 1.5. *If G is a countably infinite sofic group and $(K,\kappa), (L,\lambda)$ are two probability spaces, neither of which is a two-atom space, then the Bernoulli shifts $G \curvearrowright (K^G,\kappa^G)$ and $G \curvearrowright (L^G,\lambda^G)$ are isomorphic if and only if $H(K,\kappa) = H(L,\lambda)$.*

Recall that a shift-equivariant map $\phi : K^G \to L^G$ is a *factor map* between Bernoulli shifts $G \curvearrowright (K^G,\kappa^G)$ and $G \curvearrowright (L^G,\lambda^G)$ if $\phi_*\kappa^G = \lambda^G$. In this case, we say that $G \curvearrowright (K^G,\kappa^G)$ *factors* onto $G \curvearrowright (L^G,\lambda^G)$. Sinai [**Si59, Si62**] showed that the Bernoulli shift $\mathbb{Z} \curvearrowright (K^{\mathbb{Z}},\kappa^{\mathbb{Z}})$ factors onto $\mathbb{Z} \curvearrowright (L^{\mathbb{Z}},\lambda^{\mathbb{Z}})$ if and only if $H(K,\kappa) \ge H(L,\lambda)$. This result extends to countably infinite amenable groups by [**OW87**]. Non-amenable groups behave in a very different manner:

COROLLARY 1.6. *If G is a countable non-amenable group then there exists a number $0 \le r(G) < \infty$ such that if (K,κ) is any non-two-atom probability space and $H(K,\kappa) > r(G)$ then $G \curvearrowright (K^G,\kappa^G)$ factors onto every Bernoulli shift over G.*

In [**Bo11**], it is shown that if G contains a non-abelian free subgroup then every Bernoulli shift over G factors onto every Bernoulli shift over G. It is an open question whether every non-amenable group satisfies this property.

PROOF. Let $r(G)$ be the infimum over all numbers r such that there exists a probability space (K,κ) with $H(K,\kappa) = r$ such that $G \curvearrowright (K^G,\kappa^G)$ factors onto every Bernoulli shift over G. It follows from [**Ba05**] Theorem 6.4, that $r(G) < \infty$ if G is finitely generated. If G is not finitely generated then it contains a finitely generated non-amenable subgroup H. An easy co-induction argument shows that $r(G) \le r(H) < \infty$ (see e.g., [**Bo11**] which shows how $G \curvearrowright (K^G,\kappa^G)$ is co-induced from $H \curvearrowright (K^H,\kappa^H)$).

We claim that if (K,κ) is a non-two-atom probability space and (L,λ) is an arbitrary probability space with $H(K,\kappa) > H(L,\lambda)$ then $G \curvearrowright (K^G,\kappa^G)$ factors onto $G \curvearrowright (L^G,\lambda^G)$. To see this, let (M,μ) be a probability space with $H(K,\kappa) = H(M,\mu) + H(L,\lambda)$. By Theorem 1.2, $G \curvearrowright (K^G,\kappa^G)$ is isomorphic to $G \curvearrowright ((M \times L)^G, (\mu\times\lambda)^G)$ which clearly factors onto $G \curvearrowright (L^G,\lambda^G)$.

Now let (M,μ) be a non-two-atom probability space with $H(M,\mu) > r(G)$. Then there exists a probability space (K,κ) with $H(K,\kappa) < H(M,\mu)$ such that $G \curvearrowright (K^G,\kappa^G)$ factors onto every Bernoulli shift. Since $G \curvearrowright (M^G,\mu^G)$ factors onto $G \curvearrowright (K^G,\kappa^G)$, it follows that $G \curvearrowright (M^G,\mu^G)$ factors onto every Bernoulli shift as claimed. $\square$

The main ingredients of the proof of Theorems 1.2 and 1.3 are (i) Thouvenot's relative isomorphism theorem for actions of $\mathbb{Z}$ [**Th75**], (ii) the fact that the full group of any p.m.p. aperiodic equivalence relation contains an aperiodic automorphism, (iii) a co-induction argument similar in spirit to Stepin's [**St75**]. The idea to use elements in the full group of a factor to obtain isomorphism theorems originated in [**RW00**] and was applied in [**DP02**] to obtain a version of Thouvenot's relative isomorphism theorem for actions of amenable groups.

Acknowledgements. I'd like to thank Brandon Seward for sketching a proof of Theorem 2.2. Thanks also to Hanfeng Li for pointing out several errors in a previous version. Also thanks to Benjy Weiss for helpful suggestions which improved the paper and corrected an error. I am indebted to Denis Osin for explaining that Ol'shankii's monsters are still potential counterexamples to the claim that every countably infinite group is Ornstein.

2. Preliminaries

All probability spaces in this paper are standard and may be atomic or non-atomic. Often we denote a probability space by (X, μ) without referencing the sigma-algebra. All maps, functions, relations, etc., are considered up to sets of measure zero.

2.1. Entropy. Let $(X, \mathcal{B}, \mu)$ be a standard probability space. Let $\chi : X \to K$ be a measurable map. The entropy of χ is $H(\chi) = H(K, \chi_*\mu)$ (i.e., it is the entropy of the partition $\chi^{-1}(P)$ where P is the partition of K into points). If $\mathcal{C} \subset \mathcal{B}$ is a sub-sigma algebra, let $H(\chi|\mathcal{C})$ denote the relative entropy.

Let $T \in \mathrm{Aut}(X, \mathcal{B}_X, \mu)$ be an automorphism of $(X, \mathcal{B}_X, \mu)$. Let $\mathcal{C} \subset \mathcal{B}_X$ be a T-invariant sub-sigma-algebra. Let $h(T, \chi|\mathcal{C}) = \lim_{n\to\infty}(2n+1)^{-1}H(\bigvee_{i=-n}^{n} \chi \circ T^i|\mathcal{C})$ denote the relative entropy rate of χ. Let $h(T|\mathcal{C}) = \sup h(T, \chi|\mathcal{C})$ where the supremum is over all measurable maps χ with finite range.

A measurable map $\chi : X \to K$ is a *generator* for $(T, X, \mathcal{B}_X, \mu)$ if $\mathcal{B}_X$ is the smallest T-invariant sigma-algebra under which χ is measurable. In this case, if $\mathcal{C}$ is any T-invariant sub-sigma-algebra then the Kolmogorov-Sinai Theorem implies that if $H(\chi|\mathcal{C}) < \infty$ then $h(T|\mathcal{C}) = h(T, \chi|\mathcal{C})$.

2.2. Thouvenot's relative isomorphism theorem. Let $T \in \mathrm{Aut}(X, \mathcal{B}_X, \mu)$, $U \in \mathrm{Aut}(Z, \mathcal{B}_Z, \zeta)$ and suppose $\pi : X \to Z$ is a factor map. That is, π is measurable, $\pi_*\mu = \zeta$ and $\pi T = U\pi$. Then we say that $(T, X, \mathcal{B}_X, \mu)$ is *Bernoulli* relative to $(U, Z, \mathcal{B}_Z, \zeta)$ if there is a generator $\chi : X \to K$ for $(T, X, \mathcal{B}_X, \mu)$ such that the random variables $\{\chi \circ T^i : i \in \mathbb{Z}\}$ are jointly independent relative to $\pi^{-1}(\mathcal{B}_Z)$. The dependence on π in this definition is left implicit. The next result is in [**Th75**].

THEOREM 2.1 (Thouvenot). *Let $T \in Aut(X, \mathcal{B}_X, \mu)$, $S \in Aut(Y, \mathcal{B}_Y, \nu)$, $U \in Aut(Z, \mathcal{B}_Z, \zeta)$ and $\pi_X : X \to Z$, $\pi_Y : Y \to Z$ be factor maps. Suppose $(T, X, \mathcal{B}_X, \mu)$ and $(S, Y, \mathcal{B}_Y, \nu)$ are each Bernoulli relative to $(U, Z, \mathcal{B}_Z, \zeta)$ and U is ergodic. Suppose also that $h(T|\pi_X^{-1}(\mathcal{B}_Z)) = h(S|\pi_Y^{-1}(\mathcal{B}_Z))$. Then there is a measure-space isomorphism $\phi : (X, \mathcal{B}_X, \mu) \to (Y, \mathcal{B}_Y, \nu)$ such that $\phi T = S\phi$ and $\pi_X = \pi_Y \phi$.*

2.3. Measured equivalence relations. A *measurable equivalence relation* on a Borel space X is an equivalence relation E on X such that E is a Borel subset of $X \times X$. If $(x, y) \in E$ we write xEy. We say that E is *countable* if every E-equivalence class is countable. An *inner automorphism* of E is a Borel isomorphism $\phi : X \to X$ such that the graph of ϕ is contained in E. The group of all inner automorphisms is called the *full group* and denoted by $[E]$.

Now assume μ is a probability measure on X so that (X, μ) is a standard probability space. If $\phi_*\mu = \mu$ for every $\phi \in [E]$ then we say (X, μ, E) is a *probability measure preserving* (p.m.p.) equivalence relation. We also say that μ is E-invariant. For example, if $G \curvearrowright (X, \mu)$ is a probability measure-preserving action of a countable group G and $E := \{(x, gx) : x \in X, g \in G\}$ then μ is E-invariant. We say (X, μ, E) is *aperiodic* if for a.e. $x \in X$ the E-equivalence class of x is infinite. We say (X, μ, E) is *ergodic* if for every measurable set $A \subset X$ which is a union of E-classes, either $\mu(A) = 0$ or $\mu(A) = 1$.

THEOREM 2.2. *Let (X, μ, E) be an aperiodic ergodic p.m.p. equivalence relation. Then there exists an ergodic automorphism $T \in [E]$ such that for a.e. $x \in X$, $\{T^i x : i \in \mathbb{Z}\}$ is infinite.*

PROOF. This is a classical "folk" theorem. It is a special case of Corollary 3.3 of [**Me93**]. It is also proven in [**Ke10**], Theorem 3.5. □

3. Almost Ornstein groups

LEMMA 3.1. *Let G be a countably infinite group and let $(K,\kappa),(L,\lambda),(M,\mu)$ be standard probability spaces with $H(K,\kappa)=H(L,\lambda)$. In addition, suppose there exist measurable maps $\alpha : K \to M$ and $\beta : L \to M$ such that $\alpha_*\kappa = \beta_*\lambda = \mu$ and (M,μ) is nontrivial (i.e., $H(M,\mu)>0$). Then $G \curvearrowright (K^G,\kappa^G)$ is isomorphic to $G \curvearrowright (L^G,\lambda^G)$.*

PROOF. Let $\alpha^G : K^G \to M^G$ be the product map: $\alpha^G(x)(g) := \alpha(x(g))$. Define $\beta^G : L^G \to M^G$ similarly. Note $\alpha^G_*\kappa^G = \beta^G_*\lambda^G = \mu^G$.

By Theorem 2.2 there exists an ergodic $U \in \mathrm{Aut}(M^G,\mu^G)$ such that (i) for a.e. $x \in M^G$, the orbit $\{U^i x : i \in \mathbb{Z}\}$ is infinite, and (ii) for a.e. $x \in M^G$ there is a $g \in G$ such that $Ux = gx$.

Define $T : K^G \to K^G$ by $T(x) = gx$ where $g \in G$ is such that $U(\alpha^G(x)) = g\alpha^G(x)$. Similarly, let $S : L^G \to L^G$ be defined by $S(y) = gy$ where $g \in G$ is such that $U(\beta^G(y)) = g\beta^G(y)$. Note that for a.e. $(x,y) \in K^G \times L^G$ with $\alpha^G(x) = \beta^G(y)$, there is a $g \in G$ such that $Tx = gx$ and $Sy = gy$.

Let $\chi_M : M^G \to M$ be the projection map $\chi_M(x) = x(e)$. Let $\mathcal{B}_M$ be the smallest U-invariant sigma-algebra on M^G for which χ_M is measurable. Define $\chi_K : K^G \to K, \chi_L : L^G \to L, \mathcal{B}_K, \mathcal{B}_L$ similarly.

Observe that $(T,K^G,\mathcal{B}_K,\kappa^G)$ and $(S,L^G,\mathcal{B}_L,\lambda^G)$ are Bernoulli relative to $(U,M^G,\mathcal{B}_M,\mu^G)$. To see this, note, for example, that χ_K is a generator for $(T,K^G,\mathcal{B}_K,\kappa^G)$ and $\{\chi_K \circ T^i : i \in \mathbb{Z}\}$ are jointly independent relative to $(\alpha^G)^{-1}(\mathcal{B}_M)$. So

$$h(T|(\alpha^G)^{-1}(\mathcal{B}_M)) = H(\chi_K|\ (\alpha^G)^{-1}(\mathcal{B}_M)) = H(K,\kappa) - H(M,\mu).$$

Similarly, $H(L,\lambda) - H(M,\mu) = h(S|(\beta^G)^{-1}(\mathcal{B}_M))$. By Theorem 2.1, there is an isomorphism $\phi : (K^G,\mathcal{B}_K,\kappa^G) \to (L^G,\mathcal{B}_L,\lambda^G)$ such that $\phi T = S\phi$ and $\alpha^G = \beta^G\phi$.

Define $\Phi : K^G \to L^G$ by $\Phi(x)(g) = \phi(g^{-1}x)(e)$ for $g \in G$, $x \in K^G$. Observe that Φ is measurable with respect to the Borel sigma-algebras of K^G and L^G (which are, in general, larger than $\mathcal{B}_K$ and $\mathcal{B}_L$). We claim this is the required isomorphism. It is clearly shift-equivariant.

To see that Φ is invertible, define $\Psi : L^G \to K^G$ by $\Psi(y)(g) = \phi^{-1}(g^{-1}y)(e)$. Because both Φ and Ψ are shift-equivariant and $\Phi\Psi(y)(e) = y(e)$, $\Psi\Phi(x)(e) = x(e)$, it follows that Ψ is the inverse of Φ. It is easy to check that $\beta^G\Phi = \alpha^G$ and $\Phi T = S\Phi$.

To finish, we must show that $\Phi_*\kappa^G = \lambda^G$. For $z \in M^G$, let $X_z = \{x \in K^G :\ \alpha^G(x) = z\}$ and $Y_z = \{y \in L^G :\ \beta^G(y) = z\}$. Because $\beta^G\Phi = \alpha^G$, it follows that Φ maps X_z to Y_z.

For $z \in M^G$, let ζ_z be the fiber measure of κ^G over z. This family of measures is determined (up to measure zero sets) by the property that ζ_z is supported on X_z and $\kappa^G = \int \zeta_z\, d\mu^G(z)$. Similarly, let ν_z be the fiber measure of λ^G over z.

Because $\alpha^G_*\kappa^G = \beta^G_*\lambda^G = \mu^G$ and Φ maps X_z to Y_z, in order to show that $\Phi_*\kappa^G = \lambda^G$, it suffices to show that Φ restricts to an isomorphism from (X_z,ζ_z) to (Y_z,ν_z) for a.e. z.

For $g \in G$ and $n \in \mathbb{Z}$ let $\tau_z^n(g) \in G$ be the element satisfying $U^n(g^{-1}z)(e) = z(\tau_z^n(g))$. This is well-defined for a.e. $z \in M^G$. If $x \in X_z$ then $T^n(g^{-1}x)(e) = x(\tau_z^n(g))$ by definition.

Fix $z \in M^G$. Let $e = g_0, g_1, \ldots \in G$ be such that if $O_i = \{\tau_z^n(g_i) : n \in \mathbb{Z}\}$ then $O_i \cap O_j = \emptyset$ whenever $i \neq j$ and $\cup_{i=0}^{\infty} O_i = G$. So $K^G = \prod_{i=0}^{\infty} K^{O_i}$.

Let $\mathcal{B}_{K,z,0}$ be the restriction of $\mathcal{B}_K$ to X_z. This is the σ-algebra of X_z generated by $\{\chi_K \circ T^j : j \in \mathbb{Z}\}$. More generally, let $\mathcal{B}_{K,z,i}$ be the sigma-algebra on X_z generated by $\{\chi_K \circ T^j g_i^{-1} : j \in \mathbb{Z}\}$. Define σ-algebras $\mathcal{B}_{L,z,i}$ on Y_z similarly.

Because the O_i's are pairwise disjoint, the sigma-algebras $\mathcal{B}_{K,z,i}$ are independent. Because $\cup O_i = G$, these sigma-algebras generate the Borel sigma-algebra of K^G restricted to X_z. Similarly statements apply to L in place of K. Therefore, it suffices to show that Φ restricted to X_z determines an isomorphism from $(X_z, \mathcal{B}_{K,z,i}, \zeta_z)$ to $(Y_z, \mathcal{B}_{L,z,i}, \nu_z)$ for all i.

By definition, ϕ (and therefore Φ) restricted to X_z is an isomorphism from $(X_z, \mathcal{B}_{K,z,0}, \zeta_z)$ to $(Y_z, \mathcal{B}_{L,z,0}, \nu_z)$. Note that $(X_{g_i^{-1}z}, \mathcal{B}_{K,g_i^{-1}z,0}, \zeta_{g_i^{-1}z}) = g_i^{-1}(X_z, \mathcal{B}_{K,z,i}, \zeta_z)$. Because Φ is shift-equivariant, this implies Φ restricted to X_z determines an isomorphism from $(X_z, \mathcal{B}_{K,z,i}, \zeta_z)$ to $(Y_z, \mathcal{B}_{L,z,i}, \nu_z)$ for every i. $\square$

Proof of Theorem 1.2. Let G be a countably infinite group. Let (K, κ), (L, λ) be standard probability spaces with $H(K, \kappa) = H(L, \lambda)$. In addition, suppose both (K, κ) and (L, λ) are not two-atom spaces. We must show that the Bernoulli shifts $G \curvearrowright (K^G, \kappa^G)$ and $G \curvearrowright (L^G, \lambda^G)$ are isomorphic.

For $p \in [0, 1]$ let m_p be the probability measure on $\{0, 1\}$ given $m_p(\{0\}) = p$, $m_p(\{1\}) = 1 - p$. If, say (K, κ) is not purely atomic then there is some $p_0 > 0$ such that for all $p < p_0$, (K, κ) maps onto $(\{0, 1\}, m_p)$. In this case, $H(K, \kappa) = \infty = H(L, \lambda)$ which implies that for some $0 < p < p_0$, (L, λ) also maps onto $(\{0, 1\}, m_p)$. So the previous lemma implies the result.

Let us now assume that (K, κ) and (L, λ) are purely atomic. Borrowing an idea from [**KS79**] (Lemma 2), we observe that if $t > 0$ is the largest number such that $\kappa(\{k\}) = t$ for some $k \in K$ then there is a number $s > 0$ such that $\lambda(\{l\}) = s$ for some $l \in L$ and $t + s < 1$. Then there is a countable (or finite) set N with a probability measure ν so that for some $n_0, n_1 \in N$, $\nu(\{n_0\}) = t$, $\nu(\{n_1\}) = s$ and $H(N, \nu) = H(K, \kappa) = H(L, \lambda)$. In particular, both (K, κ) and (N, ν) map onto $(\{0, 1\}, m_t)$. Also (L, λ) and (N, ν) map onto $(\{0, 1\}, m_s)$. So the previous lemma implies $G \curvearrowright (K^G, \kappa^G)$ is isomorphic to $G \curvearrowright (N^G, \nu^G)$ which is isomorphic to $G \curvearrowright (L^G, \lambda^G)$. $\square$

4. Measurable subgroups

In order to prove Theorem 1.3 we extend the results of the previous section to so-called *measurable subgroups* of a group G. To be precise, let G be a countably infinite group and 2^G be the set of all subsets of G. G acts on 2^G by $(g, F) \mapsto gF = \{gf : f \in F\}$ for $g \in G$ and $F \in 2^G$. Let $\mathcal{R}$ be the orbit-equivalence relation on 2^G: $F \mathcal{R} H \Leftrightarrow \exists g \in G$ such that $F = gH$. Let 2^G_e be the set of all $F \in 2^G$ such that $e \in F$. Let $\mathcal{R}_e = \mathcal{R} \cap 2^G_e \times 2^G_e$ be the restriction of $\mathcal{R}$ to 2^G_e. A *measurable subgroup* of G is an $\mathcal{R}_e$-invariant probability measure η on 2^G_e.

To justify the definition, note that a subgroup is any subset $H \subset G$ which contains the identity and satisfies $h^{-1}H = H$ for all $h \in H$. A measurable subgroup η is the law of a random subset H which contains the identity and has the property that if $H \mapsto h_H \in H$ is a Borel assignment then $h_H^{-1}H$ has the same law as H (as long as $H \mapsto h_H^{-1}H$ is Borel-invertible). For example, if H is a subgroup then δ_H, the Dirac probability measure concentrated at $\{H\}$, is a measurable subgroup.

Let (K,κ) be a standard probability space. Let $2^G \otimes K$ be the set of all maps $x : \mathrm{Dom}(x) \to K$ with $\mathrm{Dom}(x) \subset G$. G acts on this space by $gx(f) = x(g^{-1}f)$ (for $x \in 2^G \otimes K$, $g \in G, f \in g\mathrm{Dom}(x)$). Note that $\mathrm{Dom}(gx) = g\mathrm{Dom}(x)$. Let $\mathcal{R}_K$ be the orbit-equivalence relation on $2^G \otimes K$: so $x\mathcal{R}_K y \Leftrightarrow \exists g \in G$ such that $gx = y$.

Let $2^G_e \otimes K$ be the set of all $x \in 2^G \otimes K$ with $e \in \mathrm{Dom}(x)$. Let $\mathcal{R}_{e,K}$ be $\mathcal{R}_K$ restricted to $2^G_e \otimes K$.

If η is a measurable subgroup of G then let $\eta \otimes \kappa$ be the probability measure on $2^G_e \otimes K$ defined by

$$\eta \otimes \kappa = \int \delta_H \times \kappa^H \, d\eta(H)$$

where δ_H is the Dirac measure concentrated on $\{H\}$ and κ^H is the product measure on K^H. This measure is $\mathcal{R}_{e,K}$-invariant and projects to η. It is the *Bernoulli shift* over η with base space (K,κ).

Let (L,λ) be another standard probability space. We say the two Bernoulli shifts $(2^G_e \otimes K, \eta \otimes \kappa)$ and $(2^G_e \otimes L, \eta \otimes \lambda)$ are isomorphic if there is a measurable map $\phi : 2^G_e \otimes K \to 2^G_e \otimes L$ such that

(1) $\mathrm{Dom}(\phi(x)) = \mathrm{Dom}(x)$ for a.e. x,
(2) $\phi_*\eta \otimes \kappa = \eta \otimes \lambda$,
(3) ϕ is invertible with measurable inverse,
(4) $\phi(gx) = g\phi(x)$ for a.e. x and every $g \in G$ with $g^{-1} \in \mathrm{Dom}(x)$ (i.e., $e \in \mathrm{Dom}(gx)$).

LEMMA 4.1. *Let η be an ergodic measurable subgroup of countable group G such that η-a.e. $H \in 2^G_e$ is infinite. Let $(K,\kappa), (L,\lambda), (M,\mu)$ be standard probability spaces with $H(K,\kappa) = H(L,\lambda)$. In addition, suppose there exist measurable maps $\alpha : K \to M$ and $\beta : L \to M$ such that $\alpha_*\kappa = \beta_*\lambda = \mu$ and (M,μ) is nontrivial (i.e., $H(M,\mu) > 0$). Then the two Bernoulli shifts $(2^G_e \otimes K, \eta \otimes \kappa)$ and $(2^G_e \otimes L, \eta \otimes \lambda)$ are isomorphic.*

PROOF. The proof is similar to the proof of Lemma 3.1. Let $\tilde{\alpha} : 2^G_e \otimes K \to 2^G_e \otimes M$ be the map $\tilde{\alpha}(x)(g) := \alpha(x(g))$ for $g \in \mathrm{Dom}(x)$. Define $\tilde{\beta} : 2^G_e \otimes L \to 2^G_e \otimes M$ similarly. Observe that $\tilde{\alpha}_*(\eta \otimes \kappa) = \tilde{\beta}_*(\eta \otimes \lambda) = \eta \otimes \mu$.

By Theorem 2.2 there exists an ergodic $U \in \mathrm{Aut}(2^G_e \otimes M, \eta \otimes \mu)$ such that (i) for a.e. $x \in 2^G_e \otimes M$, the orbit $\{U^i x : i \in \mathbb{Z}\}$ is infinite, and (ii) for a.e. $x \in 2^G_e \otimes M$, there exists $g \in G$ such that $Ux = gx$.

Define $T : 2^G_e \otimes K \to 2^G_e \otimes K$ by $T(x) = gx$ where $g \in G$ is such that $U(\tilde{\alpha}(x)) = g\tilde{\alpha}(x)$. Similarly, let $S : 2^G_e \otimes L \to 2^G_e \otimes L$ be defined by $S(y) = gy$ where $g \in G$ is such that $U(\tilde{\beta}(y)) = g\tilde{\beta}(y)$. Note that for a.e. $(x,y) \in 2^G_e \otimes K \times 2^G_e \otimes L$ with $\tilde{\alpha}(x) = \tilde{\beta}(y)$, there is a $g \in G$ such that $Tx = gx$ and $Sy = gy$.

Let $\chi_M : 2^G_e \otimes M \to M$ be the projection map $\chi_M(x) = x(e)$. Let $\mathcal{B}_M$ be the smallest U-invariant sigma-algebra on $2^G_e \otimes M$ for which χ_M is measurable. Define $\chi_K : K^G \to K, \chi_L : L^G \to L, \mathcal{B}_K, \mathcal{B}_L$ similarly.

Observe that $(T, 2^G_e \otimes K, \mathcal{B}_K, \eta \otimes \kappa)$ and $(S, 2^G_e \otimes L, \mathcal{B}_L, \eta \otimes \lambda)$ are Bernoulli relative to $(U, 2^G_e \otimes M, \mathcal{B}_M, \eta \otimes \mu)$. To see this, note, for example, that χ_K is a generator for $(T, 2^G_e \otimes K, \mathcal{B}_K, \eta \otimes \kappa)$ and $\{\chi_K \circ T^i : i \in \mathbb{Z}\}$ are jointly independent relative to $(\tilde{\alpha})^{-1}(\mathcal{B}_M)$. So

$$h(T|(\tilde{\alpha})^{-1}(\mathcal{B}_M)) = H(\chi_K|\ (\tilde{\alpha})^{-1}(\mathcal{B}_M)) = H(K,\kappa) - H(M,\mu).$$

Similarly, $H(L,\lambda) - H(M,\mu) = h(S|(\tilde{\beta})^{-1}(\mathcal{B}_M))$. By Theorem 2.1, there is an isomorphism $\phi : (2^G_e \otimes K, \mathcal{B}_K, \eta \otimes \kappa) \to (2^G_e \otimes L, \mathcal{B}_L, \eta \otimes \lambda)$ such that $\phi T = S\phi$ and $\tilde{\alpha} = \tilde{\beta}\phi$.

Define $\Phi : 2^G_e \otimes K \to 2^G_e \otimes L$ by $\Phi(x)(e) = \phi(g^{-1}x)(e)$ for $g \in G$, $x \in 2^G_e \otimes K$. Observe that Φ is measurable with respect to the Borel sigma-algebras of $2^G_e \otimes K$ and $2^G_e \otimes L$ (which are, in general, larger than $\mathcal{B}_K$ and $\mathcal{B}_L$). We claim this is the required isomorphism. It is clearly shift-equivariant.

To see that Φ is invertible, define $\Psi : 2^G_e \otimes L \to 2^G_e \otimes K$ by $\Psi(y)(g) = \phi^{-1}(g^{-1}y)(e)$. Because both Φ and Ψ are shift-equivariant and $\Phi\Psi(y)(e) = y(e)$, $\Psi\Phi(x)(e) = x(e)$, it follows that Ψ is the inverse of Φ. It is easy to check that $\tilde{\beta}\Phi = \tilde{\alpha}$ and $\Phi T = S\Phi$.

To finish, we must show that $\Phi_*\eta \otimes \kappa = \eta \otimes \lambda$. For $z \in 2^G_e \otimes M$, let $X_z = \{x \in 2^G_e \otimes K :\ \tilde{\alpha}(x) = z\}$ and $Y_z = \{y \in 2^G_e \otimes L :\ \tilde{\beta}(y) = z\}$. Because $\tilde{\beta}\Phi = \tilde{\alpha}$, it follows that Φ maps X_z to Y_z.

For $z \in 2^G_e \otimes M$, let ζ_z be the fiber measure of $\eta \otimes \kappa$ over z. This family of measures is determined (up to measure zero sets) by the property that ζ_z is supported on X_z and $\eta \otimes \kappa = \int \zeta_z \, d\eta \otimes \mu(z)$. Similarly, let ν_z be the fiber measure of $\eta \otimes \lambda$ over z.

Because $\tilde{\alpha}_*(\eta \otimes \kappa) = \tilde{\beta}_*(\eta \otimes \lambda) = \eta \otimes \mu$ and Φ maps X_z to Y_z, in order to show that $\Phi_*(\eta \otimes \kappa) = (\eta \otimes \lambda)$, it suffices to show that Φ restricts to an isomorphism from (X_z, ζ_z) to (Y_z, ν_z) for a.e. z.

For $g \in \mathrm{Dom}(z)$ and $n \in \mathbb{Z}$ let $\tau^n_z(g) \in G$ be the element satisfying $U^n(g^{-1}z)(e) = z(\tau^n_z(g))$. This is well-defined for a.e. $z \in 2^G_e \otimes M$. If $x \in X_z$ then $T^n(g^{-1}x)(e) = x(\tau^n_z(g))$ by definition.

Fix $z \in 2^G_e \otimes M$. Let $e = g_0, g_1, \ldots \in \mathrm{Dom}(z)$ be such that if $O_i = \{\tau^n_z(g_i) :\ n \in \mathbb{Z}\}$ then $O_i \cap O_j = \emptyset$ whenever $i \neq j$ and $\cup_{i=0}^\infty O_i = \mathrm{Dom}(z)$. So $K^{\mathrm{Dom}(z)} = \prod_{i=0}^\infty K^{O_i}$.

Let $\mathcal{B}_{K,z,0}$ be the restriction of $\mathcal{B}_K$ to X_z. This is the σ-algebra of X_z generated by $\{\chi_K \circ T^j :\ \ j \in \mathbb{Z}\}$. More generally, let $\mathcal{B}_{K,z,i}$ be the sigma-algebra on X_z generated by $\{\chi_K \circ T^j g_i^{-1} :\ j \in \mathbb{Z}\}$. Define σ-algebras $\mathcal{B}_{L,z,i}$ on Y_z similarly.

Because the O_i's are pairwise disjoint, the sigma-algebras $\mathcal{B}_{K,z,i}$ are independent. Because $\cup O_i = \mathrm{Dom}(z)$, these sigma-algebras generate the Borel sigma-algebra of K^G restricted to X_z. Similarly statements apply to L in place of K. Therefore, it suffices to show that Φ restricted to X_z determines an isomorphism from $(X_z, \mathcal{B}_{K,z,i}, \zeta_z)$ to $(Y_z, \mathcal{B}_{L,z,i}, \nu_z)$ for all i.

By definition, ϕ (and therefore Φ) restricted to X_z is an isomorphism from $(X_z, \mathcal{B}_{K,z,0}, \zeta_z)$ to $(Y_z, \mathcal{B}_{L,z,0}, \nu_z)$. Note that $(X_{g_i^{-1}z}, \mathcal{B}_{K,g_i^{-1}z,0}, \zeta_{g_i^{-1}z}) = g_i^{-1}(X_z, \mathcal{B}_{K,z,i}, \zeta_z)$. Because Φ is shift-equivariant, this implies Φ restricted to X_z determines an isomorphism from $(X_z, \mathcal{B}_{K,z,i}, \zeta_z)$ to $(Y_z, \mathcal{B}_{L,z,i}, \nu_z)$ for every i. $\square$

5. Ornstein groups

If G is a group, $C < G$ a subgroup and (K,κ) a probability space then let $(K^{G/C}, \kappa^{G/C})$ be the product space with the G-action $gx(fC) = x(g^{-1}fC)$. This is called the *generalized Bernoulli shift* over G/C with base space (K,κ).

LEMMA 5.1. *Let G be a countably infinite group and $C < G$ a finite cyclic subgroup of prime order. Let $N(C) = \{g \in G : \ gCg^{-1} = C\}$ be the normalizer of C. Suppose that (K, κ) is a nontrivial probability space. If $G/N(C)$ is infinite then the action $G \curvearrowright (K^{G/C}, \kappa^{G/C})$ is essentially free.*

PROOF. For $g \in G$, let $X_g = \{x \in K^{G/C} : \ gx = x\}$. It suffices to show that $\kappa^{G/C}(X_g) = 0$ for every $g \in G \setminus \{e\}$.

Given $g \in G$, let $I_g = \{fC \in G/C : \ gfC \neq fC\}$. It is easy to see that if $|I_g| = \infty$ then there exists a subset $I'_g \subset I_g$ with $|I'_g| = \infty$ such that for every $fC \in I'_g$, $gfC \notin I'_g$. Let $X(g, fC) = \{x \in X : \ x(gfC) = x(fC)\}$. Then the events $\{X(g, fC) : \ fC$
$in I'_g\}$ are jointly independent. So

$$\kappa^{G/C}(X_g) \leq \kappa^{G/C}\left(\bigcap_{fC \in I'_g} X(g, fC)\right) = \prod_{fC \in I'_g} \kappa^{G/C}(X(g, fC)) = 0.$$

So it suffices to show that $|I_g| = \infty$ for every $g \in G \setminus \{e\}$. (This fact was observed earlier in [**KT08**] Proposition 2.4).

If $fC \notin I_g$ then $gfC = fC$, i.e., $f^{-1}gf \in C$ which is equivalent to $g \in fCf^{-1}$. In particular, if $g \notin fCf^{-1}$ for any f, then $|I_g| = \infty$. So suppose that $g \in fCf^{-1}$ for some $f \in G$ and $g \neq e$. We claim that if $f_2C \notin I_g$ then $f_2N(C) = fN(C)$. Indeed, in this case $g \in fCf^{-1} \cap f_2Cf_2^{-1}$. Because C is a cyclic group of prime order and g is nontrivial, this implies $fCf^{-1} = f_2Cf_2^{-1}$ which implies $f_2^{-1}f \in N(C)$, i.e., $fN(C) = f_2N(C)$. Because $G/N(C)$ is infinite, I_g is infinite as required. □

If $p_1, \ldots, p_n$ are non-negative real numbers then let $H(p_1, \ldots, p_n) := -\sum_{i=1}^n p_i \log(p_i)$ where $0 \log(0) := 0$ by convention.

LEMMA 5.2. *For any real numbers t, r with $0 < t < 1$ and $H(t, 1-t) < r$, there exists a standard probability space (L, λ) with an element $l_1 \in L$ such that $\lambda(\{l_1\}) = t$ and $H(L, \lambda) = r$.*

PROOF. Let (N, ν) be a probability space with $H(t, 1-t) + (1-t)H(N, \nu) = r$. Let L be the disjoint union of $\{l_1\}$ and N. Define the measure λ on L by $\lambda(\{l_1\}) = t$ and $\lambda(B) = (1-t)\nu(B)$ for all Borel $B \subset N$. Then $H(L, \lambda) = H(t, 1-t) + (1-t)H(N, \nu) = r$. □

PROOF OF THEOREM 1.3. Let G be a countably infinite group. Suppose G contains a nontrivial element $g_0 \in G$ of finite prime order so that if $C = \langle g_0 \rangle$ then $G/N(C)$ is infinite. Let (K, κ) be a non-trivial two-atom space (e.g., $H(K, \kappa) > 0$). By Theorem 1.2, to prove G is Ornstein it suffices to prove that there is a non-two-atom space (L, λ) with $H(K, \kappa) = H(L, \lambda) > 0$ such that the Bernoulli shifts $G \curvearrowright (K^G, \kappa^G)$ and $G \curvearrowright (L^G, \lambda^G)$ are isomorphic.

Let $p > 1$ be the order of g. We claim that there is a non-two-atom space (L, λ) with $H(L, \lambda) = H(K, \kappa)$ and a standard nontrivial probability space (M, μ) and factor maps $\alpha : (K^p, \kappa^p) \to (M, \mu)$, $\beta : (L^p, \lambda^p) \to (M, \mu)$. Moreover, we require that if $\sigma_K : K^p \to K^p$ denotes the shift map $\sigma_K(x_0, \ldots, x_{p-1}) = (x_1, \ldots, x_{p-1}, x_0)$ then $\alpha\sigma_K = \alpha$. We also require that if $\sigma_L : L^p \to L^p$ is defined similarly then $\beta\sigma_L = \beta$.

To prove the claim, we may assume without loss of generality that $K = \{k_0, k_1\}$ with $\kappa(\{k_0\}) \leq \kappa(\{k_1\})$. Let $\epsilon = \kappa(\{k_0\})$. By the previous lemma, there exists a probability space (L,λ) with an element $l_1 \in L$ such that $\lambda(\{l_1\}) = (1-\epsilon^p)^{1/p}$ and $H(L,\lambda) = H(K,\kappa)$. The space (L,λ) must not be a two-atom space since otherwise $H(K,\kappa) = H(\epsilon, 1-\epsilon) = H((1-\epsilon^p)^{1/p}, 1-(1-\epsilon^p)^{1/p}) = H(L,\lambda)$ implies $1-\epsilon = (1-\epsilon^p)^{1/p}$ which contradicts $0<\epsilon<1$ and $p>1$. Let $M := \{0,1\}$, $\mu(\{0\}) := \epsilon^p$, $\mu(\{1\}) := 1-\epsilon^p$, $\alpha(x_0,\ldots,x_{p-1}) := 0$ if and only if $(x_0,\ldots,x_{p-1}) = (k_0,\ldots,k_0)$ and $\beta(x_0,\ldots,x_{p-1}) := 1$ if and only if $(x_0,\ldots,x_{p-1}) = (l_1,\ldots,l_1)$. This proves the claim.

Let $[K^p] = K^p/\sigma_K$. That is, $[K^p]$ is the set of all σ_K-orbits in K^p. Let $[\kappa^p]$ be the probability measure on $[K^p]$ obtained by pushing κ^p forward under the quotient map $K^p \to [K^p]$. Define $([L^p],[\lambda^p])$ similarly.

Let $([K^p]^{G/C},[\kappa^p]^{G/C})$ be the generalized Bernoulli shift over G/C with base space $([K^p],[\kappa^p])$. This system is a factor of $G \curvearrowright (K^G,\kappa^G)$ via the map $\pi_K : K^G \to [K^p]^{G/C}$ defined by $\pi_K(x)(gC) = [x(g), x(gg_0),\ldots,x(gg_0^{p-1})]$ which denotes the σ_K-orbit of $(x(g),x(gg_0),\ldots,x(gg_0^{p-1}))$. Similarly, define $\pi_L : (L^G,\lambda^G) \to ([L^p]^{G/C},[\lambda^p]^{G/C})$.

Let $\tilde\alpha : ([K^p]^{G/C},[\kappa^p]^{G/C}) \to (M^{G/C},\mu^{G/C})$ be the factor map $\tilde\alpha(x)(gC) = \alpha(x(gC))$. This involves a slight abuse of notation because α was defined from K^p to M instead of $[K^p]$ to M. However because $\alpha\sigma_K = \alpha$, α factors through $[K^p]$. Similarly, define $\tilde\beta : ([L^p]^{G/C},[\lambda^p]^{G/C}) \to (M^{G/C},\mu^{G/C})$.

Let $\Upsilon : M^{G/C} \to [0,1]$ be a Borel isomorphism. Let Z be the set of all $z \in M^{G/C}$ such that $\Upsilon(z) \le \Upsilon(g_0^i z)$ for all i. Because Υ is Borel, Z is a Borel set. Because the action of G on $(M^{G/C},\mu^{G/C})$ is essentially free (by Lemma 5.1), $\{Z, g_0Z,\ldots,g_0^{p-1}Z\}$ partitions $M^{G/C}$ up to a set of measure zero.

If (A,ρ) is a probability space and $B \subset A$ is Borel, then the *normalized restriction* of ρ to B is the probability measure ρ_B on B defined by

$$\rho_B(E) = \frac{\rho(E)}{\rho(B)}$$

for all Borel $E \subset B$.

Define $\omega : M^{G/C} \to 2^G$ by $\omega(z) := \{g \in G : \ g^{-1}z \in Z\}$. This map is equivariant: $\omega(fz) = f\omega(z)$. Therefore, $\omega_*\mu^{G/C}$ is an invariant measure on 2^G and its normalized restriction η to 2^G_e is a measurable subgroup. Note that $\eta = \omega_*\mu_Z^{G/C}$ where $\mu_Z^{G/C}$ is the normalized restriction of $\mu^{G/C}$ to Z.

The measure η is supported on the collection $\mathcal{C} \subset 2^G_e$ of all subsets $H \subset G$ with the property that $e \in H$ and $\{H, Hg_0,\ldots,Hg_0^{p-1}\}$ is a partition of G. In particular, η-a.e. $H \in 2^G_e$ is infinite.

Let $\mathcal{C}\otimes K^p \subset 2^G_e \otimes K^p$ be the set of all functions $x : \mathrm{Dom}(x) \to K^p$ where $\mathrm{Dom}(x) \in \mathcal{C}$. Define $\mathcal{C}\otimes L^p$ similarly. Since $\eta\otimes\kappa^p$ is supported on $\mathcal{C}\otimes K^p$, without loss of generality we may consider it as a measure on $\mathcal{C}\otimes K^p$. Similarly, $\mathcal{C}\otimes\lambda^p$ is a measure on $\mathcal{C}\otimes L^p$.

Let $X = (\tilde\alpha\pi_K)^{-1}(Z)$ and $Y = (\tilde\beta\pi_L)^{-1}(Z)$. Define $\Omega_K : X \to \mathcal{C}\otimes K^p$ by $\Omega_K(x)(g) = (x(g), x(gg_0),\ldots,x(gg_0^{p-1}))$ for $g \in \mathrm{Dom}(\Omega_K(x)) := \omega\tilde\alpha\pi_K(x) = \{g \in G : \ g^{-1}x \in X\}$. By definition of $\mathcal{C}$, $(\Omega_K)_*(\kappa^G_X) = \eta\otimes\kappa^p$ where κ^G_X is the normalized restriction of κ^G to X. Define $\Omega_L : Y \to \mathcal{C}\times L^p$ similarly.

For $x \in \mathcal{C}\otimes K^p$ and $0 \le i \le p-1$, let $x_i : \mathrm{Dom}(x) \to K$ be the projection to the i-th coordinate. Thus $x(g) = (x_0(g),\ldots,x_{p-1}(g))$ for $g \in \mathrm{Dom}(x)$. Of course, we define y_i similarly if $y \in \mathcal{C}\otimes L^p$.

By abuse of notation, we let $\Omega_K^{-1} : \mathcal{C} \otimes K^p \to K^G$ denote the map $\Omega_K^{-1}(x)(g) = x_i(g')$ where g', i are uniquely determined by: $g = g' g_0^i$, $0 \le i \le p-1$, $g' \in \mathrm{Dom}(x)$. Note that $\Omega_K^{-1}\Omega_K : X \to X$ is the identity map (but $\Omega_K \Omega_K^{-1}$ is not necessarily well-defined on the domain of Ω_K^{-1}, so Ω_K^{-1} is only a right-inverse). Define Ω_L^{-1} similarly.

By Lemma 4.1 there is an isomorphism $\phi : (\mathcal{C} \otimes K^p, \eta \otimes \kappa^p) \to (\mathcal{C} \otimes L^p, \eta \otimes \lambda^p)$.

Define $\Phi : K^G \to L^G$ as follows: if $x \in X$ then $\Phi(x) = \Omega_L^{-1}\phi\Omega_K(x)$. For $0 \le i \le p-1$, we define $\Phi(g_0^i x) = g_0^i \Phi(x)$. Because $X, g_0 X, \ldots, g_0^{p-1}X$ partitions K^G (up to a set of measure zero), this defines Φ on a full measure subset of K^G (which is all that we require). We claim that this is the desired isomorphism.

First we show that Φ is equivariant. If $x \in X, g \in G$ and $gx \in X$ then $\Phi(gx) = g\Phi(x)$ because Ω_K, ϕ and Ω_L^{-1} are all equivariant (on their domains). So for any i, j,

$$\Phi((g_0^j g g_0^{-i}) g_0^i x) = \Phi(g_0^j g x) = g_0^j g \Phi(x) = (g_0^j g g_0^{-i})\Phi(g_0^i x).$$

This implies Φ is equivariant because a.e. element $y \in K^G$ can be written as $y = g_0^i x$ (for a unique $0 \le i \le p-1$ and $x \in X$) and an arbitrary element $f \in G$ can be written as $f = g_0^j g g_0^{-i}$ for some $g \in G$ with $gx \in X$. Indeed, we let j be determined by the property that $g_0^{-j} f g_0^i x \in X$ then define $g = g_0^{-j} f g_0^i$.

Next we show that Φ is invertible. For this define $\Psi : L^G \to K^G$ as follows: if $y \in Y$ then $\Psi(x) = \Omega_K^{-1}\phi^{-1}\Omega_L(x)$. For $0 \le i \le p-1$, we define $\Psi(g_0^i y) = g_0^i \Psi(y)$. Because $Y, g_0 Y, \ldots, g_0^{p-1} Y$ partitions L^G (up to a set of measure zero), this defines Ψ on a full measure subset of L^G (which is all that we require). By an argument similar to the one above, Ψ is equivariant.

If $x \in X$ then $\Psi\Phi x = \Omega_K^{-1}\phi^{-1}\Omega_L\Omega_L^{-1}\phi\Omega_K(x) = x$, i.e., $\Psi\Phi$ restricts to the identity map on X. Because Φ and Ψ are equivariant, $\Psi\Phi$ restricts to the identity map on $g_0^i X$ for every $0 \le i \le p-1$. Because $X, g_0 X, \ldots, g_0^{p-1} X$ partitions K^G (up to a set of measure zero), this implies $\Psi\Phi$ is the identity map on K^G. Similarly, $\Phi\Psi$ is the identity map on L^G.

Finally, we claim that $\Phi_*\kappa^G = \lambda^G$. Because Φ is equivariant it suffices to prove that Φ restricted to X pushes κ_X^G forward to λ_Y^G. This is true because $(\Omega_K)_*\kappa_X^G = \eta \otimes \kappa^p$, $\phi_*(\eta \otimes \kappa^p) = \eta \otimes \lambda^p$ and $(\Omega_L^{-1})_*(\eta \otimes \lambda^p) = \lambda_Y^G$. □

References

[Ba05] K. Ball. *Factors of independent and identically distributed processes with non-amenable group actions.* Ergodic Theory Dynam. Systems 25 (2005), no. 3, 711–730. MR2142942 (2006d:37012)

[Bo10] L. Bowen. *Measure conjugacy invariants for actions of countable sofic groups.* J. Amer. Math. Soc. 23 (2010), 217–245. MR2552252 (2011b:37010)

[Bo11] L. Bowen. *Weak isomorphisms between Bernoulli shifts.* To appear in Israel Journal of Mathematics. MR2811154

[DP02] A. I. Danilenko and K. K. Park. *Generators and Bernoullian factors for amenable actions and cocycles on their orbits.* Ergodic Theory Dynam. Systems 22 (2002), no. 6, 1715–1745. MR1944401 (2004f:37006)

[Ke10] A. S. Kechris. *Global aspects of ergodic group actions.* Mathematical Surveys and Monographs, 160. American Mathematical Society, Providence, RI, 2010. xii+237 pp. MR2583950 (2011b:37003)

[KL1] D. Kerr and H. Li. *Entropy and the variational principle for actions of sofic groups.* arXiv:1005.0399

[KL2] D. Kerr and H. Li. *Bernoulli actions and infinite entropy.* arXiv:1005.5143 MR2813530

[Ko58] A. N. Kolmogorov. *A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces*. Dokl. Akad. Nauk SSSR (N.S.) 119 1958 861–864. MR0103254 (21:2035a)

[Ko59] A. N. Kolmogorov. *Entropy per unit time as a metric invariant of automorphisms*. Dokl. Akad. Nauk SSSR 124 1959 754–755. MR0103255 (21:2035b)

[KS79] M. Keane and M. Smorodinsky. *Bernoulli schemes of the same entropy are finitarily isomorphic*. Ann. of Math. (2) 109 (1979), no. 2, 397–406. MR528969 (80f:28024)

[KT08] A. Kechris and T. Tsankov. *Amenable actions and almost invariant sets*. Proc. Amer. Math. Soc. 136(2) (2008), 687–697. MR2358510 (2009a:28035)

[Me93] R. Mercer. *The full group of a countable measurable equivalence relation*. Proc. Amer. Math. Soc. 117 (1993), no. 2, 323–333. MR1139480 (93d:28029)

[Ol91] A. Yu. Ol'shanskii. *Geometry of defining relations in groups*. Translated from the 1989 Russian original by Yu. A. Bakhturin. Mathematics and its Applications (Soviet Series), 70. Kluwer Academic Publishers Group, Dordrecht, 1991. xxvi+505 pp. MR1191619 (93g:20071)

[Or70a] D. Ornstein. *Bernoulli shifts with the same entropy are isomorphic*. Advances in Math. 4 (1970) 337–352. MR0257322 (41:1973)

[Or70b] D. Ornstein. *Two Bernoulli shifts with infinite entropy are isomorphic*. Advances in Math. 5 (1970) 339–348. MR0274716 (43:478a)

[Os11] D. Osin, private communication.

[OW87] D. Ornstein and B. Weiss. *Entropy and isomorphism theorems for actions of amenable groups*. J. Analyse Math. 48 (1987), 1–141. MR910005 (88j:28014)

[RW00] D. J. Rudolph and B. Weiss. *Entropy and mixing for amenable group actions*. Ann. of Math. (2) 151 (2000), no. 3, 1119–1150. MR1779565 (2001g:37001)

[Si59] Ya. G. Sinaĭ. *On the concept of entropy for a dynamic system*. Dokl. Akad. Nauk SSSR 124 (1959) 768–771. MR0103256 (21:2036a)

[Si62] Ya. G. Sinaĭ. *A weak isomorphism of transformations with invariant measure*. Dokl. Akad. Nauk SSSR 147 (1962) 797–800. MR0161960 (28:5164a)

[St75] A. M. Stepin. *Bernoulli shifts on groups*. Dokl. Akad. Nauk SSSR 223 (1975), no. 2, 300–302. MR0409769 (53:13521)

[St76] A. M. Stepin. *Bernoulli shifts on groups and decreasing sequences of partitions*. Proceedings of the Third Japan-USSR Symposium on Probability Theory (Tashkent, 1975), pp. 592–603. Lecture Notes in Math., Vol. 550, Springer, Berlin, 1976. MR0492185 (58:11331)

[Th75] J-P. Thouvenot. *Quelques propriétés des systèmes dynamiques qui se décomposent en un produit de deux systèmes dont l'un est un schéma de Bernoulli*. Conference on Ergodic Theory and Topological Dynamics (Kibbutz, Lavi, 1974). Israel J. Math. 21 (1975), no. 2-3, 177–207. MR0399419 (53:3263)

Department of Mathematics, Texas A&M University, College Station, Texas 77843
E-mail address: `lpbowen@math.tamu.edu`

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11236

Fair dice-like hyperbolic systems

Leonid A. Bunimovich

Abstract. We introduce Fair Dice-Like (FDL) hyperbolic systems which form a rather narrow subclass on the class of hyperbolic dynamical systems. Some classical examples, e.g., the tent, von Neumann-Ulam and baker's maps are homogeneously hyperbolic.

A reason to consider this class is that for FDL hyperbolic systems one can make finite time predictions of their dynamics. This should be contrasted with a standard approach (in case when no small parameters are involved) to study dynamical systems (and random processes) where only asymptotic in time properties, like a rate of convergence to a stationary distribution, decay of correlations, etc., are addressed.

It is likely that analogous finite time dynamical properties can be studied for broader than FDL hyperbolic systems classes of dynamical systems.

1. Introduction

A main reason for dynamical systems theory to develop was the understanding that generally one cannot solve analytically differential equations and obtain formulae $x_t = f(x_0, t)$ which describe time evolution (dynamics) of a system, where x_0 is the initial state (vector of generalized coordinates) of a system and x_t is its state at a moment t. Certainly such formulae (if available) allow to compute the state of the system at any moment of time. Therefore a focus of research has been shifted to an asymptotic analysis of dynamics when time $t \to \infty$. Typical questions in this approach are what are the asymptotic states, their stability, rates of convergence to the stable states, attractors, etc.

While dealing with dynamical systems with stochastic (very irregular, chaotic) behavior one is interested in invariant measures (stationary distributions), their (stochastic) stability, rates of convergence to stationary (or equilibrium) distributions, decay of correlations, etc. Again, all these questions assume that the limit $t \to \infty$ is taken.

It was shown recently though [**BY**] that one can answer some important and interesting questions about a finite time dynamics. As in [**BY**] we will be concerned here with the first hitting times for various subsets of the phase space. The first hitting time is the primary characteristics of open dynamical systems.

2010 *Mathematics Subject Classification.* 28A80, 37D35.

Key words and phrases. Markov partition, hyperbolicity, uniform distribution, first hitting time, open dynamical systems, correlations of words.

Partially supported by NSF grant DMS-0900945.

Let $T : M \to M$ be an ergodic dynamical system with invariant probability (normalized) measure μ. Consider a measurable subset $H \subset M$, $\mu(H) > 0$, which will be called a hole. One gets an open dynamical system by assuming that an orbit escapes (disappears) when it hits the hole H. A standard and natural approach to study open dynamical systems is to use properties of the dynamics of the closed system $T : M \to M$ from which an open system was built [**PY, DY**].

The author suggested a new approach to study closed dynamical systems. The idea of this approach is that by opening up different holes H in the phase space M of a closed dynamical system and by analyzing the corresponding open systems one can get new information about the dynamics of the closed system of interest. This idea was inspired by the experiments with atomic billiards [**MHCR, FKCD**]. Therefore naturally it was first applied to billiard systems with regular [**BD1**] and with chaotic behavior [**BD2**].

Then several general and often counterintuitive results were obtained in [**BY**]. These results deal with holes of a finite size as well as with infinitesimally small holes. The case of infinitesimally small holes falls into the standard approach where one considers the limit when time tends to infinity. Even for this standard approach of the open dynamical systems theory it was possible to derive a new formula [**BY**] which relates the size of a (very small) hole to the escape rate through this hole. Naturally the experts in the open systems theory proved this formula for more general classes of hyperbolic dynamical systems [**KL, FP**]. In the present paper we will formulate this result in a slightly corrected form. Namely, this formula in fact holds almost everywhere rather than everywhere.

The main aim of the present paper though is to introduce a new class of hyperbolic dynamical systems which we call Fair Dice-Like (FDL) hyperbolic systems and to prove for these systems an analog of the main theorem in [**BY**].

A measure preserving dynamical system is called FDL hyperbolic if it has such a Markov partition ξ that all its elements have the same measures and all entries of the transition probabilities matrix are equal to each other. Thus, in a sense, one can think of FDL hyperbolic systems as of an analog of independent identically distributed (iid) random variables with the uniform distribution. Among the examples of FDL hyperbolic dynamical systems are e.g., tent map, von Neumann-Ulam map, baker's map, and rational maps of degree $d \geq 2$.

Consider a refinement ξ_m of the Markov partition ξ. All elements of such refinement also have the same measures. Let the number of elements of ξ_m equals k. Consider now k open dynamical systems built from the initial closed system by making the elements of ξ_m (one at a time) holes H_i, $1 \leq i \leq k$, of the corresponding open systems. Because any FDL hyperbolic system is ergodic the survival probabilities $P_n(H_i)$ for any one of these open systems tend to zero as time $n \to \infty$.

Our main result says that the question through which hole the escape is faster can be answered. The answer is completely determined by the positions of periodic orbits in the phase space M. Namely, consider periodic points of minimal period in the holes H_i and H_j $(i \neq j)$. Suppose that the minimal period in H_i is less than in H_j. Then, starting with some finite moment of time n^* survival probabilities for the hole H_j will always (for all $n > n^*$) be smaller than survival probabilities for the hole H_i. If the minimal periods in H_i and H_j are equal then one should compare the next (minimal) periods, etc.

Therefore for FDL hyperbolic dynamical systems a character of dynamics is essentially determined by the distribution of periodic orbits in the phase space. It is worthwhile to mention that this result is rather counterintuitive because one would rather expect a faster escape near periodic orbits of small periods. In fact, it is all the way around. It is also very surprising to obtain such results for the first hitting time which is a random variable that behaves rather irregularly.

A proof of this statement is based on technique developed in the theory of combinatorics on words [**GO1, GO2, E, M**] and is similar to the one in [**BY**]. Therefore, it does not demonstrate a dynamical mechanism behind it. A character of such dynamical mechanism was revealed in [**AB**].

Interestingly the analogous result occurs to be new even for Markov chains in the theory of random processes [**BB**]. These results provide a promise that some important questions on a finite time dynamics could be rigorously answered.

2. Definitions and main results

Let $T : M \to M$ be a hyperbolic dynamical system preserving a Borel probability measure μ.

Definition 1. A hyperbolic dynamical system $T : M \to M$ is called fair dice-like (FDL) hyperbolic if there exists a finite Markov partition ξ of the phase space M such that for any integer $m \geq 1$ and any j_i, $1 \leq j_i \leq r$, $\mu(C_\xi^{(j_0)} \cap T^{-1}C_\xi^{(j_i)} \cap \cdots \cap T^{-m+1}C_\xi^{(j_{m-1})}) = 1/r^m$, where r is a number of elements on the partition ξ.

It is clear from this definition that a symbolic representation of a FDL hyperbolic system is a full shift with a Bernoulli (product) measure and the uniform distribution of symbols. Therefore, FDL hyperbolic systems are, in particular, exponentially mixing.

Observe that in the definition of FDL hyperbolic dynamical systems there are no conditions on invariant measure μ (besides that it is a probabilistic one). However, for hyperbolic dynamical systems a natural class of measures to consider are the ones which are consistent with the topology and hyperbolic structure on M. Therefore although we assume throughout that μ is only a Borel probability measure, in all interesting examples (see some below) μ is absolutely continous with respect to Lebesgue measure or μ is absolutely continuous with respect to Lebesgue measure on unstable manifolds. We consider several examples of FDL hyperbolic systems.

1. Tent map. In this case $M = [0, 1]$ with the Lebesgue measure

$$Tx = \begin{cases} 2x, & \text{if } 0 \leq x \leq 1/2 \\ 2 - 2x, & \text{if } 1/2 \leq x \leq 1 \end{cases}.$$

Markov partition ξ consists of two elements $[0, 1/2]$ and $[1/2, 1]$.

2. von Neumann-Ulam map. $Tx = 4x(1 - x) \pmod 1$. The phase space and the Markov partition here are the same as in the first example. The invariant measure μ is absolutely continuous with the density $f(x) = 1/\pi[x(1 - x)]^{1/2}$.

3. baker's map. The phase space M is the unit square.

$$T(x,y) = \begin{cases} \left(\frac{1}{2}x, 2y\right) & \text{if } 0 \le y \le \frac{1}{2} \\ \left(\frac{1}{2}(x+1), 2y-1\right), & \text{if } \frac{1}{2} \le y \le 1 \end{cases}.$$

The Lebesgue measure is invariant and one gets a Markov partition ξ with required properties by taking the elements $C^{(1)} = \{(x,y) : 0 \le x \le 1, 0 \le y \le 1/2\}$ and $C^{(2)} = \{(x,y) : 0 \le x \le 1, 1/2 \le y \le 1\}$.

4. Hyperbolic Julia sets. Let $\widehat{\mathbb{C}}$ be the Riemann sphere and T a rational map of degree $d \ge 2$. Then M is the Julia set of $\widehat{\mathbb{C}}$, i.e., the closure of the repelling periodic points of T, and μ is the non-atomic SRB measure. It is known [**L**] that any hyperbolic rational map restricted to the Julia set is a quotient of a full shift. One satisfies the homogeneity condition by taking a Markov partition of M into d^m, $m \ge 1$, elements of the same measure.

Let $\mathcal{B}$ be the Borel σ-algebra of M with respect to the measure μ. For $n \ge 0$ and $H \in \mathcal{B}$ we define the following measurable sets.

$$\Omega_n(H) = \left\{x \in M : \exists\, j \in \mathbb{N}, 0 \le j \le n, T^j x \in H\right\} = \bigcup_{i=0}^{n} T^{-i}(H)$$
$$\Theta_n(H) = \left\{x \in M : T^n x \in H, T^j x \notin H, j = 0, \dots, n-1\right\},$$

where $T^{-i}H$ is a complete preimage of H under T^i.

It is easy to see that $\Omega_n(H)$ consists of all points whose orbits enter H after no more than n iterates and $\Omega_n(H) = \bigcup_{i=0}^{n} \Theta_i(H)$.

Denote $M_H = M \backslash H$. An open dynamical system with a "hole" H is the map $T_H : M_H \to M$ where $T_H := T|_{M\backslash H}$ is the restriction of T to M_H.

Definition 2. A survival probability $P_n(H)$ for a hole H and invariant measure μ equals $P_n(H) = \mu(M\backslash\Omega_n(H)) = 1 - \mu(\Omega_n(H))$.

Thus $P_n(H)$ is the measure of the set that does not escape into the hole by time n (i.e., in n iterations).

A central notion in the theory of open dynamical systems [**PY, DY**] is the escape rate which represents the average rate at which orbits (or points of the phase space M) enter the hole.

Definition 3. The escape rate $\rho(H)$ into the hole H is

$$\rho(H) = -\lim_{n\to\infty} \frac{1}{n} \ln P_n(H) \tag{1}$$

if this limit exists

Observe that the escape rate is an asymptotic in time ($n \to \infty$) characteristic of an open dynamical system, while the survival probability $P_n(H)$ is well defined for any finite moment of time n for any open dynamical system. Therefore $P_n(H)$ describes a finite time dynamics whereas $\rho(H)$ does not. Clearly $P_n(H) \underset{n\to\infty}{\longrightarrow} 0$ if $\sum_{s=0}^{\infty} \mu(\Theta_s(H)) = 1$.

We assume from now on that $T : M \to M$ is a FDL hyperbolic system. Let ξ_0 be a corresponding Markov partition (see Definition 1). Then for any $m > 0$ the partition $\xi_m = \xi_0 \vee T^{-1}\xi_0 \vee \cdots \vee T^{-m+1}\xi_0$ is a Markov partition of M into r^m elements $C_i(\xi_m)$, $i = 1, 2, \ldots, r^m$, of the same measure.

Consider now r^m open dynamical systems

$$T_i : M \backslash C_i \to M \quad i = 1, 2, \ldots, r^m \tag{2}$$

Observe that each of these holes $H_i = C_i(\xi_m)$ is an element of a Markov partition and therefore we will call them Markov holes. In what follows we consider closures of Markov holes without changing their names and notations. Denote for any $n \geq 0$

$$\Omega_n(C_i(\xi_m)) = \Omega_n(H_i) = \{x \in M : \exists\, k, 0 \leq k \leq n, T^k x \in H_i\}. \tag{3}$$

For each Markov hole H_i, $i = 1, 2, \ldots, r^m$, define an infinite sequence of positive integers $q_i^{(1)} < q_i^{(2)} < \cdots < q_i^{(n)} < \cdots$ according to the following procedure.

(α_1) Find a periodic point $x_1 \in H_i$ of minimal period $q_i^{(1)}$, i.e., $T^{q_i^{(1)}} x_1 = x_1$ such that there exists no other periodic point in H_i with period smaller than $q_i^{(1)}$. (Generally H_i can have several periodic points of the same period which belong to different periodic orbits.)

(α_2) Find a periodic point $x_2 \in H_i$ such that $T^{q_i^{(2)}} x_2 = x_2$ and there exist no other periodic point in H_i with a period greater than $q_i^{(1)}$ and smaller than $q_i^{(2)}$.

(α_3) Continue the same procedure and for any $q_i^{(j)}$, $j \geq 1$, find a periodic point $x_{j+1} \in H_i$ such that $T^{q_i^{(j+1)}} x_{j+1} = x_{j+1}$ and there exist no periodic point $\widehat{x} \in H_i$ such that $T^{\widehat{q}_i}\widehat{x} = \widehat{x}$ and $q_i^{(j)} < \widehat{q}_i < q_i^{(j+1)}$. Therefore $q_i = \{q_i^{(1)}, q_i^{(2)}, \ldots, q_i^{(n)}, \ldots\}$ is the sequence of all periods corresponding to all periodic points which belong to H_i.

Theorem 1 (Main Theorem). *Let $T : M \to M$ be a FDL hyperbolic dynamical system. Take two Markov holes H_1 and H_2 which are elements of the same Markov partition ξ_m, $m \geq 0$, and consider two corresponding open dynamical systems $T_1 : M_{H_1} \to M$ and $T_2 : M_{H_2} \to M$. Let q_1 and q_2 be the sequences of periods corresponding to H_1 and H_2. Then the following inequality holds for survival probabilities*

$$P_n(H_1) > P_n(H_2) \tag{4}$$

for all $n \geq n^*(H_1, H_2)$, where $n^*(H_1, H_2) = n^*$ is the minimal integer such that

$$q_1^{(n^*)} < q_2^{(n^*)} \tag{5}$$

and for all $p < n^*$ corresponding periods are equal, i.e., $q_1^{(p)} = q_2^{(p)}$.

Remark 1. Observe that the statement of the Main Theorem is rather counterintuitive. In fact, one would expect that the fastest escape from the phase space occurs near periodic orbits of small periods. A result though is that, to the contrary in vicinity of "short" periodic orbits escape is the slowest.

Remark 2. In the theory of dynamical systems we are used to the statements about the asymptotic properties (when time tends to infinity) or about some (time) averages of some observables (functions on the phase space). The Main Theorem states that the distribution functions (not the averages!) of the first hitting time for the subsets of the phase space ("holes") H_1 and H_2 never change an order (in value) after some finite and well defined moment $n^*(H_1, H_2)$. Another way to express that is the absence of intersections of (the "curves" of) survival probabilities $P_n(H_1)$ and $P_n(H_2)$ when $n > n^*(H_1, H_2)$.

REMARK 3. The Main Theorem points to an important (and not noticed before) role of periodic orbits in transport in the phase spaces of dynamical systems with chaotic behavior. Although the set of periodic points in such systems has μ-measure zero they can make seemingly homogeneous in phase space dynamics to be in fact quite intermittent one. Therefore truly equivalent in the spirit of the Main Theorem are such dynamical systems which are topologically conjugate and metrically conjugate via the same conjugation map (as the tent and von Neumann-Ulam maps). Indeed, just a metrical conjugation "does not see" periodic points.

A proof of the Main Theorem follows the same scheme as in [**BY**]. It is essentially based on the results in the combinatorics on words [**GO1, GO2, E, M**]. In fact the properties (see the Definition 1) of FDL hyperbolic dynamical systems allow to reduce the analysis just to a counting of a number of sequences which avoid some fixed words (which correspond to the "holes").

Let $W(r)$ be a finite alphabet (collection of different symbols) of size r. A word w is a sequence of symbols from $W(r)$ of a finite or infinite length, $w = \{w_i\}_{i=1}^k$, $w_i \in W(r)$. Denote the length of a word w by $|w|$.

Consider now all one-sided infinite words which form the set $\Lambda^+_{W(r)} = \{\{w_i\}_{i=1}^\infty : w_i \in W(r)\}$. For any word $\widehat{w} = \{\widehat{w}_i\}_{i=1}^k$ of finite length $|\widehat{w}| = k$ and a positive integer n we define a cylinder set $C_{\widehat{w}}(n) \subset \Lambda^+_{W(r)}$ where

$$C_{\widehat{w}}(n) = \{w \in \Lambda^+_{W(r)} : w_{n+i-1} = \widehat{w}_i, i = 1, 2, \ldots, |\widehat{w}|\}.$$

Let λ be a Bernoulli measure with equal $(1/r)$ probabilities on the collection of cylinder sets. The measure λ can be extended to the σ-algebra generated by the cylinders. The resulting measure of the cylinder $C_w(1)$ equals $\lambda(C_w(1)) = r^{-|w|}$.

As usually the shift map of $\Lambda^+_{W(r)}$ into itself is defined as $(\sigma(w))_i = w_{i+1}$. It is easy to see that σ preserves the Bernoulli measure, and therefore the triplet $(\Lambda^+_{W(r)}, \sigma, \lambda)$ is a measure-preserving dynamical system which serves as symbolic representation of FDL hyperbolic dynamical systems. The periodic points of $T : M \to M$ correspond to infinite periodic words in the symbolic space.

Our major tool to prove the Main Theorem will be the autocorrelation function of a word introduced by J. Conway (see [**GO2**]).

DEFINITION 4. Let w be a finite word and $|w| = k$. A correlation function $\mathrm{cor}(w)$ is the binary sequence $[b_1 b_2 \cdots b_k]$ which is determined in the following way. Place a copy of the word w under the original and shift it to the right by ℓ digits (positions). If the overlapping parts match, then $b_{\ell+1} = 1$. Otherwise $b_\ell = 0$. In particular, we always have $b_1 = 1$, since the word matches itself.

With some abuse of notations we let $\mathrm{cor}(w)$ denote both the sequence and its value as a binary number. For example, if $W = \{A, C, G, T\}$, that is, the DNA alphabet, and $w = AGCATAGCA$, then $\mathrm{cor}(w) = (100001001)$ and if it is viewed as a binary number, then $\mathrm{cor}(w) = 2^8 + 2^3 + 2^0$.

It is easy to see that $\mathrm{cor}(w)$, among other things, describes periodicities in the words. In fact, consider a cylinder $C_w(1)$ generated by the word w and suppose that $C_w(1)$ contains a periodic point $\widehat{w} = \{\widehat{w}_i\}_{i=1}^\infty \in C_w(1)$ of period p, $p < |w|$. Then in $\mathrm{cor}(w) = [b_1 \cdots b_k]$ one has $b_{p+1} = 1$.

Clearly, in order to compute survival probability $P_n(H)$, where a hole H corresponds in the symbolic representation to the word w, one just needs to compute the number of words of length (size) $n > |w|$ which avoid w, i.e., do not contain w as a subword. Denote by $a_n(w)$ the number of words of length n avoiding a given (finite) word w, $|w| < n$.

The following result (Eriksson conjecture [**E**]) was proven in [**M**].

LEMMA 1 (Main Lemma). *Let w and w' be the words of length k. If $cor(w) > cor(w')$ then $a_w(n) > a_{w'}(n)$ from the first n where equality no longer holds.*

From this result easily follows

LEMMA 2. *Suppose that w and u are two words of the same length with $cor(w) > cor(u)$. Let $cor(w) = [b_1 b_2 \cdots b_k]$, $cor(u) = [a_1 a_2 \cdots a_k]$ and $C_w(1)$ and $C_u(1)$ be two cylinder sets generated by these two words. Then for all $n \geq \min\{i : a_i \neq b_i\} - 1$ one has $\lambda(\Omega_n(C_w(1))) < \lambda(\Omega_n(C_u(1)))$, where the sets $\Omega_n(\cdot)$ are constructed with respect to the measure λ and the shift transformation σ of the space of one-sided sequences $\Lambda^+_{W(r)}$.*

The Main Theorem immediately follows from the definition of the correlation on words, Main Lemma and Lemma 2. Indeed, in the formulation of the Main Theorem

$$P_n(H_1) = 1 - \mu(\Omega_n(H_1)) = 1 - \lambda(\Omega_n(C_w(1)))$$

and

$$P_n(H_2) = 1 - \mu(\Omega_n(H_2)) = 1 - \lambda(\Omega_n(C_u(1)))$$

where the words w and u define the Markov holes H_1 and H_2 respectively. Therefore $P_n(H_1) > P_n(H_2)$ for all $n \geq n^*(H_1, H_2)$.

COROLLARY 1. *$\rho(H_1) < \rho(H_2)$.*

The Main Theorem shows that the escape slows down in the vicinity of periodic orbits with small periods. It would be interesting to get a *quantitative* result which shows how strongly a periodic orbit slows down the process of escape.

Such formula was first proven in [**BY**] for a more narrow class of dynamical systems than the FDL hyperbolic systems.

However, its proof for the entire class of FDL hyperbolic systems goes without any changes. The same formula was proven in [**KL**] and [**FP**] for larger classes of dynamical systems. To simplify exposition we assume that T is an one-dimensional map.

THEOREM 2 (Theorem on Local Escape). *Let $T : M \to M$ be a FDL hyperbolic dynamical system and assume that T is continuously differentiable almost everywhere and has left and right derivatives in all points. For a point $x \in M$ consider a sequence of the sets $C_n(x) \ni x$ which are the elements of Markov partitions ξ_n, $n \to \infty$.*

The following relations hold.

(a) If x is a periodic point of period p then

$$\lim_{n\to\infty} \frac{\rho(C_n(x))}{\mu(C_n(x))} = 1 - \frac{1}{\prod_{k=0}^{p-1} |T'(T^k x)|}$$

(b) For almost all non-periodic points x

$$\lim_{n\to\infty} \frac{\rho(C_n(x))}{\mu(C_n(x))} = 1$$

Observe that the Theorem on Local Escape is the asymptotic (in time) statement because it deals with escape rates defined under the condition that time $n \to \infty$. Therefore this statement per se does not tell us anything about a finite time behavior of a dynamical system. However, as an accompanying statement to the Main Theorem it is useful.

One can prove the analogous result for any sequence of shrinking subsets $A_n(x) \in M$ such that $A_n(x) \supset A_{n+1}(x)$ and $\bigcap_{k=1}^{\infty} A_k(x) = x$ (see [**BY, KL, FP**]). Observe that $A_n(x)$ are not required to be elements of Markov partitions.

It should be noted that in [**KL**] (the relation (3.4)) and [**FP**] (Theorem 1.1) this statement is formulated for all nonperiodic points x while in [**BY**] it was claimed just for almost all nonperiodic points x. The last seems to be correct. Indeed, take two finite words $\widehat{w}_1$ and $\widehat{w}_2$ which correspond to some periodic orbits of the map T. Now consider the infinite word $\widehat{w} = (\underbrace{\widehat{w}_1 \cdots \widehat{w}_1}_{3\text{ times}} \underbrace{\widehat{w}_2 \cdots \widehat{w}_2}_{3^2\text{ times}} \cdots \underbrace{\widehat{w}_1 \cdots \widehat{w}_1}_{3^n\text{ times}} \underbrace{\widehat{w}_2 \cdots \widehat{w}_2}_{3^{n+1}\text{ times}} \cdots)$. Then for the word $\widehat{w}$ (and therefore for the point $x \in M$ which correspond to $\widehat{w}$) the Theorem on Local Escape does not hold because the quantity $\frac{\rho(C_n(x))}{\mu(C_n(x))}$ oscillates between the limiting value for the periodic point (sequence) built by repeating $\widehat{w}_1$ and the periodic point built by repeating $\widehat{w}_2$. We assume here that $\prod_{k=0}^{p_j-1} |T'(T^k x)|$, $j = 1, 2$, are different for $\widehat{w}_1$ and $\widehat{w}_2$, where p_1 and p_2 are the corresponding periods.

Concluding remarks

We introduced the class of FDL hyperbolic systems. It has a remarkable property that some important features of a finite time dynamics of these systems could be predicted. Thus, the studies of FDL hyperbolic systems open up an opportunity to understand finite time dynamics of more general classes of dynamical systems.

FDL hyperbolic systems should not be confused with uniformly hyperbolic systems. In fact the distortion of von Neumann-Ulam map is infinite. On another hand, it is not clear even whether the Arnold's cat (algebraic isomorphism of the unit torus), given by the matrix $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, is a FDL hyperbolic system. Indeed, to be a FDL hyperbolic system the Arnold's cat must have a finite generating Markov partition into the elements with equal measures. We conjecture that such Markov partition for Arnold's cat does not exist.

It is also worthwhile to mention that all the proofs in the present paper (as in [**BY**]) are purely combinatorial and therefore they do not reveal a dynamical mechanism behind the results.

Such mechanism was discovered by the author and exploited in [**AB**]. Another generalization to more general classes of systems was addressed in [**BB**] where Markov chains, instead of the systems (random processes) derived from i.i.d.'s, as here and in [**BY**], were studied.

It seems that these papers open up a new direction of research in the theory of dynamical systems which addresses a finite time rather than asymptotic (in time) properties of dynamics. Indeed, our goal is to study properties of the distribution functions of the random variables of interest for dynamics rather than just their asymptotics and averages.

It is easy to give the examples, where the escape through a smaller hole is faster than through a bigger one (see [**BY**]). Thus the dynamics may play for escape more important role than the size of a hole. We conjecture that the process of escape (or the first visits) to different subsets of a phase space is completely determined by the size (measure) of a hole, the distribution of the periodic points and by distortion in a chaotic dynamical system under study.

References

[AB] V. S. Afraimovich, L. A. Bunimovich, Which hole is leaking the most: a topological approach to study open systems, Nonlinearity **23** (2010), 643–656. MR2593912 (2011e:37080)

[BB] Yu. Bakhtin, L. A. Bunimovich, The optimal sink and the best source in a Markov chain, *J. Stat. Phys.* **143** (2011), 943–954. MR2811467

[BD1] L. A. Bunimovich and C. Dettmann, Open circular billiards and the Riemann hypothesis, Phys. Rev. Lett. **94** (2005), 100201–4.

[BD2] L. A. Bunimovich and C. Dettmann, Peeping at chaos: Nondestructive monitoring of chaotic systems by measuring long-time escape rates, Europhys. Letts. **80** (2007), 40001–6. MR2354135 (2008g:82066)

[BY] L. A. Bunimovich, A. Yurchenko, Where to place a hole to achieve a maximal escape rate, Israel J. Math. **182** (2011), 229–252. MR2783972

[DY] M. Demers and L.-S. Young, Escape rates and conditionally invariant measures, Nonlinearity **19** (2006), 377–397. MR2199394 (2006i:37051)

[E] K. Eriksson, Autocorrelation and the enumeration of strings avoiding a fixed string, Combinatorics, Probability and Computing **6** (1997), 45–48. MR1436718 (98m:05005)

[FKCD] N. Friedman, A. Kaplan, D. Carasso and N. Davidson, Observation of chaotic and regular dynamics in atom-optics billiards, Phys. Rev. Lett. **86** (2001), 1518–1521.

[FP] Andrew Ferguson and Mark Pollicott, Escape rates for Gibbs measures, preprint, 2010, arXiv 1009.0086v2.

[GO1] L. J. Guibas and A. M. Odlyzko, Periods in strings, J. of Combinatorial Theory, Series A **30** (1981), 19–42. MR607037 (83h:68102)

[GO2] L. J. Guibas and A. M. Odlyzko, String overlaps, pattern matching and nontransitive games, J. of Combinatorial Theory, Series A **30** (1981), 183–208. MR611250 (82g:05007)

[KH] A. Katok and B. Hasselblatt, Introduction to the modern theory of dynamical systems, Cambridge Univ. Press. 1998. MR1326374 (96c:58055)

[KL] G. Keller, C. Liverani, Rare events, escape rates and quasistationarity: some exact formulae. J. Stat. Phys. **135** (2009), 519–534. MR2535206 (2011a:37012)

[L] M. Yu. Lyubich, Dynamics of rational transformations: a topological picture, Russian Math. Surveys **41** (1986), 43–117. MR863874 (88g:58094)

[M] M. Mänsson, Pattern avoidance and overlap in strings, Combinatorics, Probability and Computing **11** (2002), 393–402. MR1918724 (2003g:68100)

[MHCR] V. Milner, J. L. Hanssen, W. C. Campbell and M. Raizen, Optical billiards for atoms, Phys. Rev. Lett. **86** (2001), 1514–1517.

[PY] G. Pianigiani and J. Yorke, Expanding maps on sets which are almost invariant: decay and chaos, Trans. of the AMS **252** (1979), 351–366. MR534126 (80i:58030)

School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332-0160

E-mail address: bunimovh@math.gatech.edu

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11235

Complexity and heights of tori

Gautam Chinta, Jay Jorgenson, and Anders Karlsson

Abstract. We prove detailed asymptotics for the number of spanning trees, called complexity, for a general class of discrete tori as the parameters tend to infinity. The proof uses in particular certain ideas and techniques from an earlier paper by the authors appearing in the *Nagoya Math. J.* **198** (2010), 121–177. Our asymptotic formula provides a link between the complexity of these graphs and the height of associated real tori, and allows us to deduce some corollaries on the complexity thanks to certain results from analytic number theory. In this way we obtain a conjectural relationship between complexity and regular sphere packings.

1. Introduction

The number of spanning trees $\tau(G)$, called *the complexity*, of a finite graph G is an invariant which is of interest in several sciences: network theory, statistical physics, theoretical chemistry, etc. Via the well-known matrix-tree theorem of Kirchoff, the complexity equals the determinant of the combinatorial Laplacian Δ_G divided by the number of vertices.

For compact Riemannian manifolds M there is an analogous invariant $h(M)$, *the height*, defined as the negative of the logarithm of the zeta-regularized determinant of the Laplace-Beltrami operator, and which is of interest for quantum physics. The analogy between the height and complexity has been commented on by Sarnak in [**S90**].

In statistical physics it is of interest to study the asymptotics of the complexity, and other spectral invariants, for certain families of graphs. Important cases to study are various subgraphs of the standard lattice $\mathbb{Z}^d$. An instance of this is to study discrete tori, corresponding to periodic boundary conditions, as the parameters tend to infinity, see [**DD88**], [**CJK10**], and references therein. It is shown in [**CJK10**] that in the asymptotics of the complexity of discrete tori, the height of an associated real torus appears as a constant.

2010 *Mathematics Subject Classification.* Primary 58J52, 05C30; Secondary 05C90, 11F55, 11M35, 58J35, 58J37.

Professor Chinta was supported by grants from the National Science Foundation and the Professional Staff Congress of the City University of New York.

Professor Jorgenson was supported by grants from the National Science Foundation and the Professional Staff Congress of the City University of New York.

Professor Karlsson was supported by SNSF grant 200021_132528/1.

In the present paper we study discrete tori of a more general type, defined as follows. Let Λ be an invertible $r \times r$ matrix with all entries being integers. This matrix defines a lattice $\Lambda\mathbb{Z}^r$ in $\mathbb{R}^r$. We associate to the group quotient with standard generators

$$\Lambda\mathbb{Z}^r\backslash\mathbb{Z}^r$$

its Cayley graph, which we call a *discrete torus.* In other words, two elements x and y in $\Lambda\mathbb{Z}^r\backslash\mathbb{Z}^r$ are *adjacent*, denoted $x \sim y$, if they differ by ± 1 in exactly one of the coordinates and equal everywhere else (everything mod $\Lambda\mathbb{Z}^r$ of course).

Let $0 = \lambda_0 < \lambda_1 \leq ... \leq \lambda_{|\det\Lambda|-1}$ be the eigenvalues of the combinatorial Laplacian $\Delta_{\Lambda\mathbb{Z}^r\backslash\mathbb{Z}^r}$ of the discrete torus – see (2.1) for a definition of $\Delta_{\Lambda\mathbb{Z}^r\backslash\mathbb{Z}^r}$. Define $\det^* \Delta_{\Lambda\mathbb{Z}^r\backslash\mathbb{Z}^r}$ to be the product of the nonzero eigenvalues of the Laplacian:

$$\det{}^*\Delta_{\Lambda\mathbb{Z}^r\backslash\mathbb{Z}^r} := \lambda_1\lambda_2...\lambda_{|\det\Lambda|-1}.$$

Note that the trivial eigenvalue is removed. We will nevertheless refer to $\det^*\Delta_{\Lambda\mathbb{Z}^r\backslash\mathbb{Z}^r}$ as the determinant of the Laplacian. For simplicity we will mostly assume that $\det\Lambda > 0$.

THEOREM 1.1. *Let $\{\Lambda_n\}$ be a sequence of $r \times r$ integer matrices. Suppose that $\det\Lambda_n \to \infty$ and $\Lambda_n/(\det\Lambda_n)^{1/r} \to A \in SL_r(\mathbb{R})$. Then as $n \to \infty$,*

$$\log\det{}^*\Delta_{\Lambda_n\mathbb{Z}^r\backslash\mathbb{Z}^r} = c_r \det\Lambda_n + \frac{2}{r}\log\det\Lambda_n + \log\det{}^*\Delta_{A\mathbb{Z}^r\backslash\mathbb{R}^r} + o(1)$$

where

$$c_r = \log 2r - \int_0^\infty e^{-2rt}(I_0(2t)^r - 1)\frac{dt}{t}.$$

The definition of $\log\det^*\Delta_{A\mathbb{Z}^r\backslash\mathbb{R}^r}$ will be recalled in section 3 below.

Our earlier paper [**CJK10**] treats the case when the Λ_n are diagonal matrices. The present paper uses several facts that are established in that paper. Thanks to the fact that the first two terms in the asymptotics are universal in the sense that they only depend on Λ_n via $\det\Lambda_n$, the theorem gives a close connection between the complexity of certain graphs and the height of an associated manifold. We emphasize that this is not at all obvious: while it is true that the appropriately rescaled eigenvalues of the discrete tori converge to the eigenvalues of $\Delta_{A\mathbb{Z}^r\backslash\mathbb{R}^r}$, this convergence is certainly not uniform. Moreover, the height cannot be defined as the product of eigenvalues, there is a regularization in the definition. An explicit expression for the height of flat tori of volume 1, $h(A\mathbb{Z}^r\backslash\mathbb{R}^r) := -\log\det\Delta_{A\mathbb{Z}^r\backslash\mathbb{R}^r}$, can be found in Theorem 2.3 of [**Ch97**]. Deninger and Lück have informed us that the constant c_r also has an interpretation as a determinant, namely the Fuglede-Kadison determinant of the Laplacian on $\mathbb{Z}^r$, see [**L**].

We turn now to a connection between our results and sphere packings. The problem of finding the densest packing of ordinary space with spheres of equal radii is an old one with practical importance even in dimensions greater than 3. One type of packings is *regular sphere packings* which means that the spheres are centered at the points of a lattice $A\mathbb{Z}^r$. Gauss showed that the face centered lattice (fcc) D_3 is optimal among regular packings in three dimensions. By work of Thue and Toth one knows that the hexagonal lattice A_2 is densest in dimension 2. In fact, Thue and Toth proved that A_2 gives the densest packing among all packings — lattice as well as non-lattice. In dimension 24 it is known that the Leech lattice is optimal among unimodular lattices. We refer to the book of Conway-Sloane [**CS99**] and the paper of Cohn-Kumar [**CK09**] for more information.

Conjecturally the height of $A\mathbb{Z}^r\backslash\mathbb{R}^r$ has a global minimum when $A\mathbb{Z}^r$ is the densest regular sphere packing. This conjecture, attributed to Sarnak, appears as Conjecture 4.5 in [**Ch97**]. Extremal metrics for heights has been studied in [**OPS88**] in dimension 2, and for tori in higher dimensions notably in [**Ch97**] and [**SS06**]. In these papers, the question is phrased as the study of the derivative of Epstein zeta functions at $s = 0$. From this theory we can deduce the following corollaries from our main theorem:

Corollary 1.2. *Let $\{\Lambda_n\}$ be a sequence of $r\times r$ integer matrices with $\det\Lambda_n \to \infty$. Suppose that $\{\Lambda_n/(\det\,\Lambda_n)^{1/r}\}$ belongs to a compact subset of $SL_r(\mathbb{R})$, $r = 2, 3$, avoiding lattices equivalent to A_2, resp. D_3. Assume that there is a sequence $\{L_n\}$ with $\det L_n = \det\Lambda_n$ such that $\{L_n/(\det\,L_n)^{1/r}\}$ converges to A_2, resp. D_3. Then $L_n\mathbb{Z}^r\backslash\mathbb{Z}^r$ has more spanning trees than $\Lambda_n\mathbb{Z}^r\backslash\mathbb{Z}^r$ for all sufficiently large n.*

Corollary 1.3. *Let Λ_n be a sequence of $r \times r$ integer matrices with $\det\Lambda_n \to \infty$ Suppose that $\{\Lambda_n/(\det\,\Lambda_n)^{1/r}\}$ stays in a compact subset of $SL_r(\mathbb{R})$. For all sufficiently large n we have that*

$$\tau(\Lambda_n\mathbb{Z}^r\backslash\mathbb{Z}^r) \le \frac{(\det\Lambda_n)^{2/r-1}}{4\pi}\exp(c_r\det\Lambda_n + \gamma + 2/r),$$

where γ is Euler's constant and c_r is as in the theorem.

In the trivial case $r = 1$, this estimate gives a value close to the truth:

$$\det\Lambda_n = \tau(\Lambda_n\mathbb{Z}^r\backslash\mathbb{Z}^r) \le 1.05\det\Lambda_n.$$

Upper bounds for the number of spanning trees have been considered in the combinatorics literature since 1970s at least. For regular graphs there is a rather sharp estimate by Chung and Yau [**CY99**] improving on an earlier result of McKay [**M**]. In general, it is an open problem to decide which simple graph on n vertices and e edges has the maximal complexity. This is of interest to communication network theory since this graph invariant appears as a measure of reliability.

It would be of interest to also go in the other direction: proving results on the extrema of families of Epstein zeta functions via a better understanding of the number of spanning trees of discrete tori.

2. Spectral preliminaries for discrete tori

Let Λ be an invertible $r \times r$ integer matrix. This matrix defines a lattice $\Lambda\mathbb{Z}^r$ in $\mathbb{R}^r$. We denote by $DT(\Lambda)$ the *discrete torus*, or Cayley graph of the quotient group $\Lambda\mathbb{Z}^r\backslash\mathbb{Z}^r$ with standard generating set: two elements x and y in $\Lambda\mathbb{Z}^r\backslash\mathbb{Z}^r$ are *adjacent*, denoted $x \sim y$, if they differ by ±1 in exactly one of the coordinates and equal everywhere else (everything mod $\Lambda\mathbb{Z}^r$ of course).

The associated (combinatorial) *Laplacian* is defined by

$$\Delta_{DT(\Lambda)}f(x) = \sum_{y \text{ s.t. } y\sim x} (f(x) - f(y)) \tag{2.1}$$

on functions $f : \Lambda\mathbb{Z}^r\backslash\mathbb{Z}^r \to \mathbb{R}$.

The *dual lattice* $\Lambda^*\mathbb{Z}^r$ is as usual all the points v in $\mathbb{R}^r$ such that $(v, x) \in \mathbb{Z}$ for all $x \in \Lambda\mathbb{Z}^r$, where $(\cdot,\cdot)$ denotes the usual scalar product. Since $\mathbb{Z}^r$ is self-dual and

$\Lambda\mathbb{Z}^r$ is a subgroup of $\mathbb{Z}^r$ it follows that $\mathbb{Z}^r$ is a subgroup of $\Lambda^*\mathbb{Z}^r$. Note that the respective indices are

$$[\mathbb{Z}^r : \Lambda\mathbb{Z}^r] = [\Lambda^* : \mathbb{Z}^r] = |\det \Lambda| .$$

PROPOSITION 2.1. *The eigenfunctions of $\Delta_{DT(\Lambda)}$ are given by*

$$f_v(x) = e^{2\pi i(x,v)},$$

for each $v \in \mathbb{Z}^r \backslash \Lambda^*\mathbb{Z}^r$, with corresponding eigenvalue given by

$$\lambda_v = 2r - 2\sum_{k=1}^{r} \cos(2\pi v_k),$$

where v_k denotes the kth coordinate of v.

PROOF. The operator $\Delta_{DT(\Lambda)}$ is a semipositive, symmetric matrix and hence we are looking for $|\det \Lambda|$ number of eigenfunctions and eigenvalues. The proof is a trivial calculation:

$$\Delta e^{2\pi i(x,v)} = 2re^{2\pi i(x,v)} - \sum_{y\sim x} e^{2\pi i(y,v)} =$$

$$= \left(2r - \sum_{z\sim 0} e^{2\pi i(z,v)}\right) e^{2\pi i(x,v)}.$$

□

The *heat kernel* $K^{\Lambda}(t,x) : \mathbb{R}_{\geq 0} \times DT(\Lambda) \to \mathbb{R}$ is the unique bounded function which satisfies

$$\left(\Delta_{DT(\Lambda)} + \frac{\partial}{\partial t}\right) K^{\Lambda}(t,x) = 0$$

$$K^{\Lambda}(0,x) = \delta_0(x),$$

where $\delta_0(x) = 1$ if $x = 0$ and 0 otherwise. The existence and uniqueness of heat kernels in a general graph setting is established in [**DM06**]. Recall from e.g. [**CJK10**] that

$$K^{\mathbb{Z}^r}(t,z) = \prod_{k=1}^{r} K^{\mathbb{Z}}(t,z_k)$$

for $z = (z_k)$ and $K^{\mathbb{Z}}(t,w) = e^{-2t}I_w(2t)$, where I_w is the I-Bessel function of order w.

We have the following *theta inversion formula* (cf. [**CJK10**]).

PROPOSITION 2.2. *The following formula holds for $x \in \Lambda\mathbb{Z}^r \backslash \mathbb{Z}^r$ and $t \in \mathbb{R}_{\geq 0}$*

$$\sum_{y\in\Lambda\mathbb{Z}^r} K^{\mathbb{Z}^r}(t, x-y) = \frac{1}{|\det \Lambda|} \sum_{\nu\in\mathbb{Z}^r\backslash\Lambda^*\mathbb{Z}^r} e^{-t\lambda_v} f_v(x).$$

In particular,

$$\theta_{\Lambda}(t) := |\det \Lambda| \sum_{y\in\Lambda\mathbb{Z}^r} e^{-2rt} I_{y_1}(2t)...I_{y_r}(2t) = \sum_{\nu\in\mathbb{Z}^r\backslash\Lambda^*\mathbb{Z}^r} e^{-t\lambda_v}$$

PROOF. Since both sides of the equation satisfy the conditions for being the heat kernel, this follows from the uniqueness of heat kernels. The second formula is the special case $x = 0$. □

3. Spectral preliminaries for continuous tori

An r-dimensional (continuous) torus is given as a quotient of $\mathbb{R}^{\mathrm{r}}$ by a lattice $A\mathbb{Z}^{\mathrm{r}}$ where $A \in GL_r(\mathbb{R})$. The metric structure and the standard (positive) Laplace-Beltrami operator $-\sum_i \partial^2/\partial x_i^2$ on $\mathbb{R}^{\mathrm{r}}$ projects to the torus. The volume is $|\det A|$. Let A^* be the matrix defining the dual lattice $A^*\mathbb{Z}^{\mathrm{r}}$, and so $A^* = \left(A^{-1}\right)^t$. The eigenfunctions of the Laplacian on the torus in question are $f_v(x) = \exp(2\pi i v^t x)$ where v are the vectors in the dual lattice. The corresponding eigenvalues are $\lambda_v = 4\pi^2 \|v\|^2$ or with a different indexing: $\lambda_m = (2\pi)^2 (A^*m)^t (A^*m)$, where m runs through $\mathbb{Z}^r$. We have the associated theta function

$$\Theta_A(t) = \sum_{m \in \mathbb{Z}^r} e^{-(2\pi)^2 (A^*m)^t (A^*m)\cdot t}.$$

The theta inversion formula, which is equivalent to the Poisson summation formula in this case, yields

$$\Theta_A(t) = \frac{1}{(4\pi t)^{r/2}} \sum_{x \in A\mathbb{Z}^r} e^{-|x|^2/4t}.$$

The associated spectral zeta function, which in this case also goes under the name of the Epstein zeta function, is defined as

$$Z_A(s) = \sum_{m \neq 0} \lambda_m^{-s} = \frac{1}{(2\pi)^{2s}} \sum_{v \neq 0} \frac{1}{\|v\|^{2s}}.$$

Classically, one can prove the meromorphic continuation of $Z_A(s)$ to all $s \in \mathbb{C}$, showing that its continuation is holomorphic at $s = 0$. From this, one defines the spectral determinant $\det^* \Delta_{A\mathbb{Z}^r \backslash \mathbb{R}^r}$ by

$$\log \det{}^* \Delta_{A\mathbb{Z}^r \backslash \mathbb{R}^r} = -Z_A'(0).$$

4. Asymptotics

Let

$$\mathcal{I}_r(s) = -\int_0^\infty \left(e^{-s^2 t} e^{-2rt} I_0(2t)^r - e^{-t}\right) \frac{dt}{t}$$

and

$$\mathcal{H}_\Lambda(s) = -\int_0^\infty \left(e^{-s^2 t}\left[\theta_\Lambda(t) - |\det \Lambda| \cdot e^{-2rt} I_0(2t)^r - 1\right] + e^{-t}\right) \frac{dt}{t}.$$

Everything in section 3 of [**CJK10**] carries over with only notational changes, even though the eigenvalues are different and the theta identity is hence somewhat different. These differences are not essentially used. In particular the first order term as $t \to 0$ in the trace of the heat kernel is still (in the present notation) $|\det \Lambda| \cdot e^{-2rt} I_0(2t)^r$ since it corresponds to the trivial eigenvalue. In particular the following extension of Theorem 3.6 in [**CJK10**] holds:

Theorem 4.1. *For any $s \in \mathbb{C}$ with $\mathrm{Re}(s^2) > 0$, we have the relation*

$$\sum_{\lambda_v \neq 0} \log(s^2 + \lambda_\nu) = |\det \Lambda| \cdot \mathcal{I}_r(s) + \mathcal{H}_\Lambda(s).$$

Letting $s \to 0$ we have the identity

$$\log\Big(\prod_{\lambda_v \neq 0} \lambda_\nu\Big) = |\det \Lambda| \cdot \mathcal{I}_r(0) + \mathcal{H}_\Lambda(0).$$

Section 4 of [**CJK10**] is an independent section on uniform bounds on I-Bessel functions. We recall the following statements, slightly adapted to the present context (keeping in mind that $I_{-y} = I_y$ for integers, and that b may now be real):

PROPOSITION 4.2. *For any $t > 0$ and $b \geq 0$, there is a constant C such that*

$$0 \leq \sqrt{b^2 t} e^{-b^2 t} I_0(b^2 t) \leq C < 1$$

Fix $t \geq 0$ and integers y, $n_0 \geq 0$. Then for all $b \geq n_0$ we have the uniform bound

$$0 \leq \sqrt{b^2 t} \cdot e^{-b^2 t} I_y(b^2 t) \leq \left(1 + \frac{y}{b n_0 t}\right)^{-n_0 y/2b}.$$

PROPOSITION 4.3. *Let $N(u)$ be a sequence of positive integers parametrized by $u \in \mathbb{Z}_+$ such that $N(u)/u \to \alpha > 0$ as $u \to \infty$. Then for any $t > 0$ and integer k, we have*

$$\lim_{u \to \infty} N(u) e^{-2u^2 t} I_{N(u)k}(2u^2 t) = \frac{\alpha}{\sqrt{4\pi t}} e^{-(\alpha k)^2/4t}.$$

From now on we fix a sequence $\{\Lambda_n\}$ of integer matrices with $0 < \det \Lambda_n \to \infty$ satisfying

$$\frac{1}{(\det \Lambda_n)^{1/r}} \Lambda_n \to A \text{ as } n \to \infty,$$

for some $A \in \mathrm{SL}_r(\mathbb{R})$. From the previous propositions we will deduce the following:

PROPOSITION 4.4. *For each fixed $t > 0$ we have the pointwise convergence*

$$\theta_{\Lambda_n}(\det(\Lambda_n)^{2/r} t) \to \theta_A(t)$$

as $n \to \infty$.

PROOF. For any $v \in \mathbb{Z}^r$ and $\Lambda \in \mathrm{GL}_r(\mathbb{R})$, let

$$\mathbf{I}_{v,\Lambda}(t) = \prod_{i=1}^{r} I_{(\Lambda v)_i}(t)$$

where $(\Lambda v)_i$ denotes the i-th component of Λv. Let $u_n = \det(\Lambda_n)^{1/r}$ and $a_i = (Av)_i$. Note that $(\Lambda v)_i / u_n \to a_i$. We have

$$\theta_{\Lambda_n}(u_n^2 t) = \sum_{v \in \mathbb{Z}^r} u_n^r e^{-2r u_n^2 t} \mathbf{I}_{v,\Lambda_n}(2u_n^2 t).$$

From Proposition 4.3 (with $k = 0$ or ± 1) we have for any $t > 0$ and $v \in \mathbb{Z}^r$ that

$$u_n^r e^{-2r u_n^2 t} \mathbf{I}_{v,\Lambda_n}(2u_n^2 t) \to \frac{1}{\left(\sqrt{4\pi t}\right)^r} e^{-a_1^2/4t} \dots e^{-a_r^2/4t}$$

as $n \to \infty$. This means that the proposition will be proved if we can interchange the limit and the infinite sum. We show that for fixed t, the sum is convergent uniformly in u_n (or equivalently, in n).

We can rewrite the sum $\theta_{\Lambda_n}(u_n^2 t)$ in $r + 1$ sums depending on how many components of the $\Lambda_n v$ are zero. Pick n_0 sufficiently large so that

$$|a_i|/2 \leq |(\Lambda_n v)_i| / u_n \leq 2|a_i|$$

for all $v \in \mathbb{Z}^r$ and $n \geq n_0$ and $a_i \neq 0$. Recall that $I_{-n} = I_n$. Let us look at a term with k zeros in the y_is and estimate with the help of Proposition 4.2:

$$\begin{aligned} u_n^r e^{-2ru_n^2 t} \mathbf{I}_{v,\Lambda_n}(2u_n^2 t) &\leq \left(\frac{1}{\sqrt{2t}}\right)^r \prod_{(\Lambda_n v)_i \neq 0} \left(1 + \frac{|(\Lambda_n v)_i|}{u_n n_0 2t}\right)^{-n_0|(\Lambda_n v)_i|/2u_n} \\ &\leq \left(\frac{1}{\sqrt{2t}}\right)^r \prod_{a_i \neq 0} \lambda^{|a_i|} \end{aligned}$$

for all n large and where

$$\lambda := \left(1 + \frac{a}{n_0 4t}\right)^{-n_0/4} < 1$$

and a is the smallest nonzero absolute value of all the entries in $A\mathbb{Z}^r$. The whole theta series is therefore bounded by $r+1$ sums of a product of convergent geometric series. This shows that the infinite sum is uniformly convergent and the proof is complete. $\square$

LEMMA 4.5. *Given a sequence* $\{\Lambda_n\}$ *satisfying* $(\det \Lambda_n)^{-1/r}\Lambda_n \to A$ *as above, there is a constant* $d > 0$ *such that for all sufficiently large* n

$$\theta_{\Lambda_n}(u_n^2 t) \leq 1 + \sum_{j=1}^{\infty} e^{-dtj},$$

for all $t > 0$.

PROOF. Let $u_n = \det(\Lambda_n)^{1/r}$. We have

$$\begin{aligned} \theta_{\Lambda_n}(t) &= \sum_{\nu \in \mathbb{Z}^r \setminus \Lambda_n^* \mathbb{Z}^r} e^{-t\lambda_v} = \sum_{\nu \in \mathbb{Z}^r \setminus \Lambda_n^* \mathbb{Z}^r} e^{-t\left(2r - 2\sum_{k=1}^r \cos(2\pi v_k)\right)} \\ &= 1 + \sum_{\substack{\nu \in \mathbb{Z}^r \setminus \Lambda_n^* \mathbb{Z}^r \\ v \neq 0}} \prod_{k=1}^r e^{-4t\sin^2(\pi v_k)}, \end{aligned}$$

so that

$$\theta_{\Lambda_n}(u_n^2 t) = 1 + \sum_{\substack{\nu \in \mathbb{Z}^r \setminus \Lambda_n^* \mathbb{Z}^r \\ v \neq 0}} \prod_{k=1}^r e^{-4tu_n^2 \sin^2(\pi v_k)} .$$

We use the elementary bounds $\sin x \geq x - x^3/6$ and $\sin(\pi - x) \geq x - x^3/6$ for $x \in [0, \pi/2]$ and get

$$\begin{aligned} u_n \sin(\pi v_k) &\geq u_n \pi v_k (1 - \pi^2 v_k^2/6) > c\pi u_n v_k \text{ if } v_k \leq 1/2 \\ u_n \sin(\pi v_k) &\geq u_n \pi v_k (1 - \pi^2 v_k^2/6) > c\pi u_n (1 - v_k) \text{ if } v_k > 1/2 \end{aligned}$$

for some positive constant c for all n sufficiently large. Note also that for every $v \neq 0$, the values $u_n v_k$ range over the integers times an entry in A as $n \to \infty$ because of the convergence of $(\det \Lambda_n)^{-1/r}\Lambda_n$ to A in $\mathrm{SL}_r(\mathbb{R})$. We then conclude there is a constant $d > 0$ such that for all sufficiently large n

$$\theta_{\Lambda_n}(u_n^2 t) \leq 1 + \sum_{j=1}^{\infty} e^{-dtj}.$$

$\square$

Now we can show:

PROPOSITION 4.6. *With the notation as above and* $u_n := \det(\Lambda_n)^{1/r}$, *we have that*

$$\int_1^\infty \left(\theta_{\Lambda_n}(u_n^2 t) - u_n^r e^{-2ru_n^2 t} I_0(2u_n^2 t)^r - 1 + e^{-u_n^2 t}\right) \frac{dt}{t}$$
$$= \int_1^\infty (\theta_A(t) - 1)\frac{dt}{t} - \frac{2}{r}(4\pi)^{-r/2} + o(1)$$

as $n \to \infty$.

PROOF. Write

$$\int_1^\infty \left(\theta_{\Lambda_n}(u_n^2 t) - u_n^r e^{-2ru_n^2 t} I_0(2u_n^2 t)^r - 1 + e^{-u_n^2 t}\right) \frac{dt}{t}$$
$$= \int_1^\infty \left(\theta_{\Lambda_n}(u_n^2 t) - 1\right) \frac{dt}{t} - \int_1^\infty u_n^r e^{-2ru_n^2 t} I_0(2u_n^2 t)^r \frac{dt}{t} + \int_1^\infty e^{-u_n^2 t}\frac{dt}{t}.$$

In the last row, the third integral clearly goes to zero as $n \to \infty$. For the first integral in the same row we have

$$\int_1^\infty \left(\theta_{\Lambda_n}(u_n^2 t) - 1\right) \frac{dt}{t} \to \int_1^\infty (\theta_A(t) - 1)\frac{dt}{t}$$

in view of the pointwise convergence from Proposition 4.4 and the uniform integrable upper bound from Lemma 4.5.

The middle integral

$$\int_1^\infty u_n^r e^{-2ru_n^2 t} I_0(2u_n^2 t)^r \frac{dt}{t}$$

converges to

$$\int_1^\infty (4\pi t)^{-r/2}\frac{dt}{t} = \frac{2}{r}(4\pi)^{-r/2}$$

in view of the heat kernel convergence from Proposition 4.2 and Proposition 4.3, so then we may appeal to the Lebesgue dominated convergence theorem. □

Next we show:

PROPOSITION 4.7. *With the notation as above and* $u_n := \det(\Lambda_n)^{1/r}$, *we have that*

$$\int_0^1 \left(\theta_{\Lambda_n}(u_n^2 t) - u_n^r e^{-2ru_n^2 t} I_0(2u_n^2 t)^r\right) \frac{dt}{t} \to \int_0^1 \left(\theta_A(t) - (4\pi t)^{-r/2}\right) \frac{dt}{t}$$

as $n \to \infty$.

PROOF. For fixed t we have the pointwise convergence as $n \to \infty$

$$\theta_{\Lambda_n}(u_n^2 t) - u_n^r e^{-2ru_n^2 t} I_0(2u_n^2 t)^r \to \theta_A(t) - (4\pi t)^{-r/2}.$$

It remains therefore to exhibit uniform (for $n >> 1$) integrable bounds on the integrands. This can be done in the same way as in the proof of Proposition 4.4. In order to make the bound obtained there, in terms of λ, integrable for $0 \le t \le 1$, we just need to choose n_0 large so that $n_0 a/4 > r/2$. □

Finally we recall from [**CJK10**]:

PROPOSITION 4.8. *For $u \in \mathbb{R}$ we have the asymptotic formula*

$$\int_0^1 (e^{-u^2 t} - 1)\frac{dt}{t} = \Gamma'(1) - 2\log(u) + o(1)$$

as $u \to \infty$.

We now turn to the proof of our main result, Theorem 1.

In view of Theorem 4.1 we have

$$\log\det \Delta_{DT(\Lambda_n)} = \det \Lambda_n \cdot c_r - \int_0^\infty \left(\theta_{\Lambda_n}(t) - \det \Lambda_n \cdot e^{-2rt} I_0(2t)^r - 1 + e^{-t}\right)\frac{dt}{t}.$$

After the change of variables $t \to u_n^2 t$, the second term becomes

$$\begin{aligned}
&-\int_0^\infty \left(\theta_{\Lambda_n}(u_n^2 t) - \det \Lambda_n \cdot e^{-2ru_n^2 t} I_0(2u_n^2 t)^r - 1 + e^{-u_n^2 t}\right)\frac{dt}{t}\\
&= -\left[\int_0^1 + \int_1^\infty\right]\left(\theta_{\Lambda_n}(u_n^2 t) - \det \Lambda_n \cdot e^{-2ru_n^2 t} I_0(2u_n^2 t)^r - 1 + e^{-u_n^2 t}\right)\frac{dt}{t}.
\end{aligned}$$

In view of Propositions 4.6, 4.7, and 4.8, this integral equals

$$\begin{aligned}
&-\int_1^\infty (\theta_A(t) - 1)\frac{dt}{t} + \frac{2}{r}(4\pi)^{-r/2} - \int_0^1 \left(\theta_A(t) - (4\pi t)^{-r/2}\right)\frac{dt}{t}\\
&- \Gamma'(1) + 2\log(u_n) + o(1).
\end{aligned}$$

Keeping in mind that $u_n = (\det \Lambda_n)^{1/r}$ and identifying the constant terms appearing in the meromorphic continuation of $-\zeta'(0) = \log\det \Delta_{A\mathbb{Z}^r\backslash\mathbb{R}^r}$, the main theorem is proved; cf. equation (15) of [**CJK10**] with $V(A) = \det A = 1$.

5. Proof of the corollaries

To prove the corollaries in the introduction we recall the statements from the literature that we use.

For Corollary 1.2 note that the height has a global minimum for the hexagonal lattice A_2 in dimension 2 as is well-known and for the f.c.c. lattice $A_3 \cong D_3$ in dimension 3 by the rigorous numerics of Sarnak-Strömbergsson in [**SS06**]. Hence for any unimodular lattice L in the respective dimensions

$$\begin{aligned}
h(L) &\geq h(A_2), \text{ so } \log\det \Delta_L \leq \log\det \Delta_{A_2}\\
h(L) &\geq h(D_3), \text{ so } \log\det \Delta_L \leq \log\det \Delta_{D_3}.
\end{aligned}$$

As already remarked the two leading terms in the asymptotics in Theorem 1.1 are shape-independent, and so Corollary 1.2 follows from these remarks arguing with convergent subsequences in view of the compactness.

For Corollary 1.3 note that Corollary 1 on p. 119 in [**SS06**] implies that

$$\log\det \Delta_M < \gamma - \log 4\pi + \frac{2}{r} < -0.95$$

where M is an r-dimensional flat torus of volume 1 and γ is Euler's constant $\gamma \approx 0.577$. In view of this statement, let us replace Λ_n by a convergent subsequence. Then by Theorem 1.1 and the matrix-tree theorem, we have

$$\tau(DT(\Lambda_n)) = \frac{\det' \Delta_{DT(\Lambda_n)}}{\det \Lambda_n} \leq \frac{(\det \Lambda_n)^{2/r-1}}{4\pi}\exp(\det \Lambda_n \cdot c_r + \gamma + 2/r)$$

for all sufficiently large n. This concludes the proof of Corollary 1.3.

Finally, it may be of interest to mention another estimate in [**SS06**]:

$$h(L) \geq 4\sqrt{\frac{\pi}{r}} \left(\frac{\sqrt{r/2\pi e}}{m(L)} \right)^r (1 + o(1)), \quad \text{as } r \to \infty$$

where $m(L)$ is the length of the shortest non-zero vector in the lattice L. Recall that being the densest regular packing is equivalent to being the lattice with co-volume 1 which maximizes the length of the shortest nonzero vector.

References

[CJK10] Chinta, G., Jorgenson, J., Karlsson, A., Zeta functions, heat kernels and spectral asymptotics on degenerating families of discrete tori. Nagoya Math. J. 198 (2010), 121-172 MR2666579 (2011i:58052)

[Ch97] Chiu, P., Height of flat tori. Proc. Am. Math. Soc. 125, 723–730 (1997) MR1396970 (97g:58173)

[CY99] Chung, F. and Yau, S.-T., Coverings, heat kernels and spanning trees. Electronic J. Combin. 6 (1999) Research Paper 12, 21 pp. MR1667452 (2000g:35079)

[CK09] Cohn, H. and Kumar, A., Optimality and uniqueness of the Leech lattice among lattices. Ann. of Math. (2) 170 (2009), no. 3, 1003–1050. MR2600869 (2011c:11106)

[CS99] Conway, J.H., Sloane, N.J.A.: Sphere Packings, Lattices and Groups, (3rd edn) Grundl. Math. Wiss. 290, New York: Springer Verlag 1999 MR1662447 (2000b:11077)

[DM06] Dodziuk, J., Mathai, V., Kato's inequality and asymptotic spectral properties for discrete magnetic Laplacians. In: The ubiquitous heat kernel, Contemporary Math. AMS, 398, (2006), 269–297.

[DD88] Duplantier, B., David, F., Exact partition functions and correlation functions of multiple Hamiltonian walks on the Manhattan lattice. Journal of Statistical Physics, 51 (1988), 327–434. MR952941 (89m:82005)

[L] Lück, Wolfgang, L^2-invariants: theory and applications to geometry and K-theory. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, 44. Springer-Verlag, Berlin, 2002. xvi+595 pp. ISBN: 3-540-43566-2 MR1926649 (2003m:58033)

[M] McKay, Brendan D., Spanning trees in regular graphs. European J. Combin. 4 (1983), no. 2, 149–160. MR705968 (85d:05194)

[OPS88] Osgood, B., Phillips, R., Sarnak, P., Extremals of determinants of Laplacians. J. Funct. Anal. 80, 148–211 (1988) MR960228 (90d:58159)

[S90] Sarnak, P., Determinants of Laplacians; Heights and Finiteness. In: Analysis, et cetera, pp. 601–622, ed. by P.H. Rabinowitz, E. Zehnder. Boston, MA: Academic Press 1990 MR1039364 (91d:58260)

[SS06] Sarnak, P. and Strömbergsson, A., Minima of Epstein's Zeta function and heights of flat tori. Invent. math. 165, 115–151 (2006) MR2221138 (2008c:11059)

Department of Mathematics, The City College of New York, New York, New York 10031

E-mail address: `chinta@sci.ccny.cuny.edu`

Department of Mathematics, The City College of New York, New York, New York 10031

E-mail address: `jjorgenson@mindspring.com`

Section de mathématiques, Université de Genève, 1211 Genève 4, Switzerland

E-mail address: `Anders.Karlsson@unige.ch`

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11231

Flows with uncountable but meager group of self-similarities

Alexandre I. Danilenko

Dedicated to A. M. Stepin on the occasion of his 70-th birthday

Abstract. Given an ergodic probability preserving flow $T = (T_t)_{t\in\mathbb{R}}$, let $I(T) := \{s \in \mathbb{R}^* \mid T \text{ is isomorphic to } (T_{st})_{t\in\mathbb{R}}\}$. A weakly mixing Gaussian flow T is constructed such that $I(T)$ is uncountable and meager. For a Poisson flow T, a subgroup $I_{\mathrm{Po}}(T) \subset I(T)$ of Poissonian self-similarities is introduced. Given a probability measure κ on $\mathbb{R}^*_+$, a Poisson flow T is constructed such that $I_{\mathrm{Po}}(T)$ is the group of κ-quasi-invariance.

0. Introduction

Throughout the paper $\mathbb{R}^*$ denotes the multiplicative group of reals different from 0 and $\mathbb{R}^*_+$ denotes the multiplicative group of positive reals. Let $T = (T_t)_{t\in\mathbb{R}}$ be an ergodic free measure preserving flow on a standard non-atomic probability space $(X, \mathfrak{B}, \mu)$. Given $s \in \mathbb{R}^*$, we denote by $T \circ s$ the flow $(T_{st})_{t\in\mathbb{R}}$. Let

$$I(T) := \{s \in \mathbb{R}^* \mid T \circ s \text{ is isomorphic to } T\}.$$

It is easy to see that $I(T)$ is a multiplicative subgroup of $\mathbb{R}^*$. It is called the group of *self-similarities* of T. If $I(T) \neq \{1, -1\}$ then T is called *self-similar*. There are a lot of problems related to the self-similarities of ergodic flows. We refer the reader to the recent papers [**FrLe**], [**DaRy**] and references therein for details. In particular, it was shown there that given any countable subgroup G in $\mathbb{R}^*$, there is a weakly mixing flow T with $I(T) = G$. In this note we consider Problem 2 from the "Open problems" section in [**FrLe**]:

— is there an ergodic flow T for which the group $I(T)$ is uncountable but has zero Lebesgue measure?

Examples of such flows were constructed recently in [**DaRy**, Section 2]. However the construction there incorporates essentially some subtle facts from the measurable orbit theory such as the theorem on outer conjugacy for groups of automorphisms of continuous ergodic equivalence relations [**VeFe**] and Ratner's theory on joinings of horocycle flows [**Ra**]. Our purpose here is to provide other (simpler) examples, independent of [**VeFe**] and [**Ra**].

Theorem 0.1. *There exist weakly mixing Gaussian flows with uncountable but meager group of self-similarities.*

2010 *Mathematics Subject Classification.* Primary 37A10.
Key words and phrases. Ergodic flow, self-similarity, Poisson flow.

In this connection we recall that

— $I(T)$ is a Borel subset of $\mathbb{R}^*$ for each flow T [**DaRy**] and

— a Borel subgroup of $\mathbb{R}^*$ is meager if and only if it has zero Lebesgue measure.

It should be noted (see Remark 1.2 below) that the flows constructed in Theorem 0.1 are different (in fact, disjoint in the sense of Furstenberg [**Fu**]) from the mixing flows with uncountable meager group of self-similarities constructed in [**DaRy**, Section 2].

Next, for each Poisson flow T, we define a subgroup $I_{\mathrm{Po}}(T) \subset I(T)$ of Poissonian self-similarities. It consists of those $s \in \mathbb{R}^*$ such that $T \circ s$ is conjugate to T via a Poisson transformation. Given a probability measure τ on a locally compact second countable Abelian group G, let $H_G(\tau)$ be the set of all $g \in G$ such the translation of τ by g is equivalent to τ. Then $H_G(\tau)$ is a Borel subgroup in G (see [**Na**] and references therein). It is called the *group of τ-quasi-invariance*.

Recently, Ryzhikov [**Ry**] constructed a class of Poisson flows T with *total self-similarity*, i.e. $I(T) = \mathbb{R}^*$[1]. Generalizing his construction, we prove the following claim.

THEOREM 0.2. *For each probability measure κ on $\mathbb{R}^*_+$, there is a weakly mixing Poisson flow $\widetilde{T}$ with $I_{Po}(\widetilde{T}) = H_{\mathbb{R}^*_+}(\kappa)$.*

The author thanks E. Roy and the anonymous referees for the useful remarks.

1. Gaussian examples

Fix an onto homomorphism[2] $\pi : \mathbb{R} \to \mathbb{T}$. Then for each probability measure ρ on $\mathbb{T}$, there is a probability measure ρ' on $\mathbb{R}$ such that $\rho' \circ \pi^{-1} \sim \rho$ and $H_{\mathbb{R}}(\rho') := \pi^{-1}(H_{\mathbb{T}}(\rho))$. We call ρ' a *standard lift* of ρ to $\mathbb{R}$. It is easy to verify that the equivalence class of the standard lift of ρ is defined uniquely by the equivalence class of ρ.

Consider an isomorphism $\vartheta : \mathbb{R} \to \mathbb{R}^*_+$. Let σ denote the only non-atomic symmetric probability measure on $\mathbb{R}$ such that $\sigma \restriction \mathbb{R}^*_+ = \frac{1}{2}\rho' \circ \vartheta^{-1}$. We use here a natural (set theoretical only, not group theoretical) embedding $\mathbb{R}^*_+ \subset \mathbb{R}$. Then

$$H_{\mathbb{R}^*}(\sigma) = \vartheta(H_{\mathbb{R}}(\rho')) \cup (-\vartheta(H_{\mathbb{R}}(\rho'))).$$

It follows that if $H_{\mathbb{T}}(\rho)$ is uncountable and meager in $\mathbb{T}$ then $H_{\mathbb{R}^*}(\sigma)$ is uncountable and meager in $\mathbb{R}^*$.

Suppose that ρ is non-atomic and ergodic with respect to the non-singular action of $H_{\mathbb{T}}(\rho)$ on $(\mathbb{T}, \rho)$ by translations. We recall (see [**AaNa**] and [**Na**]) that $H_{\mathbb{T}}(\rho)$ admits a unique natural topology in which it is a Polish group. Then ρ is called $H_{\mathbb{T}}(\rho)$-ergodic if it is G-ergodic for some (and hence for every) dense countable subgroup G of $H_{\mathbb{T}}(\rho)$. It follows that ρ' is $H_{\mathbb{R}}(\rho')$-ergodic. We also note that $H_{\mathbb{R}}(\rho')$ is dense in $\mathbb{R}$. Next, σ is ergodic with respect $H_{\mathbb{R}^*}(\sigma)$ acting on $\mathbb{R}^*$ by multiplication and $H_{\mathbb{R}^*}(\sigma)$ is dense in $\mathbb{R}^*$.

Let T denote the Gaussian flow associated with σ (see [**Co-Si**]).

PROPOSITION 1.1. *If ρ is a non-atomic probability measure on $\mathbb{T}$ such that*

[1]The Bernoulli flow with infinite entropy and horocycle flows are other examples of flows with total self-similarities. Ryzhikov's examples are disjoint from them.

[2]The group homomorphisms (and isomorphisms) considered in this note are topological, i.e. continuous.

(1) *$H_{\mathbb{T}}(\rho)$ is uncountable and meager and*
(2) *ρ is ergodic with respect to $H_{\mathbb{T}}(\rho)$*

then $I(T) \supset H_{\mathbb{R}^*}(\sigma)$ but $\mathbb{R}^*_+ \not\subset I(T)$.

Proof. Since T is a Gaussian flow associated with σ, it follows that $I(T) \supset H_{\mathbb{R}^*}(\sigma)$ [**FrLe**]. On the other hand, it is well known that the reduced maximal spectral type of T is given by the measure $\exp' \sigma := \sum_{p\geq 1} \frac{\sigma^{*p}}{p!}$, where σ^{*p} denotes the p-th convolution power of σ [**Co-Si**]. Hence $I(T) \subset H_{\mathbb{R}^*}(\exp' \sigma)$. We claim that $I(T) \neq \mathbb{R}^*$. Indeed, otherwise $H_{\mathbb{R}^*}(\exp' \sigma) = \mathbb{R}^*$ and hence $\exp' \sigma$ is equivalent to Lebesgue measure $\lambda_{\mathbb{R}^*}$. Since $\lambda_{\mathbb{R}^*}$ and σ are both ergodic with respect to $H_{\mathbb{R}^*}(\sigma)$ and $\lambda_{\mathbb{R}^*} \sim \exp' \sigma \succ \sigma$, it follows that $\lambda_{\mathbb{R}^*} \sim \sigma$ and hence $H_{\mathbb{R}^*}(\sigma) = \mathbb{R}^*$, a contradiction. In a similar way we can verify that $I(T) \not\supset \mathbb{R}^*_+$. □

Proof of Theorem 0.1. Examples of ρ satisfying (1) and (2) are given in [**AaNa**, Theorem 4.1] and [**Pa**]. Now it suffices to note that

— T is weakly mixing whenever σ is non-atomic,
— if a subgroup of $\mathbb{R}^*$ does not contain $\mathbb{R}^*_+$ then it is meager in $\mathbb{R}^*$

and apply Proposition 1.1. □

We now briefly explain how to construct measures satisfying (1) and (2). Suppose we have a sequence of positive integers $(n_j)_{j\geq 1}$ and a sequence of complex numbers $(a_j)_{j\geq 1}$ such that $n_j > 2(n_1 + \cdots + n_{j-1})$ and $|a_j| \leq 1$ for all j. We let $P_j(z) := 1 + a_j z^{n_j}/2 + \overline{a_j} z^{-n_j}/2$ for $z \in \mathbb{T}$. Then the polynomial P_j is non-negative and $\int_{\mathbb{T}} \prod_{k=1}^{j} P_k \, d\lambda_{\mathbb{T}} = 1$ for each j, where $\lambda_{\mathbb{T}}$ is the normed Haar measure on $\mathbb{T}$. The sequence of probability measures $(\prod_{k=1}^{j} P_k)\lambda_{\mathbb{T}}$ converges weakly as $j \to \infty$ to a probability measure ρ on $\mathbb{T}$ which is called the *Riesz product* associated with $(n_j)_{j\geq 1}$ and $(a_j)_{j\geq 1}$ [**Na**]. The Fourier transform of ρ is explicitly known: $\widehat{\rho}(n) = 0$ except if n is a finite sum $\sum_j k_j n_j$ with $k_j \in \mathbb{Z}$ and $|k_j| \leq 1$[3], in which case $\widehat{\rho}(n) = \prod_j a_j^{k_j}$. Using standard methods of harmonic analysis and classical facts about equivalence of Riesz products, F. Parreau shows in [**Pa**] that

— $H(\rho) = \{z \in \mathbb{T} \mid \sum_{j\geq 1} |a_j|^2 |1 - a_j z^{n_j}|^2 < \infty\}$ and
— if the lacunary condition $\sum_{j\geq 1} |a_j|^2 (n_j/n_{j+1})^2 < \infty$ holds then ρ is ergodic with respect to $H(\rho)$. Moreover, in this case $H(\rho)$ is uncountable.
— If $\sum_i |a_j|^2 = \infty$ then $H(\rho) \neq \mathbb{T}$.

Thus we see that if ρ is the Riesz product associated with $n_j = j!$ and $a_j = 1$ for all j then ρ satisfies (1) and (2).

Remark 1.2. We note that the flows constructed in Proposition 1.1 are quite different from flows with uncountable but meager group of self-similarities constructed in [**DaRy**, §3]. To explain this precisely we recall some concepts from the theory of joinings of dynamical systems (see [**Th**], [**dR**]). Given two ergodic flows (X, μ, T) and (Y, ν, S), a joining of T and S is a $T \times S$-invariant measure on $X \times Y$ whose marginals on X and Y are μ and ν respectively. For instance, $\mu \times \nu$ is a joining of T and S. If there are no other joinings then T and S are called disjoint [**Fu**]. Of course, disjoint flows are non-isomorphic. A flow (X, μ, T) is called 2-fold

[3]Due to the condition imposed on $(n_j)_{j\geq 1}$, if n can be written as $\sum_j k_j n_j$ then this representation is unique. Hence $\widehat{\rho}(n)$ is well defined.

quasi-simple if for every ergodic joining $\lambda \neq \mu \times \mu$ of T with itself, the two marginal projections $(X \times X, \lambda) \to (X, \mu)$ are finite-to-one $\mod \lambda$.[4] The examples from [**DaRy**, §3] are 2-point extensions of some horocycle flows. Hence they are quasi-simple according to [**RyTh**] and [**dJLe**]. By [**RyTh**] and [**dJLe**], a weakly mixing Gaussian flow is disjoint with each quasi-simple flow. Thus the examples constructed in Theorem 0.1 are disjoint with the flows constructed in [**DaRy**, §3].

2. Poisson flows

Let X be a locally compact non-compact second countable space. Denote by $\widetilde{X}$ the vector space of Radon measures on X. We endow $\widetilde{X}$ with the $*$-weak topology, i.e. the weak topology generated by the natural duality of $\widetilde{X}$ and the space $C_c(X)$ of continuous functions with compact support on X. Since the Borel structure $\widetilde{\mathfrak{B}}$ generated by the $*$-weak topology is the same as the Borel structure generated by the strong topology on $\widetilde{X}$ and the strong topology is Polish, it follows that $\widetilde{\mathfrak{B}}$ is standard Borel. We fix a homeomorphism T of X. It induces canonically a homeomorphism $\widetilde{T}$ of $\widetilde{X}$. Denote by $\mathcal{M}_{+,T}(X) \subset \widetilde{X}$ the closed cone of non-negative (σ-finite) Radon T-invariant measures and denote by $\mathcal{M}_{+,1,\widetilde{T}}(\widetilde{X})$ the set of probability $\widetilde{T}$-invariant measures on $(\widetilde{X}, \widetilde{\mathfrak{B}})$. Given a non-negative measure μ on X, a probability measure $\widetilde{\mu}$ on $\widetilde{X}$ is well defined by the following two conditions:

- (c1) $\widetilde{\mu}([K,j]) = \exp(-\mu(K))\mu(K)^j/j!$ for each compact subset $K \subset X$ and each non-negative integer j,
- (c2) $\widetilde{\mu}([K,j] \cap [K',j']) = \widetilde{\mu}([K,j])\widetilde{\mu}([K',j'])$ whenever $K \cap K' = \emptyset$,

where $[K,j]$ stands for the cylinder $\{\omega \in \widetilde{X} \mid \omega(K) = j\}$. If μ is T-invariant then $\widetilde{\mu}$ is $\widetilde{T}$-invariant. The dynamical system $(\widetilde{X}, \widetilde{\mu}, \widetilde{T})$ is called the *Poisson suspension* of (X, μ, T). This definition can be naturally extended to the case where (X, μ, T) is an arbitrary standard Borel μ-preserving dynamical system with μ infinite and σ-finite (see, e.g., [**Ro**], [**Ja–dR**]). If a probability preserving transformation is measure theoretically isomorphic to a Poisson suspension then it is called a *Poisson* transformation.

We recall an interesting example by Ryzhikov from his recent work [**Ry**]. Let $V = (V_t)_{t\in\mathbb{R}}$ be an ergodic conservative measure preserving flow on an infinite measure space (Z, ν). We now define a conservative measure preserving flow $T = (T_t)_{t\in\mathbb{R}}$ on the infinite measure space $(X, \mu) := (\mathbb{R}^* \times Z, \lambda_{\mathbb{R}^*} \times \nu)$ by

$$T_t(s, z) := (s, V_{st}z).$$

Since T has no invariant subsets of finite measure, $\widetilde{T}$ is weakly mixing [**Ro**]. It is easy to see that $I(T) = \mathbb{R}^*$. The self-similarity group for the infinite measure preserving flows is defined in the very same way as for the probability preserving flows. Hence for the Poisson suspension $\widetilde{T} = (\widetilde{T}_t)_{t\in\mathbb{R}}$ of T, we also obtain $I(\widetilde{T}) = \mathbb{R}^*$.

We now consider a generalization of Ryzhikov's example (see the construction in the proof of Theorem 0.2). Let $\widetilde{T} = (\widetilde{T}_t)_{t\in\mathbb{R}}$ be a Poisson suspension of a measure preserving flow $T = (T_t)_{t\in\mathbb{R}}$ on a σ-finite measure space (X, μ). We let $I_{\mathrm{Po}}(\widetilde{T}) := I(T)$. Then $I_{\mathrm{Po}}(\widetilde{T})$ is obviously a subgroup of $I(\widetilde{T})$. We call it the *group of Poissonian self-similarities* of $\widetilde{T}$.

[4]More generally, the flow is called 2-fold quasi-simple if the aforementioned marginal extensions are isometric [**RyTh**] or even distal [**dJLe**].

Proof of Theorem 0.2. Let V be an ergodic conservative measure preserving flow V on an infinite σ-finite measure space (Y, ν). Let σ stand for a measure of the maximal spectral type of V. For $s > 0$, we denote by σ_s the image of σ under the mapping $\mathbb{R} \ni t \mapsto ts \in \mathbb{R}$. Then σ_s is a measure of the maximal spectral type of $V \circ s$. Suppose that $\sigma_s \perp \sigma$ for each $1 \neq s > 0$ (see [**DaRy**] for examples of such flows). Fix a probability measure κ on $\mathbb{R}_+^*$. We now define a measure preserving flow $T = (T_t)_{t \in \mathbb{R}}$ on the infinite measure space $(X, \mu) := (\mathbb{R}_+^* \times \mathbb{R} \times Z, \kappa \times \lambda_{\mathbb{R}} \times \nu)$ by

$$T_t(s, y, z) := (s, y, V_{st} z).$$

It follows from our assumption on V that if $V \circ s = RV \circ s' R^{-1}$ for $s, s' > 0$ and a measure preserving transformation R of (Y, ν) then $s = s'$.

Since T has no invariant subsets of finite measure, it follows that the Poisson suspension $\widetilde{T}$ of T is weakly mixing. We claim that $I_{\text{Po}}(\widetilde{T}) = H_{\mathbb{R}_+^*}(\kappa)$. It is easy to see that $H_{\mathbb{R}_+^*}(\kappa) \subset I(T)$. Indeed, if $h \in H_{\mathbb{R}_+^*}(\kappa)$ then we let

$$Q(s, y, z) := (hs, y \cdot \frac{d\kappa}{d\kappa_h}(s), z).$$

Then Q is a μ-preserving transformation of X and $Q^{-1} T Q = T \circ h$. Now we prove the converse inclusion $I(T) \subset H_{\mathbb{R}_+^*}(\kappa)$. Take $h \in I(T)$. Then there is a measure preserving transformation Q of X such that $Q^{-1} T Q = T \circ h$. It is easy to write the T-ergodic decomposition of μ:

$$\mu = \int_{\mathbb{R}_+^* \times \mathbb{R}} \delta_s \times \delta_y \times \nu \, d\kappa(s) dy \tag{2-1}$$

The restriction of T to $(X, \delta_s \times \delta_y \times \nu)$ is isomorphic to $V \circ s$. Therefore two ergodic components $(X, \delta_s \times \delta_y \times \nu, T)$ and $(X, \delta_{s'} \times \delta_{y'} \times \nu, T)$ are isomorphic if and only if $s = s'$. We note that (2-1) is also the $T \circ h$-ergodic decomposition of μ. However the ergodic component $(X, \delta_s \times \delta_y \times \nu, T \circ h)$ is isomorphic to $V \circ (hs)$. Since Q conjugates each ergodic component of T with an ergodic component of $T \circ h$, we obtain that

$$Q(s, y, z) = (h^{-1} s, \bullet, \bullet).$$

Thus the marginal projection $X \ni (s, y, z) \mapsto s \in \mathbb{R}_+^*$ intertwines Q with the mapping $\mathbb{R}_+^* \ni s \mapsto h^{-1} s \in \mathbb{R}_+^*$. Since Q is μ-preserving, it follows that this mapping is κ-nonsingular, i.e. $h \in H_{\mathbb{R}_+^*}(\kappa)$. □

3. Concluding remarks and problems

(1) (Ryzhikov's question) Are there non-mixing weakly mixing flows with uncountable group of self-similarities? The examples from [**DaRy**, Section 2] are mixing. We conjecture that the class of Gaussian flows constructed in Theorem 0.1 contains non-mixing flows.

(2) Are there *prime* weakly mixing flows with uncountable but meager group of of self-similarities? The flows constructed in [**DaRy**, Section 2] and Theorem 0.1 have non-trivial factors.

(3) We do not know whether $I_{\text{Po}}(\widetilde{T}) = I(\widetilde{T})$ for the Poisson flows $\widetilde{T}$ constructed in Theorem 0.2. Are there some extra conditions on κ and V which imply the equality $I_{\text{Po}}(\widetilde{T}) = I(\widetilde{T})$? It is easy to deduce from [**Ro**, Proposition 5.2] that this equality holds if the maximal spectral

type τ of T is orthogonal to $\sum_{j>1}\frac{1}{j!}\tau^{*j}$. A direct calculation shows that $\tau=\int_{\mathbb{R}^*}\sigma_s\,d\kappa(s)$, where σ is a measure of the maximal spectral type of V.

(4) On the other hand, it is also interesting to find a Poisson flow $\widetilde{T}$ with $I_{\mathrm{Po}}(\widetilde{T})\neq I(\widetilde{T})$.

References

[AaNa] J. Aaronson and M. Nadkarni, *L^∞-eigenvalues and L^2-spectra of non-singular transformations*, Proc. London Math. Soc. **55** (1987), 538–570. MR907232 (88j:28012)

[Co-Si] I. Cornfeld, S. Fomin, Ya. G. Sinai, *Ergodic Theory*, Springer-Verlag, New York, 1982. MR832433 (87f:28019)

[DaRy] A. I. Danilenko and V. V. Ryzhikov, *On self-similarities of ergodic flows*, Proc. London Math. Soc. (to appear).

[dJLe] A. del Junco and M. Lemańczyk, *Joinings of distally simple automorphisms*, in preparation.

[dR] T. de la Rue, *Joinings in ergodic theory*, Encyclopedia of Complexity and Systems Science (Robert A. Meyers, ed.), Springer, 2009, pp. 5037–5051.

[FrLe] K. Frączek and M. Lemańczyk, *On the self-similarity problem for ergodic flows*, Proc. London Math. Soc. **99** (2009), 658–696. MR2551467 (2011a:37006)

[Fu] H. Furstenberg, *Disjointness in ergodic theory, minimal sets and diophantine approximation*, Math. Syst. Th. **1** (1967), 1–49. MR0213508 (35:4369)

[Ja–dR] E. Janvresse, T. Meyerovitch, E. Roy, T. de la Rue, *Poisson suspensions and entropy for infinite transformations*, Trans. Amer. Math. Soc. **362** (2010), 3069–3094. MR2592946 (2011b:37012)

[Na] M. G. Nadkarni, *Spectral theory of dynamical systems*, Birkhäuser Advanced Texts: Basler Lehrbücher, Birkhäuser Verlag, Basel, 1998. MR1719722 (2001d:37001)

[Pa] F. Parreau, *Ergodicité et pureté des produits de Riesz*, Ann. Inst. Fourier **40** (1990), 391–405. MR1070833 (91g:42009)

[Ra] M. Ratner, *Ergodic theory in hyperbolic space*, Contemp. Math., vol. 26, Amer. Math. Soc., Providence, R. I., pp. 309–334. MR737411 (85h:58140)

[Ro] E. Roy, *Poisson suspensions and infinite ergodic theory*, Erg. Th. & Dyn. Sys. **29** (2009), 667–683. MR2486789 (2011h:37011)

[Ry] V. V. Ryzhikov, *On mixing rank one infinite transformations*, preprint, arXiv:1106.4655.

[RyTh] V. V. Ryzhikov and J.-P. Thouvenot, *Disjointness, divisibility and quasi-simplicity of measure-preserving actions*, Funct. Anal. Appl. **40** (2006), 237–240. MR2265691 (2007i:37005)

[Th] J.-P. Thouvenot, *Some properties and applications of joinings in ergodic theory*, Ergodic theory and its connections with harmonic analysis (Alexandria, 1993), 207–235, London Math. Soc. Lecture Note Ser., 205, Cambridge Univ. Press, Cambridge, 1995. MR1325699 (96d:28017)

[VeFe] A. S. Veksler and A. L. Fedorov, *Conjugacy of homomorphisms of locally compact groups into the outer group of measurable equivalence relation*, Funct. Anal. Applic. **22** (1988), 318–319. MR976999 (90c:28032)

Institute for Low Temperature Physics & Engineering of National Academy of Sciences of Ukraine, 47 Lenin Ave., Kharkov, 61164, UKRAINE

E-mail address: `alexandre.danilenko@gmail.com`

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11237

The universal minimal space of the homeomorphism group of a h-homogeneous space

Eli Glasner and Yonatan Gutman

To Anatoliĭ Stepin with great respect.

ABSTRACT. Let X be a h-homogeneous zero-dimensional compact Hausdorff space, i.e. X is a Stone dual of a homogeneous Boolean algebra. It is shown that the universal minimal space $M(G)$ of the topological group $G = \mathrm{Homeo}(X)$, is the space of maximal chains on X introduced in the paper by V. V. Uspenskij. If X is metrizable then clearly X is homeomorphic to the Cantor set and the result was already known (see the paper by Glasner and Weiss in *Fund. Math.* **176**(3)). However many new examples arise for non-metrizable spaces. These include, among others, the generalized Cantor sets $X = \{0,1\}^\kappa$ for non-countable cardinals κ, and the *corona* or *remainder* of ω, $X = \beta\omega \setminus \omega$, where $\beta\omega$ denotes the Stone-Čech compactification of the natural numbers.

1. Introduction

The existence and uniqueness of a universal minimal G dynamical system, corresponding to a topological group G, is due to Ellis (see [**Ell69**], for a new short proof see [**GL11**]). He also showed that for a discrete infinite G this space is never metrizable, and the latter statement was generalized to the locally compact non-compact case by Kechris, Pestov and Todorcevic in the appendix to their paper [**KPT05**]. For Polish groups this is no longer the case and we have such groups with $M(G)$ being trivial (groups with the fixed point property or extremely amenable groups) and groups with metrizable, easily described $M(G)$, like $M(G) = S^1$ for the group $G = \mathrm{Homeo}_+(S^1)$ ([**Pes98**]) and $M(G) = LO(\omega)$, the space of linear orders on a countable set, for $S_\infty(\omega)$ the permutation group of the integers ([**GW02**]).

Following Pestov's work Uspenskij has shown in [**Usp00**] that the action of a topological group G on its universal minimal system $M(G)$ (with cardinality $M(G) \geq 3$) is never 3-transitive so that, e.g., for manifolds X of dimension > 1 as well as for $X = Q$, the Hilbert cube, and $X = K$, the Cantor set, $M(G)$ can not coincide with X. Uspenskij proved his theorem by introducing the space of maximal chains $\Phi(X)$ associated to a compact space X. In [**GW03**] the authors then showed that

2010 *Mathematics Subject Classification.* Primary 37B05, 54H10, 54H20, 03G05, 06E15.

Key words and phrases. Universal minimal space, h-homogeneous, homogeneous Boolean algebra, maximal chains, generalized Cantor sets, corona, remainder, Parovičenko space, collapsing algebra, dual Ramsey Theorem.

The first named author's research was supported by Grant No 2006119 from the United States-Israel Binational Science Foundation (BSF).

for X the Cantor set and $G = \mathrm{Homeo}(X)$, in fact, $M(G) = \Phi$. It turns out that this group G is a closed subgroup of $S_\infty(\omega)$ and in [**KPT05**] Kechris, Pestov and Todorcevic unified and extended these earlier results and carried out a systematic study of the spaces $M(G)$ for many interesting closed subgroups of S_∞.

In the present work we go back to [**GW03**] and generalize it in another direction. We consider the class of zero-dimensional compact h-homogeneous spaces X and show that for every space in this class the universal minimal space $M(G)$ of the topological group $G = \mathrm{Homeo}(X)$ is again Uspenskij's space of maximal chains on X. If X is metrizable then clearly X is homeomorphic to the Cantor set and the result of [**GW03**] is retrieved (although even in this case our proof is new, as we make no use of a fixed point theorem). However, many new examples arise when one considers non-metrizable spaces. These include, among others, the generalized Cantor sets $X = \{0,1\}^\kappa$ for non-countable cardinals κ, and the widely studied **corona** or **remainder of** ω, $X = \beta\omega\setminus\omega$, where $\beta\omega$ denotes the Stone-Čech compactification of the natural numbers. As in [**GW03**] the main combinatorial tool we apply is the dual Ramsey theorem.

1.1. H-homogeneous spaces and homogeneous Boolean algebras. The following definitions are well known (see e.g. [**HNV04**] Section H-4):

(1) A zero-dimensional compact Hausdorff topological space X is called **h-homogeneous** if every non-empty clopen subset of X is homeomorphic to the entire space X.
(2) A Boolean algebra B is called **homogeneous** if for any nonzero element a of B the relative algebra $B|a = \{x \in B : x \leq a)$ is isomorphic to B.

Using Stone's Duality Theorem (see [**BS81**] IV§4) a zero-dimensional compact Hausdorff h-homogeneous space X is the Stone dual of a homogeneous Boolean Algebra, i.e. any such space is realized as the space of ultrafilters B^* over a homogeneous Boolean algebra B, equipped with the topology given by the base $N_a = \{U \in B^* : a \in U\}$, $a \in B$. Here are some examples of h-homogeneous spaces (see [**ŠR89**]):

(1) The countable atomless Boolean algebra is homogeneous. It corresponds by Stone duality to the Cantor space $K = \{0,1\}^{\mathbb{N}}$.
(2) Every infinite free Boolean algebra is homogeneous. These Boolean algebras correspond by Stone duality to the generalized Cantor spaces, $\{0,1\}^\kappa$, for infinite cardinals κ
(3) Let $P(\omega)$ be the Boolean algebra of all subsets of ω (the integers) and let $fin \subset P(\omega)$ be the ideal comprising the finite subsets of ω. Define the equivalence relations $A \sim_{fin} B$, $A, B \in P(\omega)$, if and only if $A \triangle B$ is in fin. The quotient Boolean algebra $P(\omega)/fin$ is homogeneous. This Boolean algebra corresponds by Stone duality to the **corona** $\omega^* = \beta\omega\setminus\omega$, where $\beta\omega$ denotes the Stone-Čech compactification of ω.
(4) A topological space X is called a **Parovičenko space** if:
    (a) X is a zero-dimensional compact space without isolated points and with weight $\mathbf{c}$,
    (b) every two disjoint open F_σ subsets in X have disjoint closures, and
    (c) every non-empty G_δ subset of X has non-empty interior.

    Under CH Parovičenko proved that every Parovičenko space is homeomorphic to ω^* ([**Par63**]).

In [**DM78**] van Douwen and van Mill show that under $\neg$ CH, there are two non-homeomorphic Parovičenko spaces. Their second example of a Parovičenko space is the remainder $X = \beta Y \setminus Y$, where Y is the σ-compact space $\omega \times \{0,1\}^{\mathbf{c}}$. It is not hard to see that in Y the clopen sets are of the form $L = \bigcup_{a \in A} \{a\} \times C_a$ for some $A \subset \omega$, where for all $a \in A$, $C_a \subset \{0,1\}^c$ is non-empty and clopen. If $|A| = \infty$ then $L \cong Y$ and if $|A| < \infty$ then $Cl_{\beta Y}(L) \subset Y$. These facts imply in a straightforward manner that X is h-homogeneous. In [**DM78**] it is pointed out that under MA, X is not homeomorphic to ω^*. Thus under $\neg$ CH+MA, this example provides another weight $\mathbf{c}$ h-homogeneous space.

(5) Let κ be a cardinal. By a well-known theorem of Kripke ([**Kri67**]) there is a homogeneous countably generated complete Boolean algebra, the so called **collapsing algebra** $C(\kappa)$ such that if A is a Boolean algebra with a dense subset of power at most κ, then there is a complete embedding of A in $C(\kappa)$.
(6) It is not hard to check that the product of any number of h-homogeneous spaces is again h-homogeneous.

1.2. The universal minimal space. A compact Hausdorff G-space X is said to be **minimal** if X and $\emptyset$ are the only G-invariant closed subsets of X. By Zorn's lemma each G-space contains a minimal G-subspace. These minimal objects are in some sense the most basic ones in the category of G-spaces. For various topological groups G they have been the object of intensive study. Given a topological group G one is naturally interested in describing all of them up to isomorphism. Such a description is given (albeit in a very weak sense) by the following construction: as was mentioned in the introduction one can show there exists a minimal G-space $M(G)$ unique up to isomorphism such that if X is a minimal G-space then X is a factor of $M(G)$, i.e., there is a continuous G-equivariant mapping from $M(G)$ onto X. $M(G)$ is called the **universal minimal G-space**. Usually this minimal universal space is huge and an explicit description of it is hard to come by.

1.3. The space of maximal chains. Let K be a compact Hausdorff space. We denote by $\mathrm{Exp}(K)$ the space of closed subsets of K equipped with the Vietoris topology. A subset $C \subset \mathrm{Exp}(K)$ is a **chain** in $\mathrm{Exp}(K)$ if for any $E, F \in C$ either $E \subset F$ or $F \subset E$. A chain is **maximal** if it is maximal with respect to the inclusion relation. One verifies easily that a maximal chain in $\mathrm{Exp}(K)$ is a closed subset of $\mathrm{Exp}(K)$, and that $\Phi(K)$, the space of all maximal chains in $\mathrm{Exp}(K)$, is a closed subset of $\mathrm{Exp}(\mathrm{Exp}(K))$, i.e. $\Phi(K) \subset \mathrm{Exp}(\mathrm{Exp}(K))$ is a compact space. Note that a G-action on K naturally induces a G-action on $\mathrm{Exp}(K)$ and $\Phi(K)$. This is true in particular for $K = M(G)$. As the G-space $\Phi(M(G))$ contains a minimal subsystem it follows that there exists an injective continuous G-equivariant mapping $f : M(G) \to \Phi(M(G))$. By investigating this mapping Uspenskij in [**Usp00**] showed that for every topological group G, the action of G on the universal minimal space $M(G)$ is not 3-transitive. As a direct consequence of this theorem only rarely the natural action of the group $G = \mathrm{Homeo}(K)$ on the compact space K coincides with the universal minimal G-action (as is the case for $\mathrm{Homeo}_+(S^1)$). In [**Gut08**] it was shown that for $G = \mathrm{Homeo}(X)$, where X belongs to a family of spaces that contain the closed manifolds of dimension 2 or higher and the Hilbert Cube.

It is easy to see that every $c \in \Phi(K)$ has a first element F which is necessarily of the form $F = \{x\}$. Moreover, calling $x \triangleq r(c)$ the **root** of the chain c, it is clear that the map $\pi : \Phi(K) \to K$, sending a chain to its root, is a homomorphism of dynamical systems.

1.4. The main result. In [**GW03**] it was shown that the universal minimal space of the group of homeomorphisms of the Cantor set, equipped with the compact-open topology, is the space of maximal chains over the Cantor set. Our goal is to prove the following generalization:

THEOREM. *Let X be a h-homogeneous zero-dimensional compact Hausdorff topological space. Let $G = \mathrm{Homeo}(X)$ equipped with the compact-open topology, then $M(G) = \Phi(X)$, the space of maximal chains on X.*

1.5. Acknowledgements. Part of this work was conducted during the Thematic Program on Asymptotic Geometric Analysis at the Fields Institute. The authors would like to thank the Fields Institute for its hospitality. The second author would like to thank Dana Bartošová and Stevo Todorcevic for a helpful discussion regarding h-homogeneous spaces and homogeneous Boolean algebras. Finally, thanks are due to the referee for several helpful remarks.

2. Preliminaries

2.1. Clopen covers. Let X be a zero-dimensional compact Hausdorff space. Denote by $\mathcal{D}$ ($\tilde{\mathcal{D}}$) the directed set (semilattice) consisting of all finite ordered (unordered) clopen partitions of X which are necessarily of the form $\alpha = (A_1, A_2 \ldots, A_m)$ ($\tilde{\alpha} = \{A_1, A_2 \ldots, A_m\}$), where $\dot\cup_{i=1}^{n} A_i = X$ (disjoint union). The relation is given by refinement: $\alpha \preceq \beta$ ($\tilde{\alpha} \preceq \tilde{\beta}$) iff for any $B \in \beta$ ($B \in \tilde{\beta}$), there is $A \in \alpha$ ($A \in \tilde{\alpha}$) so that $B \subset A$. The join (least upper bound) of α and β, $\alpha \vee \beta = \{A \cap B : A \in \alpha, B \in \beta\}$, where the ordering of indices is given by the lexicographical order on the indices of α and β ($\tilde{\alpha} \vee \tilde{\beta} = \{A \cap B : A \in \tilde{\alpha}, B \in \tilde{\beta}\}$). It is convenient to introduce the notations $\mathcal{D}_k = \{\alpha \in \mathcal{D} : |\alpha| = k\}$ and $\tilde{\mathcal{D}}_k = \{\alpha \in \mathcal{D} : |\tilde{\alpha}| = k\}$. We denote the natural map $(A_1, A_2 \ldots, A_m) \mapsto \{A_1, A_2 \ldots, A_m\}$ by $\tilde{t} : \mathcal{D} \to \tilde{\mathcal{D}}$. There is a natural G-action on $\mathcal{D}$ ($\tilde{\mathcal{D}}$) given by $g(A_1, A_2 \ldots, A_m) = (g(A_1), g(A_2) \ldots, g(A_m))$ ($g\{A_1, A_2 \ldots, A_m\} = \{g(A_1), g(A_2) \ldots, g(A_m)\}$). Let S_k denote the group of permutations of $\{1, \ldots, k\}$. S_k acts naturally on $\mathcal{D}_k$ by $\sigma(B_1, B_2 \ldots, B_k) = (B_{\sigma(1)}, B_{\sigma(2)} \ldots, B_{\sigma(k)})$ for any $\beta = (B_1, B_2 \ldots, B_k) \in \mathcal{D}_k$ and $\sigma \in S_k$. This action commutes with the action of G, i.e. $\sigma g \beta = g \sigma \beta$ for any $\sigma \in S_k$ and $g \in G$. Notice that one can identify $\tilde{\mathcal{D}}_k = \mathcal{D}_k / S_k$.

2.2. Partition homogeneity. Let us introduce a new definition:

DEFINITION 2.1. A zero-dimensional compact Hausdorff space X is called **partition-homogeneous** if for every two finite ordered clopen partitions of the same cardinality, $\alpha, \beta \in \mathcal{D}_m$, $\alpha = (A_1, A_2 \ldots, A_m)$, $\beta = (B_1, B_2 \ldots, B_m)$ there is $h \in \mathrm{Homeo}(X)$ such that $hA_i = B_i$, $i = 1, \ldots, k$.

PROPOSITION 2.2. *Let X be an infinite zero-dimensional compact Hausdorff space. X is h-homogeneous iff X is partition-homogeneous.*

PROOF. Assume X is h-homogeneous. Let $\alpha, \beta \in \mathcal{D}_m$, $\alpha = (A_1, A_2 \ldots, A_m)$, $\beta = (B_1, B_2 \ldots, B_m)$. Select homeomorphisms $h_{A,i}, h_{B,i}$, $i = 1, \ldots, m$ with $h_{A,i}$:

$A_i \to X$, $h_{B,i} : B_i \to X$. Define $g \in \mathrm{Homeo}(X)$ by $g(x) = h_{B,i}^{-1} \circ h_{A,i}(x)$ for $x \in A_i$. Trivially $gA_i = B_i$. Assume now X is partition-homogeneous. Let $A \neq X$ be a clopen set in X. We distinguish between two cases:

(1) A is a singleton. As X is partition-homogeneous there exists $h \in \mathrm{Aut}(X)$ with $hA = A^c$ and $hA^c = A$. We conclude X is a two point space contradicting the assumption that X is infinite.
(2) A is not a singleton. Because X is a compact Hausdorff zero-dimensional space we can find disjoint clopen sets A_1, A_2 such that $A = A_1 \cup A_2$. Let $h_1 \in G$ so that $h_1 A_1 = A_1 \cup A^c$ and $h_1 A_1^c = A_2$. Define the homeomorphism $h : A \to X$.

$$h(x) = \begin{cases} h_1(x) & x \in A_1 \\ x & x \in A_2 \end{cases}$$

□

3. Basic properties of h-homogeneous spaces

3.1. Induced orders. Let X be a compact Hausdorff zero-dimensional h-homogeneous space and denote $G = \mathrm{Homeo}(X)$. As X is either trivial or infinite, we will assume from now onward, w.l.og. that X is infinite. Let $v \in \Phi(X)$ and $D \subset X$ a closed set. Define

$$D_v = \bigcap_{A \in v : A \cap D \neq \emptyset} A$$

By maximality of v, one has $D_v \in v$. By a standard compactness argument $D_v \cap D \neq \emptyset$ and trivially it is the minimal element of v that intersects D. Similarly for $D \subset X$ a closed set with $r(v) \in D$, define:

$$D^v = \overline{\bigcup_{A \in v : A \subset D} A}$$

The maximal element of v that is contained in D.

Definition 3.1. Let $v \in \Phi(X)$ and $\tilde{\alpha} = \{A_1, A_2 \ldots, A_m\} \in \tilde{\mathcal{D}}$. Define $<_{v|\tilde{\alpha}} = <_v$, the **induced order** on $\tilde{\alpha}$ by v:

$$A_i <_v A_j \Leftrightarrow (A_i)_v \subseteq (A_j)_v$$

Similarly for $v \in \Phi(X)$ and $\alpha \in \mathcal{D}$, define the induced order $<_{v|\alpha} = <_{v|\tilde{t}(\alpha)}$. Denote by $t_v^* : \tilde{\mathcal{D}} \to \mathcal{D}$ the map $\{A_1, A_2 \ldots, A_m\} \mapsto (A_1, A_2 \ldots, A_m)$ where $i < j$ if and only if $A_i <_{v|\alpha} A_j$. For $\beta \in \mathcal{D}$, define $t_v^*(\beta) = t_v^*(\tilde{t}(\beta))$. Notice that for all $\sigma \in S_k$, $v \in \Phi(X)$ and $\beta \in \mathcal{D}$,

$$t_v^*(\sigma t_v^*(\beta)) = t_v^*(\beta)$$

Lemma 3.2. $g t_v^*(\tilde{\beta}) = t_{gv}^*(g\tilde{\beta})$.

Proof. Let $\alpha = (A_1, A_2 \ldots, A_m) = t_v^*(\tilde{\beta})$. By definition $i < j$ if and only if $(A_i)_v \subseteq (A_j)_v$. Notice $(gA_i)_{gv} = \bigcap_{gA \in gv : gA \cap gA_i \neq \emptyset} gA = g \bigcap_{A \in v : A \cap A_i \neq \emptyset} A = g(A_i)_v$. Therefore $i < j$ if and only if $(gA_i)_{gv} \subseteq (gA_j)_{gv}$, and we conclude $g\alpha = t_{gv}^*(g\tilde{\beta})$. □

PROPOSITION 3.3. *Let $\upsilon \in \Phi(X)$ and $\tilde{\alpha} = \{A_1, A_2 \ldots, A_m\} \in \tilde{\mathcal{D}}$. $<_{\upsilon|\tilde{\alpha}}$ is a linear order on $\tilde{\alpha}$. The ordering $A_{i_1} <_{\upsilon|\tilde{\alpha}} A_{i_2} <_{\upsilon|\tilde{\alpha}} \ldots <_{\upsilon|\tilde{\alpha}} A_{i_m}$ is characterized by $(A_{i_k})_\upsilon \setminus (A_{i_1} \cup \ldots \cup A_{i_{k-1}})^\upsilon = \{x_k\}$ for $k = 1, 2, \ldots m$ and suitable $x_k \in A_k$.*

PROOF. Let $D \subset X$ be clopen so that $r(\upsilon) \in D$, then it is easy to see that $\upsilon_{|D^c} \triangleq \{A \setminus D | D^\upsilon \subsetneq A \in \upsilon\}$ is a maximal chain in D^c and in particular has a root $r(\upsilon_{|D^c}) = x_0 \in D^c$. Let i_1 be such that $r(\upsilon) \in A_{i_1}$. Inductively let i_{k+1} be such that $r(\upsilon_{|(A_{i_1} \cup A_{i_2} \cup \ldots \cup A_{i_k})^c}) \in A_{i_{k+1}}$. It is easy to see $A_{i_1} <_\upsilon A_{i_2} <_\upsilon \ldots <_\upsilon A_{i_m}$. This implies both that $<_\upsilon$ is a linear order and $(A_{i_k})_\upsilon \setminus (A_{i_1} \cup \ldots \cup A_{i_{k-1}})^\upsilon = \{x_k\}$ for some $x_k \in A_k$, $k = 1, 2, \ldots m$. $\square$

3.2. Minimality and proximality of natural actions. The basis for the Vietoris topology for the compact Hausdorff space $\mathrm{Exp}(X)$ is given by open sets of the form:

$$\mathcal{U} = \langle A_1, \ldots, A_k \rangle = \{F \in \mathrm{Exp}(X) : \ \forall i \, F \cap A_i \neq \emptyset \text{ and } F \subset \bigcup A_i\}$$

where $A_i \subset X$ is clopen. It is easy to see that a basis of clopen neighborhood of a maximal chain $\upsilon \in \Phi(X)$ is given by

$$\mathfrak{U}_\alpha = \langle \mathcal{U}_1, \ldots, \mathcal{U}_n \rangle$$

where $\alpha = (A_1, A_2 \ldots, A_n) \in \mathcal{D}$ and

$$\mathcal{U}_j = \langle A_1, \ldots, A_j \rangle, \qquad j = 1, 2, \ldots, n,$$

The following lemma is straightforward:

LEMMA 3.4. *Let $\alpha = (A_1, A_2 \ldots, A_n) \in \mathcal{D}$ and $\upsilon \in \Phi(X)$. Let $<_{\upsilon|\alpha}$ be the induced order of υ on α, then $\upsilon \in \mathfrak{U}_\alpha$ if and only if $<_\upsilon = <$, where $<$ is the usual order on $\{1, 2 \ldots, n\}$. In particular $\upsilon \in \mathfrak{U}_{t^*_\upsilon(\alpha)}$.*

THEOREM 3.5.

(1) *The system (X, G) is minimal.*
(2) *The system (X, G) is extremely proximal; i.e. for every closed set $\emptyset \neq F \subsetneq X$ there exists a net $\{g_i\}_{i \in I}$ in G such that we have $\lim_{i \in I} g_i F = \{x_0\}$ for some point $x_0 \in X$ (see* **[Gla74]***).*
(3) *The minimal system (X, G) is not isomorphic to the universal minimal system $(M(G), G)$.*
(4) *$(\Phi(X), G)$ is minimal.*
(5) *$(\Phi(X), G)$ is proximal.*

PROOF.

(1) Since X is h-homogeneous, then by Proposition 2.2, G acts transitively on non-trivial (i.e. not $\emptyset, X$) clopen sets. Since G acts transitively on the above mentioned basis, it follows that for every $U \in \mathcal{U}$ we have $\cup\{\alpha(U) : \alpha \in G\} = X$. This property is equivalent to the minimality of the system (X, G).
(2) Fix some x_0 in X such that $x_0 \notin F$. For an arbitrary basic clopen neighborhood $U = A$ of x_0 which is disjoint from F choose $\alpha_U \in G$ such that $\alpha_U(A^c) = A$. Then α satisfies $\alpha_U(F) \subset U$. Clearly now $\{\alpha_U : U \text{ a neighborhood of } x_0\}$ is the required net.

(3) As the system (X, G) is certainly 3-transitive this claim follows from Uspenskij's theorem [**Usp00**]. For completeness we provide a direct proof. Suppose (X, G) is isomorphic to the universal minimal G system. Let $Y \subset \Phi$ be a minimal subset of Φ. Then, by the coalescence of the universal minimal system (every G-endomorphism $\phi : (M(G), G) \to (M(G), G)$ (which is necessarily onto) is an isomorphism, see [**GL11**] and [**Usp00**]), the restriction $\pi : Y \to X$, sending a chain to its root, is an isomorphism. Fix $c_0 \in Y$ and let $p_0 \in X$ be its root; i.e. $\pi(c_0) = p_0$. Let $H = \{\alpha \in G : \alpha p_0 = p_0\}$, the stabilizer of p_0. Since π is an isomorphism we also have $H = \{\alpha \in G : \alpha c_0 = c_0\}$. Choose $F \in c_0$ such that $\{p_0\} \subsetneqq F \subsetneqq X$ and let $p_0 \neq a \in F$ (recall X is infinite). Choose a clopen partition of (P, A, B) of X with $B \cap F = \emptyset$, $P \cap F \neq \emptyset$ and $A \cap F \neq \emptyset$. Using the fact that X is partition homogeneous, one can find $g \in G$ so that $gP = P$, $gA = B$ and $gB = A$. One redefines g so that $g_{|P} = Id$. As $g(A \cup P) \cap A = \emptyset$, we have $F \setminus gF \neq \emptyset$. As $gA = B$ we have $gF \setminus F \neq \emptyset$. Conclude that F and gF are not comparable. On the other hand $g(p_0) = p_0$ means $g \in H$ whence also $gc_0 = c_0$. In particular $gF \in c_0$ and as c_0 is a chain one of the inclusions $F \subset gF$ or $gF \subset F$ must hold. This contradiction shows that (X, G) cannot be the universal minimal G-system.

(4) Let $\upsilon', \upsilon \in \Phi(X)$ and $\upsilon' \in \mathfrak{U}_\alpha$ for some $\alpha = (A_1, A_2 \ldots, A_n) \in \mathcal{D}$. Let $<_\upsilon$ be the induced order of υ on α. Let $\sigma \in S_n$ be such that for any $i < j$, $A_{\sigma(i)} <_\upsilon A_{\sigma(j)}$. As X is partition homogeneous we can choose $g \in G$ so that $gA_{\sigma(i)} = A_i$. Clearly $g\upsilon \in \mathfrak{U}_\alpha$.

(5) Let $\upsilon_1, \upsilon_2 \in \Phi(X)$. Fix some $\upsilon' \in \mathfrak{U}_\alpha$ for some $\alpha = (A_1, A_2 \ldots, A_n) \in \mathcal{D}$. Let $<$ be the usual order on $\{1, 2 \ldots, n\}$. Inductively we will construct $g \in G$ so that $<_{g\upsilon_1|\alpha} = <_{g\upsilon_2|\alpha} = <$. Using Lemma 3.4, this implies $g\upsilon_1 \in \mathfrak{U}_\alpha$ and $g\upsilon_2 \in \mathfrak{U}_\alpha$. As $\mathfrak{U}_\alpha$ is arbitrary, this establishes proximality. Indeed let $g_1 \in G$ so that $g_1(r(\upsilon_1)), g_1(r(\upsilon_2)) \in A_1$. Assume we have constructed $g_k \in G$. Define $g_{k+1} \in G$ so that $g_{k+1|A_1 \cup A_2 \cup \ldots \cup A_k} = g_{k|A_1 \cup A_2 \cup \ldots \cup A_k}$ and $g_{k+1}(r((g_k\upsilon_1)_{|(A_{i_1} \cup A_{i_2} \cup \ldots \cup A_{i_k})^c}), g_{k+1}(r(g_k\upsilon_2)_{|(A_{i_1} \cup A_{i_2} \cup \ldots \cup A_{i_k})^c}) \in A_{i_{k+1}}$. It is easy to see that $g = g_n$ has the desired properties.

□

4. Calculation of the universal minimal space

4.1. Overview. The goal of this section is to generalize the main theorem of [**GW03**]: the universal minimal space of the group of homeomorphisms of the Cantor set, equipped with the compact-open topology, is the space of maximal chains over the Cantor set. We prove the following theorem:

THEOREM 4.1. *Let X be a h-homogeneous zero-dimensional compact Hausdorff topological space. Let $G = \mathrm{Homeo}(X)$ equipped with the compact-open topology, then $M(G) = \Phi(X)$.*

The proof borrows heavily from the proof in [**GW03**]. The new ideas (that build on ideas in [**GW03**]) are presented in subsections 4.2, 4.3, 4.5

4.2. Order topology. Recall that given a set Y and a linear order $<$ on Y there is a topology generated by the basis of open intervals $(a, b) = \{y \in Y : a < y < b\}$ where $a, b \in Y$ and equality is allowed on the left (right) if a (b) is the smallest (biggest) element of Y. This topology is called the **order topology** on $(Y,<)$. For more details see [**Mun75**] Section 2.3. One of the most important ingredients in the proof in [**GW03**] is the fact that the topology on the Cantor set K is the order topology associated with the natural order $<$ on $K \subset [0,1]$). A natural approach to generalizations of the result in the case of $X = \omega^*$ the corona, is to look for an order that will generate the topology on the corona. However, as the following proposition shows this is impossible.

PROPOSITION 4.2. *The topology on ω^* is not an order topology.*

PROOF. Assume for a contradiction that the topology on ω^* is an order topology associated with a linear order $<$. As ω^* has no isolated points we can find (with no loss of generality) an increasing bounded sequence of points $p_1 < p_2 < p_3 < \cdots < b$. By compactness this sequence admits a least upper bound $p = l.u.b\ \{p_k : k = 1, 2, \dots\}$. It is easy to check that $p = \lim_{k\to\infty} p_k$, so that the set $\{p_k : k = 1, 2, \dots\} \cup \{p\}$ is a closed subset of ω^*. However, it is well known that the remainder ω^* has no nontrivial converging sequences; e.g. one can use the fact that the closure of the set $\{p_k : k = 1, 2, \dots\}$, like the closure of any infinite discrete countable set in $\beta\omega$, is homeomorphic to $\beta\omega$ (see e.g. [**Eng78**, Theorem 3.6.14]). $\square$

4.3. The spaces Ω_k and $\tilde{\Omega}_k$ and a cocycle equation. The following subsection is a generalization of Section 3 of [**GW03**]. Fix $\alpha = (A_1, A_2 \dots, A_k) \in \mathcal{D}_k$ and define the clopen subgroup $H_\alpha = \{g \in G : gA_i = A_i, i = 1, \dots, k\} \subset G$. Consider the discrete homogeneous space of right cosets $H_\alpha \backslash G = \{H_\alpha g : g \in G\}$. There is a natural bijection $\phi : H_\alpha \backslash G \to \mathcal{D}_k$ given by $\phi(H_\alpha g) = g^{-1}\alpha$. Let $\Omega_k = \{1, -1\}^{\mathcal{D}_k}$ equipped with the product topology. This is a G-space under the action $g\omega(\beta) = \omega(g^{-1}\beta)$ for any $\omega \in \Omega_k$, $\beta \in \mathcal{D}_k$ and $g \in G$.

Set $\mathcal{T}^k = \{1, -1\}^{S_k}$. We refer to the elements of $\mathcal{T}^k$ as **tables**. Denote $\tilde{\Omega}_k = (\mathcal{T}^k)^{\tilde{\mathcal{D}}_k}$ equipped with the product topology. This is a G-space under the action $\cdot : G \times \tilde{\Omega}_k \to \tilde{\Omega}_k$ given by $g \cdot \tilde{\omega}(\tilde{\beta})(\sigma) = \tilde{\omega}(g^{-1}\tilde{\beta})(\sigma)$ for any $\omega \in \Omega_k$, $\tilde{\beta} \in \tilde{\Omega}_k$ and $g \in G$.

There is a natural family of homeomorphisms $\pi_c : \Omega_k \to \tilde{\Omega}_k$, $c \in \Phi(X)$ given by $\omega \mapsto \tilde{\omega}^c$, (also denoted $\tilde{\omega}$ when no confusion arises) where for $\tilde{\beta} = \{B_1, B_2 \dots, B_k\} \in \tilde{D}_k$ and $\sigma \in S_k$, $\tilde{\omega}(\tilde{\beta})(\sigma) = \omega(\sigma^{-1} t_c^*(\tilde{\beta}))$ ($t_c^*(\cdot)$ is defined after Definition 3.1). In order for π_c to be a G-homeomorphism we need to equip $\tilde{\Omega}_k$ with a different G-action than the natural G-action mentioned above. Namely $\bullet_c : G \times \tilde{\Omega}_k \to \tilde{\Omega}_k$, is defined by

$$g \bullet_c \tilde{\omega}(\tilde{\beta})(\sigma) = \tilde{\omega}(g^{-1}\tilde{\beta})(\rho_c(g, \tilde{\beta})\sigma) = \omega(\sigma^{-1}\rho_c(g, \tilde{\beta})^{-1} t_c^*(g^{-1}\tilde{\beta}))$$

where $\rho_c : G \times \tilde{\Omega}_k \to S_k$ is defined uniquely by the equation:

$$\rho_c(g, \tilde{\beta})^{-1} t_c^*(g^{-1}\tilde{\beta}) = g^{-1} t_c^*(\tilde{\beta})$$

As $g \bullet_c \tilde{\omega}(\tilde{\beta})(\sigma) = \omega(\sigma^{-1} g^{-1} t_c^*(\tilde{\beta}))$, we have the equality $g \bullet_c \tilde{\omega}(\tilde{\beta})(\sigma) = \widetilde{g\omega}(\tilde{\beta})(\sigma)$ which makes $\pi_c : (G, \Omega_k) \to (G, \tilde{\Omega}_k)$ a G-homeomorphism (and formally proves $g\bullet_c$ is indeed a G-action).

LEMMA 4.3. *$\rho_c : G \times \tilde{\Omega}_k \to S_k$ obeys the following* ***cocyle*** equation:

$$\rho_c(gh, \tilde{\beta}) = \rho_c(g, \tilde{\beta})\rho_c(h, g^{-1}\tilde{\beta})$$

PROOF. By definition we have $gh \bullet_c \tilde{\omega}(\tilde{\beta})(\sigma) = \widetilde{gh\omega}(\tilde{\beta})(\sigma) = g \bullet_c \widetilde{h\omega}(\tilde{\beta})(\sigma)$. Notice

$$gh \bullet_c \tilde{\omega}(\tilde{\beta})(\sigma) = \omega(\sigma^{-1}\rho_c(gh,\tilde{\beta})^{-1}t_c^*(h^{-1}g^{-1}\tilde{\beta})),$$

whereas

$$g \bullet_c \widetilde{h\omega}(\tilde{\beta})(\sigma) = h\omega(\sigma^{-1}\rho_c(g,\tilde{\beta})^{-1}t_c^*(g^{-1}\tilde{\beta})) = \omega(\sigma^{-1}h^{-1}\rho_c(g,\tilde{\beta})^{-1}t_c^*(g^{-1}\tilde{\beta})).$$

This implies

$$\rho_c(gh,\tilde{\beta})^{-1}t_c^*(h^{-1}g^{-1}\tilde{\beta}) = h^{-1}\rho_c(g,\tilde{\beta})^{-1}t_c^*(g^{-1}\tilde{\beta}).$$

As $\rho_c(h,g^{-1}\tilde{\beta})^{-1}t_c^*(h^{-1}g^{-1}\tilde{\beta}) = h^{-1}t_c^*(g^{-1}\tilde{\beta})$, we have

$$\rho_c(gh,\tilde{\beta})^{-1} = \rho_c(h,g^{-1}\tilde{\beta})^{-1}\rho_c(g,\tilde{\beta})^{-1}.$$

Taking the inverses we get $\rho_c(gh,\tilde{\beta}) = \rho_c(g,\tilde{\beta})\rho_c(h,g^{-1}\tilde{\beta})$ □

Note that in the end of Section 3 of [**GW03**] it was mistakenly claimed that $g \bullet_{c_0} \tilde{\omega}(\tilde{\beta})(\sigma) = g \cdot \tilde{\omega}(\tilde{\beta})(\sigma)$, for $c_0 = \{[0,t] \cap K\}_{t\in[0,1]}$ where K, the Cantor set, is embedded naturally in $[0,1]$.

4.4. The dual Ramsey theorem. A partition $\gamma = (C_1,\ldots,C_k)$ of $\{1,\ldots,s\}$ into k nonempty sets is **naturally ordered** if for any $1 \leq i < j \leq k$, $\min(C_i) < \min(C_j)$. We denote by $\Pi\binom{s}{k}$ the collection of naturally ordered partitions of $\{1,\ldots,s\}$ into k nonempty sets.

DEFINITION 4.4. Let $\beta = (B_1,\ldots,B_s) \in \Pi\binom{k}{s}$ and $\gamma = (C_1,\ldots,C_k) \in \Pi\binom{m}{k}$ define the **amalgamated partition** $\gamma_\beta = (G_1,\ldots,G_s) \in \Pi\binom{m}{s}$ by:

$$G_j = \bigcup_{i\in B_j} C_i$$

Notice γ_β is naturally ordered and $(\mathcal{P}_\gamma)_\beta = \mathcal{P}_{\gamma_\beta}$. Similarly for $\alpha = (A_1, A_2 \ldots, A_m) \in \mathcal{D}$ define the **amalgamated clopen cover** $\alpha_\gamma = (G_1, G_1 \ldots, G_k)$, where $G_j = \bigcup_{i\in C_j} A_i$. Notice that $(\alpha_\gamma)_\beta = \alpha_{\gamma_\beta}$.

We denote by $\tilde{\Pi}\binom{s}{k}$ the collection of unordered partitions of $\{1,\ldots,s\}$ into k nonempty sets. Notice there is a natural bijection $\tilde{\Pi}\binom{s}{k} \leftrightarrow \Pi\binom{s}{k}$.

THEOREM 4.5 (The dual Ramsey Theorem). *Given positive integers k,m,r there exists a positive integer $N = DR(k,m,r)$ with the following property: for any coloring of $\tilde{\Pi}\binom{N}{k}$ by r colors there exists a partition $\alpha = \{A_1, A_2, \ldots, A_m\} \in \tilde{\Pi}\binom{N}{m}$ of N into m non-empty sets such that all the partitions of N into k non-empty sets (i.e. elements of $\tilde{\Pi}\binom{N}{k}$) whose atoms are measurable with respect to α (i.e. each equivalence class is a union of elements of α) have the same color.*

PROOF. This is Corollary 10 of [**GR71**]. □

4.5. Minimal symbolic systems. In the beginning of Section 4 of [**GW03**] a family of mappings $\phi_T : (G,\Phi(X)) \to (G,\Omega_k), T \in \mathcal{T}^k$ are introduced. We will introduce a generalized family but using a different description.

DEFINITION 4.6. Let $\beta \in \mathcal{D}$ and $c \in \Phi(X)$, define the β**-ratio** of c, to be the unique element $\theta_\beta(c) \in S_k$ so that:

$$\theta_\beta(c)\beta = t_c^*(\beta)$$

LEMMA 4.7. *The following holds:*

(1) $\theta_\beta(c) = \theta_{g\beta}(gc)$ *for* $c \in \Phi(X)$, $g \in G$ *and* $\beta \in \mathcal{D}$.
(2) $\theta_{\sigma^{-1}t_c^*(\tilde{\beta})}(c) = \sigma$ *for* $\sigma \in S_k$, $\tilde{\beta} \in \tilde{\mathcal{D}}$ *and* $c \in \Phi(X)$.

Proof.

(1) By definition $\theta_{g\beta}(gc)g\beta = t^*_{gc}(g\beta)$. By Lemma 3.2, $gt^*_c(\beta) = t^*_{gc}(g\beta)$ and therefore one has $\theta_{g\beta}(gc)g\beta = gt^*_c(\beta)$. As the G and S_k actions commute it implies $\theta_{g\beta}(gc)\beta = t^*_c(\beta)$. By definition $\theta_\beta(c)\beta = t^*_c(\beta)$ and we conclude $\theta_\beta(c) = \theta_{g\beta}(gc)$.
(2) $\theta_{\sigma^{-1}t_c^*(\tilde{\beta})}(c)\sigma^{-1}t^*_c(\tilde{\beta}) = t^*_c(\sigma^{-1}t^*_c(\tilde{\beta}))$

□

Let $T \in \mathcal{T}^k$. Define $\phi_T : \Phi(X) \to \Omega_k$ by

$$\phi_T(c)(\beta) = T(\theta_\beta(c))$$

Lemma 4.8. *$\phi_T : \Phi(X) \to \Omega_k$ is continuous and G-equivariant.*

Proof. We start by showing that ϕ_T is continuous. Let $n \in \mathbb{N}$, $\epsilon_1, \epsilon_2, \dots \epsilon_n \in \{\pm 1\}$, $\beta_1, \beta_2, \dots \beta_n \in \mathcal{D}_k$. Let V be an open set of Ω_k so that $V = \{\omega \in \Omega_k : \omega(\beta_i) = \epsilon_i\}$ and assume $V \neq \emptyset$. Let $c_1 \in \phi_T^{-1}(V)$. Denote $\mathfrak{U} = \bigcap_{\mathrm{i=1}}^{\mathrm{n}} \mathfrak{U}_{t^*_{c_1}(\beta)}$. By Lemma 3.4 $c_1 \in \mathfrak{U}$ so $\mathfrak{U} \neq \emptyset$. We claim $\phi_T(\mathfrak{U}) \subset V$. Indeed let $c_2 \in \mathfrak{U}$ and fix i. By assumption $c_2 \in \mathfrak{U}_{t^*_{c_1}(\beta_i)}$. By Lemma 3.4 $c_2 \in \mathfrak{U}_{t^*_{c_2}(\beta_i)}$. Conclude $t^*_{c_1}(\beta) = t^*_{c_2}(\beta)$, which implies $\theta_{\beta_i}(c_1) = \theta_{\beta_i}(c_2)$. This in turn implies $\phi_T(c_1)(\beta_i) = \phi_T(c_2)(\beta_i) = \epsilon_i$.

To show G-equivariance one has to show $g\phi_T(c)(\beta) = \phi_T(c)(g^{-1}\beta) = \phi_T(gc)(\beta)$. By definition $\phi_T(c)(g^{-1}\beta) = T(\theta_{g^{-1}\beta}(c))$ whereas $\phi_T(gc)(\beta) = T(\theta_\beta(gc))$. By Lemma 4.7 $\theta_\beta(gc) = \theta_{g^{-1}\beta}(c)$. □

Let $c_0 \in \Phi(X)$. We will investigate $\pi_{c_0} \circ \phi_T$. By definition $\tilde{\omega}^{c_0}(\tilde{\beta})(\sigma) = \omega(\sigma^{-1}t^*_c(\tilde{\beta}))$ and therefore we have $\widetilde{\phi_T(c)}^{c_0}(\tilde{\beta})(\sigma) = T(\theta_{\sigma^{-1}t^*_{c_0}(\tilde{\beta})}(c))$. By Lemma 4.7

$$\widetilde{\phi_T(c_0)}^{c_0}(\tilde{\beta})(\sigma) = \sigma$$

In particular $\widetilde{\phi_T(c_0)}^{c_0}(\tilde{\beta})(\sigma)$ *does not depend* on $\tilde{\beta}$ and we denote it by $\tilde{\omega}_T$.

The following theorem is based on Theorem 4.1 of [**GW03**]:

Theorem 4.9. *Every minimal subsystem of (G, Ω_k) is a factor of $(G, \Phi(X))$.*

Proof. Fix a minimal subset $\Sigma \subset \Omega_k$. We shall construct a homomorphism $\phi : (G, \Phi(X)) \to (G, \Sigma)$. Moreover it will be shown that $\phi = \phi_T$ for some $T \in \mathcal{T}^k$. Fix a point $\omega \in \Sigma$ and $c_0 \in \Phi(X)$. We consider $\tilde{\omega}^{c_0}$ as a coloring of elements of $\tilde{\mathcal{D}_k}$ by $r = |\mathcal{T}^k|$ where the colors are the tables of $\mathcal{T}^k$. For $\tilde{\beta} \in \mathcal{D}_k$, we thus denote by $\tilde{\omega}^{c_0}(\tilde{\beta})$ the element in $\mathcal{T}^k$. Let $m \in \mathbb{N}$ and fix $\alpha \in \mathcal{D}_m$. Let $\beta \in \mathcal{D}$ such that $\alpha \preceq \beta$, $t^*_{c_0}(\beta) = \beta$ and $|\beta| = N = DR(k, m, r)$ as in Theorem 4.5.

We define the coloring map to be $f : \Pi\binom{N}{k} \to \mathcal{T}^k$ where $\gamma \hookrightarrow \tilde{\omega}^{c_0}(\tilde{t}(\beta_\gamma))$. According to Theorem 4.5 there exists $\eta \in \Pi\binom{N}{m}$ and $T_\alpha \in \mathcal{T}^k$ so that for any $\tau \in \Pi\binom{m}{k}$, $f(\eta_\tau) = T_\alpha$. Let $g_\alpha \in G$ be such that $g_\alpha^{-1}t^*_{c_0}(\alpha) = \beta_\eta$. Denote $\tilde{\omega}^{c_0}_{g_\alpha} = g_\alpha \bullet_{c_0} \tilde{\omega}^{c_0}$. Notice

$$\begin{aligned}
\tilde{\omega}^{c_0}_{g_\alpha}(\tilde{t}(t^*_{c_0}(\alpha)_\tau))(\sigma) &= \omega(\sigma^{-1}g_\alpha^{-1}(t^*_{c_0}(\alpha)_\tau)) \\
&= \omega(\sigma^{-1}(g_\alpha^{-1}t^*_{c_0}(\alpha))_\tau) \\
&= \omega(\sigma^{-1}(\beta_\eta)_\tau) = \omega(\sigma^{-1}\beta_{\eta_\tau})
\end{aligned}$$

for any $\tau \in \Pi\binom{m}{k}$. We also have

$$T_\alpha = f(\eta_\tau) = \tilde{\omega}^{c_0}(\tilde{t}(\beta_{\eta_\tau}))(\sigma) = \omega(\sigma^{-1}t^*_{c_0}(\tilde{t}(\beta_{\eta_\tau}))) = \omega(\sigma^{-1}\beta_{\eta_\tau})$$

as $t^*_{c_0}(\beta) = \beta$. Conclude $\tilde{\omega}^{c_0}_{g_\alpha}(\tilde{t}(t^*_{c_0}(\alpha)_\tau)) = T_\alpha$. Let $\tilde{\upsilon} \in \Sigma$ be an accumulation point of the net $\{\tilde{\omega}^{c_0}_{g_\alpha}\}_{\alpha\in\mathcal{D}}$. Let $\tilde{\xi}_1, \tilde{\xi}_2 \in \tilde{\mathcal{D}}_k$. Let α be a common ordered refinement. By the calculations we have just performed for any $\gamma \succeq \alpha$, $\tilde{\xi}_1 = \tilde{t}(t^*_{c_0}(\gamma)_{\tau_1})$ and $\tilde{\xi}_2 = \tilde{t}(t^*_{c_0}(\gamma)_{\tau_2})$ for some $\tau_1, \tau_2 \in \Pi(\binom{|\gamma|}{k})$, we have $\tilde{\omega}^{c_0}_{g_\gamma}(\tilde{\xi}_1) = \tilde{\omega}^{c_0}_{g_\gamma}(\tilde{\xi}_2)$. This implies there exists $T \in \mathcal{T}^k$ such that for any $\tilde{\xi} \in \tilde{\mathcal{D}}_k$, $\tilde{\upsilon}(\tilde{\xi}) = T$, i.e. $\tilde{\nu} = \tilde{\omega}_T$ defined above. We conclude $\Sigma = \phi_T(\Phi(X))$. $\square$

4.6. Calculation of the universal minimal space. We now proceed as in [**GW03**].

LEMMA 4.10. *If Y is zero-dimensional compact Hausdorff topological space then the topological group* Homeo(Y) *equipped with the compact-open topology has a clopen basis at the identity.*

PROOF. See the proof of Lemma 3.2 of [**MS01**]. The clopen basis is given by $\{H_\alpha\}_{\alpha\in\mathcal{D}}$ where H_α is defined in Subsection 4.3. $\square$

THEOREM 4.11. *Let H be a topological group. If the topology of H admits a basis for neighborhoods at the identity consisting of clopen subgroups, then $M(H)$ is zero dimensional.*

PROOF. This follows from Proposition 3.4 of [**Pes98**] where it is shown that under the same conditions the **greatest ambit** of H is zero-dimensional. $\square$

We now give the proof of Theorem 4.1:

PROOF. The proof is a reproduction of the proof appearing in [**GW03**] that $M(G) = \Phi(K)$, where K is the Cantor set and $G = \text{Homeo}(K)$ is equipped with the compact-open topology. By Theorem 3.5 $(G, \Phi(X))$ is minimal and therefore there is an epimorphism $\pi : (G, M(G)) \to (G, \Phi(X))$. Fix $c_0 \in \Phi(X)$ and let $m_0 \in M(G)$ so that $\pi(m_0) = c_0$. By Lemma 4.10 and Theorem 4.11 $M(G)$ is zero-dimensional. Let $D \subset M(G)$ be a clopen subset and define the continuous function $F_D = 2\mathbf{1}_D - \mathbf{1}$, where $\mathbf{1}_D$ is the indicator function of D. If $H = \{g \in G : gD = D\}$ then H is a clopen subgroup of G and hence it contains H_α for some $\alpha \in \mathcal{D}_k$ for some $k \in \mathbb{N}$ (see proof of Lemma 4.10). It follows that the map $\psi_D(m) = (F_D(gm))_{g\in G}$, $m \in M(G)$ can be defined as a mapping into $\{1, -1\}^{H_\alpha\backslash G} = \Omega_k$ and thus we have $\psi_D : (G, M(G)) \to (G, \Omega_k)$, so that if we set $Y_D = \psi_D(M(G))$, the system (Y_D, G) is a minimal symbolic subsystem of Ω_k. Denote $y_D = \psi_D(m_0)$.

Apply Theorem 4.9 to define a G-homomorphism $\phi_D : \Phi \to \Omega_k$, with $y'_D = \phi_D(c_0)$. Given a clopen subset $D \subset M(G)$ consider the following diagram:

$$\begin{tikzcd}
(M(G), m_0) \arrow[r, "\pi"] \arrow[d, "\psi_D"'] & (\Phi, c_0) \arrow[d, "\phi_D"] \\
(Y_D, y_D) & (Y_D, y'_D)
\end{tikzcd}$$

The image $(\psi_D \times (\phi_D \circ \pi))(M(G), m_0) = (W, (y_D, y'_D))$, with $W \subset Y_D \times Y_D$, is a minimal subset of the product system $(Y_D \times Y_D, G)$. By Theorem 3.5(5) (Y_D, G) is proximal. Therefore the diagonal $\Delta = \{(y, y) : y \in Y_D\}$ is the unique minimal subset of the product system and we conclude that $y_D = y'_D$, so that the above diagram is replaced by

$$\begin{tikzcd}
(M(G), m_0) \arrow[rr, "\pi"] \arrow[dr, "\psi_D"'] & & (\Phi, c_0) \arrow[dl, "\phi_D"] \\
& (Y_D, y_D) &
\end{tikzcd}$$

Next form the product space

$$\Pi = \prod \{Y_D : D \text{ a clopen subset of } M(G)\},$$

and let $\psi : M(G) \to \Pi$ be the map whose D-projection is ψ_D (i.e. $(\psi(m))_D = \psi_D(m)$). We set $Y = \psi(M(G))$ and observe that, since clearly the maps ψ_D separate points on $M(G)$, the map $\psi : M(G) \to Y$ is an isomorphism, with $\psi(m_0) = y_0$, where $y_0 \in Y$ is defined by $(y_0)_D = y_D$. Likewise define $\phi : \Phi(X) \to Y$ by $(\phi(m))_D = \phi_D(m)$, so that also $\phi(c_0) = y_0$. These equations force the identity $\psi = \phi \circ \pi$ in the diagram

$$\begin{tikzcd}
(M(G), m_0) \arrow[rr, "\pi"] \arrow[dr, "\psi"'] & & (\Phi, c_0) \arrow[dl, "\phi"] \\
& (Y, y_0) &
\end{tikzcd}$$

Since ψ is a bijection it follows that so are π and ϕ and the proof is complete. □

References

[BS81] Stanley Burris and H. P. Sankappanavar. *A course in universal algebra*, volume 78 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1981. MR648287 (83k:08001)

[DM78] Eric K. van Douwen and Jan van Mill. Parovičenko's characterization of $\beta\omega - \omega$ implies CH. *Proc. Amer. Math. Soc.*, 72(3):539–541, 1978. MR509251 (80b:04007)

[Ell69] Robert Ellis. *Lectures on topological dynamics*. W. A. Benjamin, Inc., New York, 1969. MR0267561 (42:2463)

[Eng78] Ryszard Engelking. *Dimension theory*. North-Holland Publishing Co., Amsterdam, 1978. Translated from the Polish and revised by the author, North-Holland Mathematical Library, 19. MR0482697 (58:2753b)

[GL11] Yonatan Gutman and Hanfeng Li. A new short proof of the uniqueness of the universal minimal space. preprint, 2011. http://arxiv.org/abs/1103.2647

[Gla74] Shmuel Glasner. Topological dynamics and group theory. *Trans. Amer. Math. Soc.*, 187:327–334, 1974. MR0336723 (49:1496)

[GR71] R. L. Graham and B. L. Rothschild. Ramsey's theorem for n-parameter sets. *Trans. Amer. Math. Soc.*, 159:257–292, 1971. MR0284352 (44:1580)

[Gut08] Yonatan Gutman. Minimal actions of homeomorphism groups. *Fund. Math.*, 198(3):191–215, 2008. MR2391011 (2008k:57061)

[GW02] E. Glasner and B. Weiss. Minimal actions of the group $\mathbb{S}(\mathbb{Z})$ of permutations of the integers. *Geom. Funct. Anal.*, 12(5):964–988, 2002. MR1937832 (2003h:37009)

[GW03] E. Glasner and B. Weiss. The universal minimal system for the group of homeomorphisms of the Cantor set. *Fund. Math.*, 176(3):277–289, 2003. MR1992824 (2004d:37010)

[HNV04] Klaas Pieter Hart, Jun-iti Nagata, and Jerry E. Vaughan, editors. *Encyclopedia of general topology*. Elsevier Science Publishers B.V., Amsterdam, 2004. MR2049453 (2005d:54001)

[KPT05] A. S. Kechris, V. G. Pestov, and S. Todorcevic. Fraïssé limits, Ramsey theory, and topological dynamics of automorphism groups. *Geom. Funct. Anal.*, 15(1):106–189, 2005. MR2140630 (2007j:37013)

[Kri67] Saul A. Kripke. An extension of a theorem of Gaifman-Hales-Solovay. *Fund. Math.*, 61:29–32, 1967. MR0220641 (36:3693)

[MS01] Michael G. Megrelishvili and Tzvi Scarr. The equivariant universality and couniversality of the Cantor cube. *Fund. Math.*, 167(3):269–275, 2001. MR1815091 (2002b:54039)

[Mun75] James R. Munkres. *Topology: a first course*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1975. MR0464128 (57:4063)

[Par63] I. I. Parovičenko. On a universal bicompactum of weight $\aleph$. *Dokl. Akad. Nauk SSSR*, 150:36–39, 1963. MR0150732 (27:719)

[Pes98] Vladimir G. Pestov. On free actions, minimal flows, and a problem by Ellis. *Trans. Amer. Math. Soc.*, 350(10):4149–4165, 1998. MR1608494 (99b:54069)

[ŠR89] Petr Štěpánek and Matatyahu Rubin. Homogeneous Boolean algebras. In *Handbook of Boolean algebras, Vol. 2*, pages 679–715. North-Holland, Amsterdam, 1989.

[Usp00] Vladimir Uspenskij. On universal minimal compact G-spaces. In *Proceedings of the 2000 Topology and Dynamics Conference (San Antonio, TX)*, volume 25, pages 301–308, 2000. MR1875600 (2002j:22024)

School of Mathematical Sciences, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel

E-mail address: glasner@math.tau.ac.il

Laboratoire d'Analyse et de Mathématiques Appliquées, Université de Marne-la-Vallée, 5 Boulevard Descartes, Cité Descartes - Champs-sur-Marne, 77454 Marne-la-Vallée cedex 2, France

E-mail address: yonatan.gutman@univ-mlv.fr

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11232

Generic eigenvalues, generic factors and weak disjointness

Wen Huang and Xiangdong Ye

This paper is dedicated to Anatolii Stepin for his 70th Birthday.

Abstract. Generic factors of topological dynamical systems are introduced and their properties are discussed. It turns out that a transitive system is weakly scattering, i.e. weakly disjoint from any minimal equicontinuous system iff it has no non-trivial equicontinuous generic factors. Other characterizations of weakly scattering are also obtained.

Meanwhile, it is proven that if a transitive system has no non-trivial almost equicontinuous generic factors, then it is strongly scattering. As applications, it is shown that for a large class of transitive systems including Banach-transitive systems, weakly scattering is equivalent to strongly scattering; and for E-systems, weakly scattering is equivalent to weakly mixing, which is a generalization of the classical result that a minimal system is weakly mixing iff it admits no non-trivial equicontinuous factors. For an almost equicontinuous system equivalent forms of weakly scattering and scattering are given. Using these equivalence forms it is proven that the existence of a complete pointed monothetic group which is minimally almost periodic and does not have the fixed point property is equivalent to that of an almost equicontinuous system which is weakly scattering yet not scattering.

1. Introduction

By a *topologically dynamical system* (X, T) (t.d.s.) we mean a compact metric space X with a continuous surjective map T from X to X. The eigenvalues of a minimal t.d.s. are important invariants to interrupt mixing properties of the system. It is well known that a minimal t.d.s. is weakly mixing iff it admits no non-trivial eigenvalues. For more general "indecomposable" systems, i.e. transitive t.d.s., the eigenvalues of the systems do not always reflect its mixing properties; for example a non-trivial transitive t.d.s. with a fixed point has no non-trivial eigenvalues but is not necessarily weakly mixing. So a natural question is how to generalize the notion of eigenvalue of a transitive t.d.s. such that "generalized eigenvalues" still characterize mixing properties of these systems.

In 1969, Keynes and Robertson [**19**] made a breakthrough. They defined the notion of generic eigenvalues for a t.d.s., which can be considered as the eigenvalues

2010 *Mathematics Subject Classification.* Primary 37B20, 37B05, 37B10, 37B40.

Key words and phrases. Generic factor, generic eigenvalue, almost equicontinuous, disjoint, weakly disjoint.

The first author is supported by NNSF of China (10911120388), Fok Ying Tung Education Foundation and the Fundamental Research Funds for the Central Universities.

The first and second authors are supported by NNSF of China (11071231).

of the system modulo the topologically negligible sets - the sets of first category or meager sets. When a system is minimal, the generic eigenvalues coincide with the eigenvalues. The authors proved that a transitive t.d.s. with an ergodic invariant measure of full support is weakly mixing iff it has no non-trivial generic eigenvalues. Moreover, they conjectured that the assumption that there is an ergodic invariant measure with full support is not necessary. Unfortunately, this conjecture is not true in general, for counterexamples see [**9, 5, 14**]. Thus it is natural to ask for which class of transitive t.d.s., is weakly mixing equivalently to having no non-trivial generic eigenvalues. In this paper, we will give a partial answer to the question, precisely we prove that any E-system (i.e. a transitive system having an invariant measure with full support) is in that class.

To this purpose, we first need to characterize transitive systems having no non-trivial generic eigenvalues using other terminology. Now, let us recall some basic definitions. Two t.d.s. are *weakly disjoint* if their product is transitive [**20**]. A t.d.s. is called (cf. [**7, 14**])

- *weakly mixing* if it is weakly disjoint from itself;
- *strongly scattering* if it is weakly disjoint from any E-system;
- *scattering* if it is weakly disjoint from any minimal system;
- *weakly scattering* if it is weakly disjoint from any minimal equicontinuous system.

It turns out that a transitive t.d.s. has no non-trivial generic eigenvalues iff it is weakly scattering. In the meantime, we show that the set of return times of a weakly scattering system is a Bohr-recurrent sequence, introduced by Katzenelson [**17**]. For the similar characterizations of scattering and strongly scattering, see [**14**].

We see easily

$$\text{weakly mixing} \subseteq \text{strongly scattering} \subseteq \text{scattering} \subseteq \text{weakly scattering}$$

from the definitions and a result of [**13**]. In [**14**], we showed that weakly mixing is strictly stronger than strongly scattering, but the question concerning the relation between strongly scattering, scattering and weakly scattering is still open. We remark that the question has an affinity with the question on the existence of a Bohr recurrent, non-recurrent sequence (cf. [**17, 23**]), as well as with the classical question on the existence of a syndetic set, S, of integers such that $S - S$ is not a Bohr neighborhood of zero (cf. [**21**]). By the characterizations of return times sets of several kinds of scattering systems, it is easy to see that if weakly scattering is not equivalent to scattering, then there exists a Bohr recurrent, non-recurrent sequence, and there exists a syndetic set, S, of integers such that $S - S$ is not a Bohr neighborhood of zero. That is, the above two open questions have negative answers. Hence to investigate the question whether scattering is equivalent to weakly scattering may help us solve finally the above two long open questions.

We make some progress on the question of whether the about three kinds of scattering properties are the same. By introducing a key notion: generic factors for a transitive t.d.s. (which follows partially from the idea of Weiss [**22**], i.e. considering the "factors" of a system under modulo the topologically negligible sets - the sets of first category or meager sets), we prove that a t.d.s. is weakly scattering iff it has no non-trivial equicontinuous generic factors; and if a transitive t.d.s. is not strongly scattering, then it has a non-trivial generic almost equicontinuous factor. As applications, we obtain that for a large class of transitive t.d.s. including Banach-transitive systems, weakly scattering is equivalent to strongly scattering; and for E-systems, weakly scattering is equivalent to weakly mixing.

Finally, we consider the question in the category of almost equicontinuous systems. By *a pointed monothetic group* we mean a pair (G, g), where G is a topological group with a compatible invariant metric and $g \in G$ generates a cyclic subgroup dense in G. A topological group G is called *minimally almost periodic* if the only continuous character on G is the trivial one, i.e. the constant function 1. There are numerous known examples of minimally almost periodic complete pointed monothetic groups (see for example [**1**]). A topological group G is said to *have the fixed point property* if each G-flow has a fixed point. A non-trivial complete pointed monothetic group having the fixed point property was found by Glasner [**11**] and (independently) Furstenberg and Weiss. The question if there exists a minimally almost periodic pointed monothetic group without the fixed point property, suggested by Glasner [**11**], remains open.

In [**5**], Akin and Glasner proved that an almost equicontinuous system (X, T) is scattering iff the set of transitive points of (X, T) forms a complete pointed monothetic group having the fixed point property under some compatible metric with the induced topology, and thus obtained a non-trivial almost equicontinuous scattering system by the existence of such a group. In this paper, we show that an almost equicontinuous system (X, T) is weakly scattering iff the set of transitive points of (X, T) forms a complete minimally almost periodic pointed monothetic group under some compatible metric with the induced topology. Moreover, we prove that the existence of a complete pointed monothetic group which is minimally almost periodic and does not have the fixed point property is equivalent to that of an almost equicontinuous system which is weakly scattering yet not scattering. Hence the question suggested by Glasner is equivalent to the existence of an almost equicontinuous system which is weakly scattering yet not scattering.

Acknowledgements: We thank the editors for inviting us to submit this paper. In particular, we thank L. Bowen for the careful reading of the manuscript and many suggestions for the writing.

2. Generic eigenvalues and weakly scattering

In this section we will characterize the transitive systems having no non-trivial generic eigenvalues.

2.1. Definitions and basic properties. Let X be a compact metric space and $GC(X)$ be the set of all bounded complex-valued functions which are continuous on a dense G_δ subset of X. For $f \in GC(X)$, let $C(f) = \{x \in X : f \text{ is continuous at } x\}$ and $||f|| = \sup\{|f(x)| : x \in C(f)\}$. It is easily seen that

$$||f|| = \sup\{|f(x)| : x \in D \subset C(f), D \text{ dense in } X\}.$$

Two functions f and g in $GC(X)$ are said to be *equal almost everywhere* ($f = g$ a.e.) if $||f - g|| = 0$ or $\{x \in X : f(x) = g(x)\}$ is a dense G_δ subset of X. Clearly, $f \neq g$ a.e. is equivalent to the existence of an opene (open and non-empty) subset U of X with $f(x) \neq g(x)$ for $x \in U$.

We say a t.d.s. (X, T) is *invertible* if T is a homeomorphism. Let (X, T) be an invertible t.d.s.. Consider the map $T^* : GC(X) \to GC(X)$ induced by T, i.e.

$T^*(f) = f \circ T$ ($f \in GC(X)$). Then T^* is onto and $||T^*f|| = ||f||$ ($f \in GC(X)$). T^* will be denoted by T when no confusion arises.

Definition 2.1. Let (X,T) be an invertible t.d.s.. A complex number $\lambda \in \mathbb{C}$ is called *a generic eigenvalue* of (X,T) if there exists $f \in GC(X)$ with $f \neq 0$ a.e. with $Tf = \lambda f$ a.e.. The function f is called *a generic eigenfunction* (associated with λ). The set of all generic eigenvalues of (X,T) is denoted by $Eig(X,T)$ or $Eig(T)$, and the set of all generic eigenfunctions associated with λ is denoted by $Eig(\lambda)$.

We remark that:

(1) The notion of a generic eigenvalue first appeared in [**18**].
(2) It is easy to see that $1 \in Eig(T)$ which is called the *trivial generic eigenvalue* of (X,T).
(3) $Eig(T) \subseteq \mathbb{T}$, where $\mathbb{T}$ is the unit circle in the complex plane. To see this let $\lambda \in Eig(T)$. Then there exists $f \in GC(X)$ with $f \neq 0$ a.e. such that $Tf = \lambda f$ a.e.. Hence $||f|| = ||Tf|| = |\lambda| ||f||$. Since $||f|| > 0$, $|\lambda| = 1$.

In the Definition 2.1, we require that T is invertible, as it may happen that $Tf \notin GC(X)$ for some $f \in GC(X)$. For a transitive t.d.s. we still can define its generic eigenvalues without the assumption of invertibility.

Recall that a t.d.s. (X,T) is *transitive* if for each pair of *opene* subsets U and V, $N(U,V) = \{n \in \mathbb{Z}_+ : U \cap T^{-n}V \neq \emptyset\}$ is non-empty. (X,T) is *(topologically) weakly mixing* if $(X \times X, T \times T)$ is transitive. (X,T) is *(topologically) strongly mixing* if for each pair of opene subsets U and V, $N(U,V)$ is cofinite. $x \in X$ is a *transitive point* if the orbit of x, $orb(x,T) = \{x, T(x), T^2(x), \ldots\}$, is dense in X. It is easy to see that if (X,T) is transitive then the set of all transitive points is a dense G_δ set of X (denoted by $Tran_T(X)$). If $Tran_T(X) = X$ then we say that (X,T) is *minimal*. It is well known that there is some minimal subsystem in any dynamical system (X,T), which is called a *minimal set* of X. Each point in a minimal set of X is called a *minimal point*. So for a transitive t.d.s. we have

Definition 2.2. Let (X,T) be a transitive t.d.s.. A complex number $\lambda \in \mathbb{C}$ is called a generic eigenvalue of (X,T) if there exists on $Tran_T(X)$ a continuous nonzero function f with $Tf = \lambda f$. The function f is called the generic eigenfunction (associated with λ). The notations $Eig(T)$ and $Eig(\lambda)$ are used in this case too.

Concerning the above definition we have the following remarks.

(1) The definition was introduced by Downarowicz in [**9**]. For an invertible transitive t.d.s., Definition 2.2 coincides with Definition 2.1.
(2) Recall that for a t.d.s. (X,T), if f is a nonzero continuous map from X to $\mathbb{C}$ with $f(Tx) = \lambda f(x)$ for any $x \in X$ and some fixed $\lambda \in \mathbb{C}$, then f is the eigenfunction of (X,T) associated with λ. It is not hard to see that for a minimal system (X,T), generic eigenvalues and generic eigenfunctions of (X,T) are just eigenvalues and eigenfunctions of (X,T).
(3) Let (X,T) be a transitive t.d.s.. Then $1 \in Eig(T)$, which is called the trivial generic eigenvalue of (X,T).
(4) $Eig(T) \subseteq \mathbb{T}$. To see this, let $\lambda \in Eig(T)$ and $f \in Eig(\lambda)$. For $x \in Tran_T(X)$ with $f(x) \neq 0$, there exists $n_k \to +\infty$ such that $T^{n_k}x \to x$. Hence $\lambda^{n_k} f(x) = f(T^{n_k}x) \to f(x)$, that is, $\lambda^{n_k} \to 1$. This shows that $|\lambda| = 1$, i.e. $\lambda \in \mathbb{T}$. Moreover, it is not hard to see that $|f(x)|$ has nonzero constant value on $Tran_T(X)$.

In 1969, the authors of [**19**] investigated the transitive systems (X,T) with $Eig(T) = \{1\}$. They proved that if a transitive t.d.s. (X,T) has a T-invariant ergodic measure μ with $\mathrm{supp}(\mu) = X$, then (X,T) is weakly mixing iff $Eig(T) = \{1\}$. Moreover, they conjectured that the conclusion is still valid for any transitive system. In the rest of the section, we characterize the transitive systems having no non-trivial generic eigenvalues, thus disproving this conjecture, see [**9, 5, 14**] for counterexamples. Further study concerning the conjecture will be carried out in Section 4.

We start with some lemmas. The first one is a result in [**9**], and we include a proof for completeness.

LEMMA 2.3. *Let (X,T) be a transitive t.d.s. and $x_0 \in Tran_T$. Then $\lambda \in Eig(X,T)$ iff $\lambda^{n_k} \to 1$ for every sequence $\{n_k\} \subset \mathbb{Z}_+$ for which $T^{n_k}x_0 \to x_0$.*

PROOF. Let $\lambda \in Eig(X,T)$ and f be a generic eigenfunction of (X,T) with the generic eigenvalue λ. For every sequence $\{n_k\} \subset \mathbb{Z}_+$ for which $T^{n_k}x_0 \to x_0$, one has

$$\lim_{k\to+\infty} \lambda^{n_k} = \lim_{k\to+\infty} \frac{f(T^{n_k}x_0)}{f(x_0)} = 1.$$

Conversely, assume $\lambda^{n_k} \to 1$ for every sequence $\{n_k\} \subset \mathbb{Z}_+$ for which $T^{n_k}x_0 \to x_0$. For any $x \in Tran_T(X)$, there exists $n_k \in \mathbb{Z}_+$ with $T^{n_k}x_0 \to x$ and $\lambda^{n_k} \to \lambda_1$. Define $f(x) = \lambda_1$. Now we show that $f(x)$ is independent of the choice of the sequence $\{n_k\}$. If fact, since $x \in Tran_T(X)$ there exist $s_l \in \mathbb{Z}_+$ and $\lambda_2 \in \mathbb{T}$ such that $T^{s_l}x \to x_0$ and $\lambda^{s_l} \to \lambda_2$. It is easy to see that there exist $k_i, l_i \in \mathbb{N}$ such that $T^{n_{k_i}+s_{l_i}}x_0 \to x_0$ and $\lambda^{n_{k_i}+s_{l_i}} \to \lambda_1\lambda_2$. By the assumption we know that $\lambda_1\lambda_2 = 1$. Hence $\lambda_1 = \frac{1}{\lambda_2}$, that is, $f(x) = \lambda_1$ is independent the choice of $\{n_k\}$. Hence $f : Tran_T(X) \to \mathbb{C}$ is well defined. Clearly, $f(Tx) = \lambda f(x), \forall x \in Tran_T(X)$. In the following we show that f is continuous.

Fix an $x \in Tran_T(X)$ and let $x_n \in Tran_T(X)$ with $x_n \to x$. For each $n \in \mathbb{N}$, there is $\{l_i^n\}$ with $T^{l_i^n}(x_0) \to x_n$. Choose a large i and set $k_n = l_i^n$ such that $d(T^{k_n}x_0, x_n) < \frac{1}{2^n}$ and $|f(x_n) - f(T^{k_n}x_0)| < \frac{1}{2^n}$. Then $T^{k_n}x_0 \to x$ and $\lim_{n\to+\infty} f(x_n) = \lim_{n\to+\infty} f(T^{k_n}x_0) = f(x)$. Therefore, f is continuous at x, and thus continuous on $Tran_T(X)$. This shows $\lambda \in Eig(X,T)$. $\square$

Let (X,T) and (Y,S) be two t.d.s.. A continuous map $\pi : X \to Y$ is called *a homomorphism or a factor map* of systems (X,T) and (Y,S) if it is onto and $\pi T = S\pi$. We say (X,T) is an *extension* of (Y,S) and (Y,S) is a *factor* of (X,T). If π is also injective then it is called an *isomorphism.* An extension $\pi : X \to Y$ is called *almost one-to-one*, if there exists a dense G_δ set $X_0 \subset X$ such that $\pi^{-1}\pi(x) = \{x\}$ for any $x \in X_0$. It is well known that if $\pi : (X,T) \to (Y,S)$ is an almost one-to-one extension, then (X,T) is transitive iff (Y,S) is transitive (cf. [**5**]). An extension $\pi : X \to Y$ is called *asymptotic*, if $\lim_{n\to+\infty} d(T^nx, T^ny) = 0$ when $\pi(x) = \pi(y)$.

The following lemma shows that under some extensions, the set of eigenvalues does not change.

LEMMA 2.4. *Let $\pi : (X,T) \to (Y,S)$ be a factor map between transitive t.d.s.. Then $Eig(Y,S) \subseteq Eig(X,T)$. If the extension π is almost one-to-one or asymptotic, then $Eig(Y,S) = Eig(X,T)$.*

PROOF. Since $\pi(Tran_T(X)) \subset Tran_S(Y)$, it follows that $Eig(Y,S) \subseteq Eig(X,T)$. Let $\lambda \in Eig(X,T)$. It remains to show that $\lambda \in Eig(Y,S)$ under the additional assumptions.

First let π be almost one-to-one. Set $X_0 = \{x \in X : \pi^{-1}\pi(x) = \{x\}\}$. Then X_0 is a dense G_δ subset of X. Take $x_0 \in X_0 \cap Tran_T(X)$ and let $y_0 = \pi(x_0)$. Obviously, $y_0 \in Tran_S(Y)$. For every sequence $\{n_k\} \subset \mathbb{Z}_+$ for which $S^{n_k} y_0 \to y_0$, because of the choice of X_0, one has $\lim_{k\to+\infty} T^{n_k} x_0 = x_0$. Moreover, since $\lambda \in Eig(X,T)$, $\lim_{k\to+\infty} \lambda^{n_k} = 1$ by Lemma 2.3. Using Lemma 2.3 again, $\lambda \in Eig(Y,S)$.

Now let π be asymptotic. Take $x_0 \in Tran_T(X)$ and let $y_0 = \pi(x_0)$. Obviously, $y_0 \in Tran_S(Y)$. Assume $\{n_k\}$ is a sequence such that $S^{n_k} y_0 \to y_0$. If $\lambda^{n_k} \not\to 1$, then there exists a subsequence $\{n_{k_i}\}$ of $\{n_k\}$ such that $T^{n_{k_i}} x_0 \to x_1$ and $\lambda^{n_{k_i}} \to \lambda_1 \neq 1$. Therefore, $\pi(x_1) = \pi(\lim_{i\to+\infty} T^{n_{k_i}} x_0) = \lim_{i\to+\infty} S^{n_{k_i}} y_0 = y_0$. That is, $x_0, x_1 \in \pi^{-1}(y_0)$, and hence $\lim_{n\to+\infty} d(T^n x_0, T^n x_1) = 0$. Thus $x_1 \in Tran_T(X)$ and $\lim_{i\to+\infty} T^{n_{k_i}} x_1 (= \lim_{i\to+\infty} T^{n_{k_i}} x_0) = x_1$. Now by Lemma 2.3, $\lambda^{n_{k_i}} \to 1$, a contradiction as $\lambda^{n_{k_i}} \to \lambda_1 \neq 1$. This shows that $\lambda^{n_k} \to 1$ for each sequence $\{n_k\} \subset \mathbb{Z}_+$ for which $S^{n_k} y_0 \to y_0$. Using Lemma 2.3 again we have $\lambda \in Eig(Y,S)$. $\square$

Let X be a compact metric space with a metric d and 2^X be the set of all nonempty closed subsets of X. For $A \in 2^X$ and $\epsilon > 0$, put $N(A,\epsilon) = \{x \in X : d(x,A) < \epsilon\}$. Now for $A, B \in 2^X$, define

$$H_d(A,B) = \inf\{\epsilon > 0 : A \subseteq N(B,\epsilon), B \subset N(A,\epsilon)\}.$$

Then $(2^X, H_d)$ is a compact metric space. When (X,T) is a t.d.s., there is a naturally induced map $\widetilde{T} : 2^X \to 2^X$ by $\widetilde{T}(A) = \{Tx : x \in A\}$, and $(2^X, \widetilde{T})$ is also a t.d.s.. If (X,T) is invertible, so is $(2^X, \widetilde{T})$. The following result is well known, see for example [**6**, Lemma 44] in chapter 14.

Lemma 2.5. *Let X, Y be compact metric spaces and $\pi : X \to Y$ be continuous and onto. Then $\pi^{-1} : Y \to 2^X$ with $y \mapsto \pi^{-1}(y)$ is continuous on a dense G_δ subset of Y.*

A t.d.s. (X,T) is *equicontinuous*, if for every $\epsilon > 0$ there exists $\delta > 0$ such that whenever $d(x,y) < \delta$, $d(T^n x, T^n y) < \epsilon$ for each $n \in \mathbb{Z}_+$. A t.d.s. (X,T) is *topologically ergodic* if $N(U,V)$ is syndetic (has bounded gaps) for each pair of opene subsets of X. It is clear that a minimal system is an E-system, and an E-system is topologically ergodic [**13**]. In the introduction we introduced the notion of scattering, strongly scattering and weakly scattering. Now recall that (cf. [**7, 14, 16**]) a t.d.s. (X,T) is

- *mildly mixing* if it is weakly disjoint from any transitive system;
- *extremely scattering* if it is weakly disjoint from any topological ergodic system;
- *totally transitive* if it is weakly disjoint from any periodic system.

By [**5, 14, 15, 16**], we have

strongly mixing $\subsetneqq$ mild mixing $\subsetneqq$ weakly mixing $\subsetneqq$ extreme scattering $\subsetneqq$
strongly scattering $\subset$ (=?)scattering $\subset$ (=?)weakly scattering $\subsetneqq$ total transitivity.

For a t.d.s. (X,T) with a metric d, we say $(\widetilde{X}, \sigma_T)$ is the *natural extension* of (X,T), if $\widetilde{X} = \{(x_1, x_2, \cdots) : T(x_{i+1}) = x_i, x_i \in X, i \in \mathbb{N}\}$, which is a subspace of the product space $\Pi_{i=1}^{\infty} X$ with the compatible metric d_T defined by

$$d_T((x_1, x_2, \cdots), (y_1, y_2, \cdots)) = \sum_{i=1}^{\infty} \frac{d(x_i, y_i)}{2^i}.$$

Moreover, $\sigma_T : \widetilde{X} \longrightarrow \widetilde{X}$ is the shift homeomorphism, i.e. $\sigma_T(x_1, x_2, \cdots) = (T(x_1), x_1, x_2, \cdots)$. Let $\pi_i : \widetilde{X} \longrightarrow X$ be the projection to the i-th coordinate for each $i \in \mathbb{N}$.

LEMMA 2.6. *Let (X,T) be a transitive t.d.s. and $\pi_1 : (\widetilde{X}, \sigma_T) \longrightarrow (X,T)$ be the natural extension of (X,T). Then $Eig(\widetilde{X}, \sigma_T) = Eig(X,T)$.*

PROOF. It is easy to see that $(\widetilde{X}, \sigma_T)$ is transitive and π_1 is an asymptotic extension. By Lemma 2.4 $Eig(\widetilde{X}, \sigma_T) = Eig(X,T)$. □

2.2. A characterization of weakly scattering via generic eigenvalues. Now we give a characterization of weakly scattering via generic eigenvalues.

THEOREM 2.7. *Let (X,T) be a transitive t.d.s.. Then the following statements are equivalent:*

(1) *(X,T) is weakly scattering.*
(2) *(X,T) is totally transitive and weakly disjoint from any irrational rotation $(\mathbb{T}, R_\lambda)$.*
(3) *(X,T) has no non-trivial generic eigenvalues, i.e. $Eig(T) = \{1\}$.*

PROOF. (1) ⇒ (2) is clear.

(2) ⇒ (3). Assume that (X,T) is totally transitive and weakly disjoint from any irrational rotation $(\mathbb{T}, R_{\lambda_1})$, where $\lambda_1 = e^{2\pi i\alpha}, \alpha \notin \mathbb{Q}$ and $R_{\lambda_1}(z) = \lambda_1 z$ for $z \in \mathbb{T}$.

Let $\lambda \in Eig(T)$, $K = \mathrm{cl}(\{\lambda^n : n \in \mathbb{Z}_+\}) \subseteq \mathbb{T}$ and $R_\lambda : K \to K$ with $R_\lambda(k) = \lambda k, \forall k \in K$. Then (K, R_λ) is either a periodic system or an irrational rotation $(\mathbb{T}, R_\lambda)$. So $(X \times K, T \times R_\lambda)$ is transitive.

Since $\lambda \in Eig(T)$, there exists a non-zero continuous map $f : Trans_T(X) \to \mathbb{C}$ with $f(Tx) = \lambda f(x), \forall x \in Tran_T(X)$. Now define $g(x,k) = \frac{f(x)}{k} : Tran_T \times K \to \mathbb{C}$. Then g is a continuous map. Take a transitive point (x_0, k_0) of $(X \times K, T \times R_\lambda)$. As $g(T^n x_0, R_\lambda^n k_0) = g(x_0, k_0)$, one has $g \equiv g(x_0, k_0)$. In particular, $g(x, k_0) \equiv g(x_0, k_0), \forall x \in Tran_T$. This shows $f(x) \equiv f(x_0)$ for all $x \in Tran_T(X)$ and thus $\lambda = 1$. That is, (X,T) has no nontrivial generic-eigenvalues.

(3) ⇒ (1) Assume that (X,T) has no non-trivial generic eigenvalues. Without loss of generality, we assume that T is a homeomorphism (otherwise consider it's natural extension and use Lemma 2.6).

If (X,T) is not weakly scattering, then there exists a minimal equicontinuous system (Y,S) such that $(X \times Y, T \times S)$ is not transitive. Thus, there exists a proper subsystem $(X_1, T \times S)$ of $(X \times Y, T \times S)$ with $\mathrm{int}(X_1) \neq \emptyset$. Let $\pi_X : X_1 \to X$ and $\pi_Y : X_1 \to Y$ be the projections to X and Y respectively. Clearly, π_X, π_Y are surjective.

Let $\pi_X^{-1} : X \to 2^{X_1}$ with $x \mapsto \pi_X^{-1}(x)$. Then π_X^{-1} is continuous on a dense G_δ set D_0 of X by Lemma 2.5. Put $D = \bigcap_{i \in \mathbb{Z}} T^{-i} D_0 \cap Tran_T(X)$. Then D is an invariant dense G_δ set of X.

Set $H = \overline{\{\pi_X^{-1}(x) : x \in D\}} \subseteq 2^{X_1}$. It is easy to see that $\widetilde{T \times S}(H) = H$ and for any $A \in H$, $\pi_X(A)$ is a singleton denoted by $\pi_X(A) = \{x_A\}$. Then $A = \{x_A\} \times \pi_Y(A)$. Let $\hat{\pi}_X : H \to X$ with $\hat{\pi}_X(A) = x_A$. By the choice of D, $\hat{\pi}_X$ is an almost one-to-one map, so $(H, \widetilde{T \times S})$ is a transitive subsystem of $(2^{X_1}, \widetilde{T \times S})$.

Let $\hat{\pi}_Y : H \to 2^Y$ with $\hat{\pi}_Y(A) = \pi_Y(A)$ and let $H_Y = \hat{\pi}_Y(H)$. Then $\hat{\pi}_Y : (H, \widetilde{T \times S}) \to (H_Y, \widetilde{S})$ is a factor map. Since Y is minimal and equicontinuous, $(2^Y, \widetilde{S})$ is equicontinuous. Thus, $(H_Y, \widetilde{S})$ is minimal and equicontinuous since $(H_Y, \widetilde{S})$ is transitive.

Take any $A = \pi_X^{-1}(x) \in H$ with $x \in D$. Since $A \subseteq X_1$, $\hat{\pi}_Y(A) \neq Y$ otherwise $\{x\} \times Y \subseteq X_1$ implies $X_1 = X \times Y$. As Y is minimal, $S(\hat{\pi}_Y(A)) \neq \hat{\pi}_Y(A)$. Hence $(H_Y, \widetilde{S})$ is a non-trivial minimal equicontinuous system.

It is well known that there exists a non-trivial eigenvalue $\lambda \neq 1$ for the minimal equicontinuous system $(H_Y, \widetilde{S})$. Particularly, $\lambda \in Eig(H_Y, \widetilde{S})$. By Lemma 2.4, $\lambda \in Eig(H, \widetilde{T \times S}) = Eig(X, T)$ which contradicts the assumption $Eig(X, T) = \{1\}$. Thus (X, T) is weakly scattering. $\square$

Definition 2.8. [**17**] Let $S \subset \mathbb{Z}_+$.

(1) S is Bohr(n)-recurrent if for every translation R_g of $\mathbb{T}^n$ and $\epsilon > 0$ there are $g_0 \in \mathbb{T}^n$ and $s \in S \setminus \{0\}$ with $d(R_g^s g_0, g_0) < \epsilon$.

(2) S is Bohr-recurrent if for every equicontinuous system (Y, T) and $\epsilon > 0$ there are $y_0 \in Y$ and $s \in S \setminus \{0\}$ with $d(T^s y_0, y_0) < \epsilon$.

Denote the set of all Bohr(n)-recurrent sequences by $\mathcal{BR}(n)$, and the set of all Bohr-recurrent sequences by $\mathcal{BR}$.

Remark 2.9. Katznelson [**17**] proved that

$$\mathcal{BR}(1) \supsetneq \mathcal{BR}(2) \supsetneq \cdots \supsetneq \mathcal{BR}(n) \supsetneq \mathcal{BR}(n+1) \supsetneq \cdots\cdots \supsetneq \mathcal{BR}.$$

Letting (X, T) be a t.d.s., $x \in X$ and U, V two opene subsets of X, we define

$$N(x, U) = \{n \in \mathbb{Z}_+ : T^n x \in U\} \text{ and } N(U, V) = \{n \in \mathbb{Z}_+ : T^n U \cap V \neq \emptyset\},$$

and call $N(U, V)$ *the set of return times* of U, V. The following lemma is easy to obtain

Lemma 2.10. *Let $(X, T), (Y, S)$ be two transitive t.d.s.. Then (X, T) is weakly disjoint from (Y, S) iff for any opene subsets U, V of X and any opene subset W of Y, $N(U, V) \cap N(W, W) \neq \emptyset$.*

Now, we can characterize the sets of return times of weakly scattering systems.

Theorem 2.11. *Let (X, T) be a t.d.s.. Then the following conditions are equivalent.*

(1) *(X, T) is weakly scattering;*

(2) *For any opene subsets U, V of X, $N(U, V)$ is a Bohr(1)-recurrent sequence;*

(3) *For any opene subsets U, V of X, $N(U, V)$ is a Bohr-recurrent sequence.*

Proof. (1) $\Rightarrow$ (3) Let (X, T) be weakly scattering. Fix opene subsets U, V of X and let (Y, T) be a minimal equicontinuous system. For any $\epsilon > 0$, choose an opene subset W of Y with $diam(W) < \epsilon$. Since $(X \times Y, T \times S)$ is transitive, $N(U \times W, V \times W)$ is infinite. This shows that there exists $n \in N(U, V) \setminus \{0\}$ such that $S^{-n} W \cap W \neq \emptyset$. For $y_0 \in S^{-n} W \cap W$, we have $S^n y_0, y_0 \in W$, which implies $d(S^n y_0, y_0) < \epsilon$. Since (Y, S) is arbitrary, $N(U, V)$ is a Bohr-recurrent sequence.

(3) $\Rightarrow$ (2) is obvious. Now we show (2) $\Rightarrow$ (1). To do this it remains to prove that (X, T) is totally transitive and weakly disjoint from every irrational rotation $(\mathbb{T}, R_\lambda)$ according to Theorem 2.7.

First, we show that (X,T) is totally transitive. Fix opene subsets U,V of X and put $\lambda_n = e^{2\pi i \frac{1}{n}}$, $n \in \mathbb{N}$. Consider the translation R_{λ_n} on $\mathbb{T}$ and let d be the usually metric on $\mathbb{T}$. Choose $0 < \epsilon < \frac{1}{2n}$. Since $N(U,V)$ is Bohr(1)-recurrent, so there exists $s \in N(U,V) \setminus \{0\}$ and $\lambda_0 \in \mathbb{T}$ such that $d(R^s_{\lambda_n}\lambda_0, \lambda_0) < \epsilon$. That is, $d(\lambda_n^s, 1) < \epsilon$. Hence $n|s$ and thus $N(U,V) \cap n\mathbb{N} \neq \emptyset$. This implies (X,T) is totally transitive.

Now we show (X,T) is weakly disjoint from any irrational rotation $(\mathbb{T}, R_\lambda)$. For any opene subsets U,V of X and any opene subset W of $\mathbb{T}$ choose $\lambda_1 \in W$ and $\epsilon > 0$ such that $B(\lambda_1, \epsilon) \subset W$. Since $N(U,V)$ is a Bohr(1)-recurrent sequence, there exists $s \in N(U,V) \setminus \{0\}$ and $\lambda_0 \in \mathbb{T}$ with $d(R^s_\lambda \lambda_0, \lambda_o) < \epsilon$. Hence $d(R^s_\lambda \lambda_1, \lambda_1) = d(R^s_\lambda \lambda_0, \lambda_0) < \epsilon$, thus $s \in N(W,W) \cap N(U,V)$. By Lemma 2.10, (X,T) is weakly disjoint from $(\mathbb{T}, R_\lambda)$. $\square$

Remark 2.12. In [**14**], we characterized the sets of return times of scattering and strongly scattering systems. Recall that

- $W \subset \mathbb{Z}_+$ is a *Poincaré sequence* if for any m.p.s. $(X, \mathcal{B}, \mu, T)$ and $A \in \mathcal{B}$ with $\mu(A) > 0$, there is $n \in W \setminus \{0\}$ with $\mu(A \cap T^{-n}A) > 0$;
- $W \subset \mathbb{Z}_+$ is a *recurrent sequence* if for any t.d.s. (X,d,T) and any $\epsilon > 0$ there are $x_0 \in X$ and $n \in W \setminus \{0\}$ with $d(T^n x_0, x_0) < \epsilon$.

Now, we have by [**14**]

(1) (X,T) is scattering iff $N(U,V)$ is a recurrent sequence for any opene subsets U,V of X;
(2) (X,T) is strongly scattering iff $N(U,V)$ is a Poincaré sequence for any opene subsets U,V of X.

By the definitions, a recurrent sequence is a Bohr recurrent one. The question if a Bohr recurrent sequence is a recurrent one is a long open problem (cf. [**17, 23**]), as is the classical question if there is a syndetic set, S, of integers such that $S-S$ is not a Bohr neighborhood of zero (cf. [**21**]). By Theorem 2.11 and the above remark, we know that if weakly scattering and scattering are not the same properties then the above two open questions have negative answers.

2.3. Disjointness and weakly disjointness. In this subsection we will show that a weakly scattering t.d.s. is disjoint from all minimal semi-distal t.d.s.. We remark that in [**7**] the authors showed that a scattering t.d.s. is disjoint from all minimal distal t.d.s..

The notion of *disjointness* of two t.d.s. was introduced by Furstenberg [**10**]. If (X,T) and (Y,S) are two t.d.s. we say $J \subset X \times Y$ is a *joining* of X and Y if J is a non-empty closed invariant set and is projected onto X and Y. If each joining is equal to $X \times Y$ then we say that (X,T) and (Y,S) are *disjoint*, denoted by $(X,T) \perp (Y,S)$ or $X \perp Y$. It is known that if $(X,T) \perp (Y,S)$ then one of them is minimal [**10**], and if (X,T) is minimal then the set of recurrent points of (Y,S) is dense [**15**]. Recall that a dynamical property is *residual* ([**5**]) it is inherited by factor maps, almost one-to-one extensions and inverse limits. We remark that weakly disjointness and disjointness are residual properties [**5**].

Recall that a t.d.s. (X,T) is *semi-distal* if each idempotent in the the adherence semigroup is minimal, or equivalently each recurrent point in $X \times X$ is minimal.

To show the next theorem we need a lemma related to semi-distality.

Lemma 2.13. *If (Y,S) is semi-distal then $(2^Y, \widetilde{S})$ is also semi-distal.*

Proof. Let (Y,S) be a semi-distal system, then $(Z,W)=\prod_{y\in Y}(Y,S)$ is also semi-distal. To see this, let $(x,y)\in Z\times Z$ be recurrent. Then there is an idempotent u such that $u(x,y)=(x,y)$. Since u is minimal we conclude that (x,y) is minimal.

Define $\Phi:\prod_{y\in Y}Y\to 2^Y$ with $\Phi\left((x(y))_{y\in Y}\right)=\overline{\{x(y):y\in Y\}}$ for $(x(y))_{y\in Y}\in\prod_{y\in Y}Y$. Clearly $\Phi:\prod_{y\in Y}(Y,S)\to(2^Y,\widetilde{S})$ is a factor map, which implies that $(2^Y,\widetilde{S})$ is semi-distal, since for each recurrent point x in the factor there is a recurrent point y in the extension which maps to x by the factor map. $\square$

We now show that

Theorem 2.14. *A weakly scattering t.d.s. is disjoint from all minimal semi-distal t.d.s..*

Proof. Assume the contrary, i.e. there are an invertible weakly scattering system (X,T) and a minimal semi-distal system (Y,S) such that $X\not\perp Y$. Thus, there exists a proper subsystem $(X_1,T\times S)$ of $(X\times Y,T\times S)$ with $\pi_X(X_1)=X$ and $\pi_Y(X_1)=Y$, where $\pi_X:X_1\to X$ and $\pi_Y:X_1\to Y$ are the projections to X and Y respectively.

Let $\pi_X^{-1}:X\to 2^{X_1}$ with $x\mapsto\pi_X^{-1}(x)$. Then π_X^{-1} is continuous on a dense G_δ set D_0 of X. Put $D=\bigcap_{i\in\mathbb{Z}}T^{-i}D_0\cap Tran_T(X)$. Then D is an invariant dense G_δ set of X.

Set $H=\overline{\{\pi_X^{-1}(x):x\in D\}}\subseteq 2^{X_1}$. It is easy to see that $\widetilde{T\times S}(H)=H$ and for any $A\in H$, $\pi_X(A)$ is a singleton denoted by $\pi_X(A)=\{x_A\}$. Then $A=\{x_A\}\times\pi_Y(A)$. Let $\hat{\pi}_X:H\to X$ with $\hat{\pi}_X(A)=x_A$. By the choice of D, $\hat{\pi}_X$ is an almost one-to-one map, so $(H,\widetilde{T\times S})$ is a transitive subsystem of $(2^{X_1},\widetilde{T\times S})$. Let $\hat{\pi}_Y:H\to 2^Y$ with $\hat{\pi}_Y(A)=\pi_Y(A)$ and let $H_Y=\hat{\pi}_Y(H)$. Then $\hat{\pi}_Y:(H,\widetilde{T\times S})\to(H_Y,\widetilde{S})$ is a factor map. Since Y is minimal semi-distal, $(2^Y,\widetilde{S})$ is semi-distal (see Lemma 2.13). Thus, $(H_Y,\widetilde{S})$ is minimal semi-distal since $(H_Y,\widetilde{S})$ is transitive. Moreover, it follows by the minimality of (Y,S) that $(H_Y,\widetilde{S})$ is a non-trivial minimal semi-distal system.

It is well known that there exists a non-trivial eigenvalue $\lambda\neq 1$ for the minimal semi-distal system $(H_Y,\widetilde{S})$. Particularly, $\lambda\in Eig(H_Y,\widetilde{S})$. By Lemma 2.4, $\lambda\in Eig(H,\widetilde{T\times S})=Eig(X,T)$ which contradicts the assumption $Eig(X,T)=\{1\}$. $\square$

Combining the above theorem and some known results (see [**8**]) we have

- a weakly scattering t.d.s. is disjoint from all minimal semi-distal t.d.s..
- a scattering t.d.s. is disjoint from all minimal HPI t.d.s..
- a weakly mixing system with dense minimal points is disjoint from all minimal PI t.d.s..
- a weakly mixing system with dense distal points is disjoint from all minimal t.d.s..

The following questions are open

(1) Is it true that a weakly scattering t.d.s. is disjoint from all minimal HPI t.d.s.?
(2) Is it true that a weakly scattering t.d.s. is weakly disjoint from all minimal PI t.d.s.?

We remark that if the second question has an affirmative answer, so does the first question (by [**7**, Lemma 4.3]).

3. Generic homomorphisms, generic factors and weak disjointness

In the study of measure-theoretical dynamical systems, we neglect sets with zero measure. The topologically negligible sets are sets of first category or meager sets. Observing this fact Weiss [**22**] introduced the notion of generic isomorphisms for two t.d.s.. He proved that for countable group actions, any two non-trivial full shifts are generically isomorphic. In this section, we will define generic homomorphisms and generic factors for a transitive system partially following the idea of Weiss, and then study the relations among generic homomorphisms, generic factors and weak disjointness.

First, we recall some notations related to a family (for details see [**10, 2, 5**]). Denote by $\mathcal{P}$ the collection of all subsets of $\mathbb{Z}_+$. A subset $\mathcal{F}$ of $\mathcal{P}$ is a *family*, if it is *hereditary upwards*. That is, $F_1 \subset F_2$ and $F_1 \in \mathcal{F}$ imply $F_2 \in \mathcal{F}$. A family $\mathcal{F}$ is called *proper* if it is a proper subset of $\mathcal{P}$, i.e. neither empty nor all of $\mathcal{P}$. It is easy to see that $\mathcal{F}$ is proper iff $\mathbb{Z}_+ \in \mathcal{F}$ and $\emptyset \notin \mathcal{F}$. Any subset $\mathcal{A}$ of $\mathcal{P}$ can generate a family $[\mathcal{A}] = \{F \in \mathcal{P} : F \supset A$ for some $A \in \mathcal{A}\}$. Let $\mathcal{F}$ be a proper family, a t.d.s. (X, T) is $\mathcal{F}$*-transitive* if for each pair of nonempty open subsets U and V of X, $N(U, V) \in \mathcal{F}$.

For a t.d.s. (X, T), we define

$$\mathcal{F}_X = [\{N(U_1, U_2) : U_1, U_2 \text{ are opene subsets of } X\}],$$

we call $\mathcal{F}_X$ *the generating family* of (X, T). It is not hard to see that (X, T) is a transitive system iff $\mathcal{F}_X$ is a proper family. Using the generating family of a system, it is easily seen that

Lemma 3.1. *Let (X, T) and (Y, S) be two transitive t.d.s.. Then (X, T) is weakly disjoint from (Y, S) iff $\mathcal{F}_X \subseteq k\mathcal{F}_Y$ (equivalently $\mathcal{F}_Y \subseteq k\mathcal{F}_X$).*

Now we define generic factors.

Definition 3.2. Let (X, T) and (Y, S) be two transitive t.d.s.. If there exists a continuous map $\phi : Tran_T(X) \to Tran_S(Y)$ with $\phi(Tx) = S\phi(x)$ for all $x \in Tran_T(X)$, then we say ϕ is a *generic homomorphism* from (X, T) to (Y, S), (Y, S) is a *generic factor* of (X, T) and (X, T) is a *generic extension* of (Y, S). It is not hard to see that if (X, T) is minimal and $\phi : (X, T) \to (Y, S)$ is a generic homomorphism then ϕ is a factor map.

We have

Lemma 3.3. *Let $\phi : (X, T) \to (Y, S)$ be a generic homomorphism between transitive t.d.s.. Then*

$$\mathcal{F}_Y \subseteq \mathcal{F}_X \text{ and } Eig(S) \subseteq Eig(T).$$

In particular, for a proper family $\mathcal{F}$, if (X, T) is $\mathcal{F}$-transitive, then so is (Y, S).

Proof. For $x \in Tran_T(X)$ we have $\phi(x) \in Tran_S(Y)$. Since $S^n\phi(x) = \phi(T^n x) \in \phi(Tran_T(X))$ for $n \in \mathbb{Z}_+$, $\phi(Tran_T(X))$ is dense in $Tran_S(Y)$. Hence for any opene subset V of Y, $\phi^{-1}(V \cap Tran_S(Y))$ is an opene subset of $Tran_T(X)$.

Since $\phi^{-1}(V_1 \cap Tran_S(Y)), \phi^{-1}(V_2 \cap Tran_S(Y))$ are opene subsets of $Tran_T(X)$ for any opene subsets V_1, V_2 of Y, there exist open subsets U_1, U_2 of X such that

$U_i \cap Tran_T(X) = \phi^{-1}(V_i \cap Trans_S(Y)), i = 1, 2$. Fixing $x_0 \in \phi^{-1}(V_1 \cap Trans_S(Y))$, we get

$$\begin{aligned} N(U_1, U_2) &= N(x_0, U_1) - N(x_0, U_2) \\ &= N(x_0, \phi^{-1}(V_1 \cap Trans_S(Y))) - N(x_0, \phi^{-1}(V_2 \cap Trans_S(Y))) \\ &\subseteq N(\phi^{-1}(V_1 \cap Trans_S(Y)), \phi^{-1}(V_2 \cap Trans_S(Y))) \subseteq N(V_1, V_2), \end{aligned}$$

where $A - B = \{a - b \geq 0 : a \in A, b \in B\}$ for $A, B \subset \mathbb{Z}_+$. Hence $N(V_1, V_2) \in \mathcal{F}_X$. Since V_1, V_2 are arbitrary, $\mathcal{F}_Y \subseteq \mathcal{F}_X$.

Now let $\lambda \in Eig(S)$. Then there exists a non-zero continuous map $f : Trans_S(Y) \to \mathbb{T}$ with $f(Ty) = \lambda f(y)$ for any $y \in Trans_S(Y)$. Then $g = f \circ \phi$ is a generic eigenfunction of (X, T) with the generic eigenvalue λ, and hence $\lambda \in Eig(T)$. $\square$

LEMMA 3.4. *Let $\phi : (X, T) \to (Y, S)$ be a generic homomorphism between transitive t.d.s.. If (X, T) is weakly disjoint from a t.d.s. (Z, R), then (Y, S) is also weakly disjoint from (Z, R).*

PROOF. It follows from Lemmas 3.1 and 3.3. $\square$

LEMMA 3.5. *Let (X, T) be a t.d.s. and $\pi_1 : (\widetilde{X}, \sigma_T) \longrightarrow (X, T)$ be the natural extension of (X, T). If (Z, R) is a t.d.s. and $\mathcal{F}$ is a proper family, then*

1. *$\mathcal{F}_X = \mathcal{F}_{\widetilde{X}}$. Hence (X, T) is $\mathcal{F}$-transitive iff $(\widetilde{X}, \sigma_T)$ is $\mathcal{F}$-transitive.*
2. *(X, T) is weakly disjoint from (Z, R) iff $(\widetilde{X}, \sigma_T)$ is weakly disjoint from (Z, R).*

PROOF. For any opene subsets U, V of $\widetilde{X}$, we can find opene subsets U', V' of X and $k \in \mathbb{N}$ such that $\pi_k^{-1} U' \subset U, \pi_k^{-1} V' \subset V$. Clearly, $N(U, V) \supseteq N(\pi_k^{-1} U', \pi_k^{-1} V') = N(U', V')$, so that $N(U, V) \in \mathcal{F}_X$ which implies $\mathcal{F}_{\widetilde{X}} \subseteq \mathcal{F}_X$. Conversely for any opene subsets U', V' of X, $N(U', V') = N(\pi_1^{-1} U', \pi_1^{-1} V') \in \mathcal{F}_{\widetilde{X}}$, that is $\mathcal{F}_X \subseteq \mathcal{F}_{\widetilde{X}}$. This shows (1). (2) follows by (1) and Lemma 3.1. $\square$

LEMMA 3.6. *Let $\pi : (X, T) \to (Y, S)$ be an almost one-to-one extension, (Z, R) be a t.d.s. and $\mathcal{F}$ be a proper family. Then*

1. *$\mathcal{F}_X = \mathcal{F}_Y$. Hence (X, T) is $\mathcal{F}$-transitive iff (Y, S) is $\mathcal{F}$-transitive.*
2. *(X, T) is weakly disjoint from (Z, R) iff (Y, S) is weakly disjoint from (Z, R).*

PROOF. Since $\pi : (X, T) \to (Y, S)$ is almost one-to-one, there exists a dense G_δ set $X_0 \subset X$ with $\pi^{-1}\pi(x) = \{x\}$ for any $x \in X_0$. For any opene subset U of X and $x \in X_0 \cap U$, one has $\pi^{-1}\pi(x) = \{x\} \subset U$. Therefore, there exists an open neighborhood U' of $\pi(x)$ with $\pi^{-1}(U') \subset U$. Hence for any opene subsets U, V of X, we can find opene subsets U', V' of Y such that $\pi^{-1} U' \subset U, \pi^{-1} V' \subset V$. Clearly, $N(U, V) \supseteq N(\pi^{-1} U', \pi^{-1} V') = N(U', V')$, so that $N(U, V) \in \mathcal{F}_Y$ and $\mathcal{F}_X \subseteq \mathcal{F}_Y$. Conversely for any opene subsets U', V' of Y, $N(U', V') = N(\pi^{-1} U', \pi^{-1} V') \in \mathcal{F}_X$, that is $\mathcal{F}_Y \subseteq \mathcal{F}_X$. This shows (1). (2) follows by (1) and Lemma 3.1. $\square$

PROPOSITION 3.7. *Let (X, T) and (Y, S) be two transitive t.d.s.. If (X, T) is weakly disjoint from (Y, S), then any common generic factor of (X, T) and (Y, S) is weakly mixing. Moreover, if (Y, S) is minimal and equicontinuous, then (X, T) is weakly disjoint from (Y, S) iff (X, T) and (Y, S) have no common nontrivial generic factor.*

Proof. Let (Z,R) be a common generic factor of (X,T) and (Y,S), then $\mathcal{F}_Z \subseteq \mathcal{F}_X \cap \mathcal{F}_Y$. Hence

$$\mathcal{F}_Z \subseteq \mathcal{F}_X \subseteq k\mathcal{F}_Y \subseteq k\mathcal{F}_Z.$$

Therefore (Z,R) is weakly disjoint from (Z,R), that is, (Z,R) is weakly mixing.

Now let (Y,S) be minimal and equicontinuous, and (X,T) be weakly disjoint from (Y,S). Assume (Z,R) is a common generic factor of (X,T) and (Y,S), then (Z,R) is weakly mixing. Since (Y,S) is minimal, (Z,R) is a factor of (Y,S). Hence (Z,R) is an equicontinuous weakly mixing system, it must be a trivial system.

Conversely, let (X,T) and (Y,S) have no common nontrivial generic factors. Assume the contrary, i.e. (X,T) is not weakly disjoint from (Y,S). Then we have

Claim: there exists $1 \neq \lambda \in Eig(T) \cap Eig(S)$.

Proof. Since (Y,S) is a minimal rotation of a compact metric abelian group, we assume $Y = G$ is the group, $S = R_{g_0}$ is the rotation by $g_0 \in G$. Let $x_0 \in Tran_T(X)$ and $e \in G$ be the unit of G. Put $P = \mathrm{cl}(\bigcup_{n\in\mathbb{Z}_+} T \times R_{g_0})^n(x_0,e))$. As $(X \times G, T \times S)$ is not transitive, $P \neq X \times G$. Thus $H = \{g \in G : (x_0, g) \in P\}$ is a closed proper subset of G.

We now show that H is a closed group of G. First, we show that H is a semigroup. Let $g_1, g_2 \in H$. Then there exist $n_k, m_k \in \mathbb{Z}_+$ such that $T^{n_k}x_0 \to x_0$, $T^{m_k}x_0 \to x_0$ and ${g_0}^{n_k} \to g_1$, ${g_0}^{m_k} \to g_2$. It is easy to see that there exist $s_i, t_i \in \mathbb{N}$ such that $T^{n_{s_i}+m_{t_i}}x_0 \to x_0$ and $g^{n_{s_i}+m_{t_i}} \to g_1g_2$. This shows $(x_0, g_1g_2) \in P$. Therefore, $g_1g_2 \in H$. Hence H is a closed semigroup. Note that for every $g \in G$ and $m \in \mathbb{Z}$, $g^m \in \mathrm{cl}(\{g^n : n \in \mathbb{Z}_+\})$. Thus if $g \in H$ then $g^{-1} \in H$ since H is a closed semigroup. This implies H is a closed group.

Since H is a closed proper group of G, G/H is a non-trivial topological group. Let γ be a non-trivial character of G/H, which can be considered as a non-trivial character of G with $\gamma(H) = \{1\}$. Using γ we can define an eigenvalue for (G, R_{g_0}) with the eigenvalue $\lambda = \gamma(g_0) \neq 1$. Now whenever $T^{n_k}x_0 \longrightarrow x_0$, $n_k \to +\infty$, we claim $\lambda^{n_k} \longrightarrow 1$. Assume the contrary that there exists $\{k_i\}$ such that $\lambda^{n_{k_i}} \longrightarrow \lambda_0 \neq 1$ and ${g_0}^{n_{k_i}} \longrightarrow g_1$. Note that $(x_0, g_1) \in P$, one has $g_1 \in H$. Moreover $\lambda_0 = \lim_{i\longrightarrow+\infty} \lambda^{n_{k_i}} = \lim_{i\longrightarrow+\infty} \gamma({g_0}^{n_{k_i}}) = \gamma(g_1) = 1$, a contradiction. Now by Lemma 2.3, $\lambda \in Eig(X,T)$. Hence $1 \neq \lambda \in Eig(T) \cap Eig(S)$. This finishes the proof of the Claim. □

Let $f : Tran_T(X) \to \mathbb{T}$ be a generic eigenvalue of (X,T) with the generic eigenvalue λ. Set $K = \mathrm{cl}(\{\lambda^n : n \in \mathbb{Z}_+\}) \subset \mathbb{T}$ and $R_\lambda : K \to K$ with $R_\lambda(k) = \lambda k, \forall k \in K$. Then (K, R_λ) is a periodic system or an irrational rotation on $(\mathbb{T}, R_\lambda)$. Since $\lambda \neq 1$, (K, R_λ) is a nontrivial minimal equicontinuous system.

Fix $x_0 \in Tran_T(X)$ and let $\phi : Tran_T(X) \to K$ be defined by $\phi(x) = \frac{f(x)}{f(x_0)}, \forall\ x \in Tran_T(X)$. Clearly $\phi : (X,T) \to (K, R_\lambda)$ is a generic factor map, i.e. (K, R_λ) is a generic factor of (X,T). Similarly, (K, R_λ) is a generic factor of (Y,S). This shows that (K, R_λ) is a non-trivial common generic factor of (X,T) and (Y,S), a contradiction. □

Now we can interrupt weakly scattering via generic factors.

Theorem 3.8. *Let (X,T) be a transitive t.d.s.. Then (X,T) is weakly scattering iff it has no non-trivial equicontinuous generic factor.*

PROOF. Since a generic factor of a minimal equicontinuous system is minimal and equicontinuous, Theorem 3.8 follows from Proposition 3.7. □

4. Relations among several scattering properties

It is well known that a minimal t.d.s. is weakly mixing iff it has no non-trivial eigenvalues iff it is weakly scattering. For a general transitive system (X,T), the property having no non-trivial generic eigenvalues is strictly weaker than weakly mixing (cf. for example [**9, 5, 14**]). Then it is natural to ask for which systems weakly scattering is equivalent to weakly mixing. In this section, we will give some sufficient conditions such that a weakly scattering system is weakly mixing. First we discuss the question of which systems weakly scattering is equivalent to strongly scattering. As applications, we show that for Banach transitive systems, weakly scattering is equivalent to strongly scattering, and for E-systems, weakly scattering is equivalent to weakly mixing.

Let (X,T) be a t.d.s.. A point $x_0 \in X$ is called an *equicontinuity point* if for every $\epsilon > 0$ there exists $\delta > 0$ such that if $d(y,x) < \delta$ then $d(T^i(y), T^i(x)) < \epsilon$ for each $i \in \mathbb{Z}_+$. A t.d.s. (X,T) is *almost equicontinuous* if it is transitive and there is an equicontinuity point (cf. [**3**]). For an almost equicontinuous system, the set of almost equicontinuous points coincides with the set of transitive points.

DEFINITION 4.1. Let $\mathcal{F}$ be a proper family.

(1) We call $\mathcal{F}$ a *non-almost equicontinuous (for short NAE) family*, if the trivial system is the unique $\mathcal{F}$-transitive almost equicontinuous system.
(2) We call $\mathcal{F}$ a minimal almost equicontinuous (for short MAE) family, if every $\mathcal{F}$-transitive almost equicontinuous system is minimal.

REMARK 4.2. Let $\mathcal{F}$ and $\mathcal{F}'$ be two proper families and $\mathcal{F}' \subseteq \mathcal{F}$. It is not hard to see that if $\mathcal{F}$ is a NAE-family (resp. a MAE-family), then so is $\mathcal{F}'$.

In the following, we will give some examples of NAE-families and MAE-families. First we give a class of NAE-families. We start with the following (the lemma first appeared in [**14**])

LEMMA 4.3. *Suppose that (X,T) is almost equicontinuous (non-trivial) with an equicontinuous point x. Assume for each i, U_i is an open neighborhood of x with $\cap_i U_i = \{x\}$. Then there is an i such that for any $p,q \in \mathbb{N}$, $N(U_i,U_i) \supset \{p-1, q-1, p+q-1\}$ does not hold.*

PROOF. Assume that for each i there are p_i, q_i such that $N(U_i,U_i) \supset \{p_i - 1, q_i - 1, p_i + q_i - 1\}$.

Take an open neighborhood U of x such that $diam(U) < \epsilon$, $diam(T(U)) < \epsilon$ and $4\epsilon < d(U, T(U))$ for some $\epsilon > 0$. As x is an equicontinuous point we can assume that $diam(T^n(U)) < \epsilon$ for each $n \in \mathbb{N}$. Let $V = T(U)$. Then $T^n(U) \cap U \neq \emptyset$ and $T^n(U) \cap V \neq \emptyset$ can not hold at the same time.

It is easy to see that there is an i such that $N(x,U) \supset N(U_i,U_i)$. Thus

$$N(U,U) = N(x,U) - N(x,U) \supset N(U_i,U_i) - N(U_i,U_i) \supset \{p_i, p_i - 1\}.$$

This implies that $T^{p_i}(U) \cap U \neq \emptyset$ and $T^{p_i}(U) \cap V \supset T(T^{p_i-1}(U) \cap U) \neq \emptyset$, a contradiction. □

Take $\mathcal{N} = \{A \subset \mathbb{Z}_+ : A \supset \{p-1, q-1, p+q-1\}\ \exists p, q \in \mathbb{N}\}$. Recall that $A \subset \mathbb{N}$ is *thick* if it contains arbitrarily long intervals. By Lemma 4.3, one easily obtains

PROPOSITION 4.4. *If a family $\mathcal{F} \subseteq \mathcal{N}$, then $\mathcal{F}$ is a NAE-family. In particular, the family consisting of thick sets is a NAE-family.*

Let S be a subset of $\mathbb{Z}_+$. The *upper Banach density of S* is

$$BD^*(S) = \limsup_{|I| \longrightarrow +\infty} \frac{|S \cap I|}{|I|},$$

where I ranges over intervals of $\mathbb{Z}_+$. Let $\mathcal{D} = \{S \subset \mathbb{Z}_+ : BD^*(S) > 0\}$.

A t.d.s. is *Banach-transitive* if the upper Banach density of $N(U, V)$ is positive for each pair of opene subsets U, V of X. The following lemma appears in [**14**] (cf. also [**12**]). For completeness, we copy the proof.

THEOREM 4.5. *If (X, T) is almost equicontinuous and Banach-transitive, then it is minimal.*

PROOF. As (X, T) is almost equicontinuous, there is an equicontinuous point p which is a transitive point. For $\epsilon > 0$ let $U = B_\epsilon(p)$. Then, there is $\delta > 0$ such that $N(p, U) \supset N(U_\delta, U_\delta)$, where $U_\delta = B_\delta(p)$. For example, we take $0 < \delta < \epsilon/2$ such that $diam(T^n(U_\delta)) < \epsilon/2$ for $n \in \mathbb{Z}_+$. At the same time, there is $\delta_1 > 0$ such that $N(p, U_\delta) \supset N(U_{\delta_1}, U_{\delta_1})$, where $U_{\delta_1} = B_{\delta_1}(p)$. Thus, we have

$$N(p, U) \supset N(U_\delta, U_\delta) \supset N(U_{\delta_1}, U_{\delta_1}) - N(U_{\delta_1}, U_{\delta_1}).$$

By [**10**, Proposition 3.19] $N(p, U)$ is syndetic, and thus p is a minimal point. This implies that (X, T) is minimal. $\square$

PROPOSITION 4.6. *If a family $\mathcal{F} \subseteq \mathcal{D}$, then $\mathcal{F}$ is a MAE-family.*

The following lemma is crucial in this section and our main results in the section are corollaries of this lemma.

LEMMA 4.7. *Let (X, T) be a transitive t.d.s., (Y, S) be a t.d.s. and ν be an S-invariant ergodic measure on (Y, S) with full support. If $(X \times Y, T \times S)$ is not transitive, then (X, T) has a non-trivial almost equicontinuous generic factor (Ω, σ).*

PROOF. We assume that S is invertible, otherwise consider its natural extension; and use Lemma 3.4 and the fact that the natural extension of (Y, S) has an invariant ergodic measure with full support. Since $(X \times Y, T \times S)$ is not transitive, there exists a proper subsystem $(J, T \times S)$ of $(X \times Y, T \times S)$ with $\mathrm{int}(J) \neq \emptyset$. Let $J_x = \{y \in Y : (x, y) \in J\}$. Then $x \mapsto J_x$ is an upper semi-continuous map from X to 2^Y, i.e. for each neighborhood U of x, $J_y \subset U$ if y is close enough to x.

Step 1. Let $\psi : Tran_T(X) \to L^1(\nu)$ be such that $\psi(x) = 1_{J_x} \in L^1(\nu)$ for $x \in Tran_T(X)$. We claim: ψ is a continuous map.

PROOF. First we show $\nu(J.)$ is a constant function on $Tran_T(X)$. Let $x_0 \in Tran_T(X)$ and for any $\epsilon > 0$, take an open neighborhood U of J_{x_0} in Y with $\nu(U \setminus J_{x_0}) < \epsilon$. For any $x_1 \in Tran_T(X)$, there exists $n \in \mathbb{N}$ such that $S^n J_{x_1} \subseteq J_{T^n x_1} \subset U$ as $x \mapsto J_x$ is upper semi-continuous. Since $\nu(J_{x_1}) = \nu(S^{n_i} J_{x_1}) \le \nu(U) \le \nu(J_{x_0}) + \epsilon$, this implies that $\nu(J_{x_1}) \le \nu(J_{x_0}) + \epsilon$. Since x_0, x_1 and ϵ are arbitrary, one gets $\nu(J_x)$ is a constant for any $x \in Tran_T(X)$.

Fix $x_0 \in Tran_T(X)$. For any $\epsilon > 0$, take an open neighborhood U of J_{x_0} with $\nu(U \setminus J_{x_0}) < \epsilon$. As $x \mapsto J_x$ is upper semi-continuous, when $x \in Tran_T(X)$ is close enough to x_0, one has $J_x \subset U$ and $\nu(J_x \setminus J_{x_0}) \le \nu(U \setminus J_{x_0}) < \epsilon$. Since $\nu(J_x \cap J_{x_0}) = \nu(J_x) - \nu(J_x \setminus J_{x_0}) \ge \nu(J_x) - \epsilon$, one has $\nu(J_{x_0} \setminus J_x) = \nu(J_{x_0}) - \nu(J_x \cap$

$J_{x_0}) \leq \nu(J_{x_0}) - (\nu(J_x) - \epsilon) = \epsilon$. Hence when $x \in Tran_T(X)$ is close enough to x_0, one has $||\psi(x_0) - \psi(x)||_{L^1(\nu)} = \nu(J_{x_0} \Delta J_x) = \nu(J_{x_0} \setminus J_x) + \nu(J_x \setminus J_{x_0}) < 2\epsilon$. Therefore, ψ is continuous at x_0 and thus continuous on $Tran_T(X)$. This ends the proof of the Claim. □

Step 2. Let $Z_0 = \psi(Tran_T(X)) \subseteq L^1(\nu)$. Define $S_1 : L^1(\nu) \to L^1(\nu)$ with $S_1 f(y) = f(S^{-1}y), \forall f \in L^1(\nu)$. Clearly, S_1 is an isometry of $L^1(\nu)$. For any $1_{J_x} \in Z_0, x \in Tran_T(X)$, since $SJ_x \subseteq J_{Tx}$ and $\nu(SJ_x) = \nu(J_x) = \nu(J_{Tx})$, one has $1_{SJ_x} = 1_{J_{Tx}}$ in $L^1(\nu)$, i.e. $S_1\phi(x) = \phi(Tx), \forall x \in Tran_T(X)$. Hence

$$\begin{array}{ccc} Tran_T(X) & \xrightarrow{T} & Tran_T(X) \\ \psi\Big\downarrow & & \psi\Big\downarrow \\ Z_0 & \xrightarrow{S_1} & Z_0 \end{array}$$

is a commutative diagram and $S_1(Z_0) \subseteq Z_0$. Now for any $1_{J_x} \in Z_0, x \in Tran_T(X)$, take $x_1 \in X$ with $Tx_1 = x$. Then $x_1 \in Tran_T(X)$ and $S_1 1_{J_{x_1}} = 1_{J_x}$ in $L^1(\nu)$. Moreover $S_1^{-1} 1_{J_x} = 1_{J_{x_1}} \in Z_0$. This shows that $S_1^{-1} Z_0 \subseteq Z_0$. Combining this fact and $S_1(Z_0) \subseteq Z_0$, one has $S_1(Z_0) = Z_0$. In the following, we will construct a compactification (Ω, σ) of (Z_0, S_1) such that it is an almost equicontinuous system which we need (for the method see the proof of [**5**, Theorem 4.7]).

On $I^{\mathbb{Z}}$, the space of bi-infinite sequences in $I = [0, 1]$, the metric d given by

$$d(a, b) = \sup\{\frac{|a_i - b_i|}{2^{|i|}} : i \in \mathbb{Z}\} \tag{4.1}$$

yields the compact product topology. Let σ be the shift homeomorphism so that $\sigma(a)_i = a_{i+1}$ for $i \in \mathbb{Z}$. Set $d_\sigma(a, b) \doteq \sup_{i \in \mathbb{Z}}\{d(\sigma^i a, \sigma^i b)\}, \forall a, b \in I^{\mathbb{Z}}$. From (4.1) we see that

$$d_\sigma(a, b) = \sup\{|a_i - b_i| : i \in \mathbb{Z}\}. \tag{4.2}$$

Since $int(J) \neq \emptyset$, there exist $x_0 \in Tran_T(X)$ and an opene subset U of Y such that $J_{x_0} \supset U$. This shows $\nu(J_{x_0}) > 0$ as ν has full support. Since $J \neq X \times Y$, $J \supset \{x_0\} \times J_{x_0}$ and $x_0 \in Tran_T(X)$, one knows $J_{x_0} \neq Y$. Hence

$$0 < \nu(J_{x_0}) < 1. \tag{4.3}$$

Now fix $g_0 = 1_{J_{x_0}} \in Z_0$ and define a map $h : Z_0 \to I^{\mathbb{Z}}$ by

$$h(g)_i = ||S_1^i g - g_0||_{L^1(\nu)} = ||g - S_1^{-i} g_0||_{L^1(\nu)} \text{ for } i \in \mathbb{Z} \text{ and } g \in Z_0. \tag{4.4}$$

By the triangle inequality it is clear that $|h(g_1)_i - h(g_2)_i| \leq ||g_1 - g_2||_{L^1(\nu)}$ for $g_1, g_2 \in Z_0$. On the other hand, we can choose a sequence $\{i_n\} \subset \mathbb{N}$ so that $S_1^{i_n} g_1 \to g_0$ in $L^1(\nu)$ (equivalently $S_1^{-i_n} g_0 \to g_1$ in $L^1(\nu)$) from which it follows that $|h(g_1)_{i_n} - h(g_2)_{i_n}| \to ||g_1 - g_2||_{L^1(\nu)}$. So from (4.2) we obtain

$$d_\sigma(h(g_1), h(g_2)) = ||g_1 - g_2||_{L^1(\nu)}. \tag{4.5}$$

Hence $h : Z_0 \to I^{\mathbb{Z}}$ is a uniformly continuous injective map.

In fact, $h^{-1} : h(Z_0) \to Z_0$ is also continuous. To see this let $h(g_n) \to h(g), g_n, g \in Z_0$. Then for any $i \in \mathbb{Z}$

$$|||g_n - S_1^{-i} g_0||_{L^1(\nu)} - ||g - S_1^{-i} g_0||_{L^1(\nu)}| = |h(g_n)_i - h(g)_i| \to 0, \text{ when } n \to +\infty.$$

Hence

$$\begin{aligned}\limsup_{n\to+\infty}||g_n-g||_{L^1(\nu)} &\le \limsup_{n\to+\infty}(||g_n-S_1^{-i}g_0||_{L^1(\nu)}+||g-S_1^{-i}g_0||_{L^1(\nu)})\\ &= 2||g-S_1^{-i}g_0||_{L^1(\nu)}, \forall i\in\mathbb{Z}.\end{aligned}$$

Since $g \in Z_0$, $\liminf_{i\to+\infty}||g - S_1^{-i}g_0||_{L^1(\nu)} = \liminf_{i\to+\infty}||S_1^i g - g_0||_{L^1(\nu)} = 0$. This shows that $\lim_{n\to+\infty}||g_n - g||_{L^1(\nu)} = 0$, and thus $h^{-1} : h(Z_0) \to Z_0$ is continuous. It follows that h is a uniformly continuous embedding of Z_0 into $I^{\mathbb{Z}}$.

Let Ω be the closure of $h(Z_0)$ in $I^{\mathbb{Z}}$. Since $hS_1 = \sigma h$, Ω is a σ-invariant subset. Obviously, (Ω,σ) is a transitive system and $h(Z_0) \subseteq Tran_\sigma(\Omega)$. Let $\phi = h\circ\psi : Tran_T(X) \to \Omega$, then ϕ is a continuous map from $Tran_T(X)$ to $Tran_\sigma(\Omega)$, i.e., $\phi(Tran_T(X)) \subset Tran_\sigma(\Omega))$ with $\phi(Tx) = \sigma\phi(x), \forall x \in Tran_T(X)$. Hence (Ω,σ) is a generic factor of (X,T).

Step 3: To end the proof we now show that (Ω,σ) is a non-trivial almost equicontinuous system. First, (Ω,σ) is non-trivial. Assume the contrary, i.e. Ω is a singleton. Since $h(g_0)_0 = ||g_0 - g_0||_{L^1(\nu)} = 0$, $\Omega = \{(\cdots 000\cdots)\}$. Particularly, $h(g_0) = (\cdots 000\cdots)$. This shows $||S_1 g_0 - g_0||_{L^1(\nu)} = h(g_0)_1 = 0$, i.e., $\nu(J_{x_0}\Delta SJ_{x_0}) = 0$. Since ν is ergodic, $\nu(J_{x_0}) = 0$ or 1 which contradicts (4.3).

Second, (Ω,σ) is an almost equicontinuous system. Let $\omega = h(g_0)$. So $\omega \in Tran_\sigma(\Omega)$. We will prove that for any $\epsilon > 0$ there exists $\delta > 0$ such that for any $i,j \in \mathbb{Z}_+$ $d(\sigma^i\omega,\omega) < \delta$ implies $d(\sigma^{i+j}\omega,\sigma^j\omega) \le \epsilon$.

For any $\epsilon > 0$, since h is uniformly continuous, there exists $\epsilon_0 > 0$ such that $||g_1-g_2||_{L^1(\nu)} < \epsilon_0$ implies $d(h(g_1),h(g_2)) < \epsilon$ for $g_1,g_2 \in Z_0$. As $h^{-1} : h(Z_0) \to Z_0$ is continuous at $\omega = h(g_0)$, there exists $\delta > 0$ such that $d(h(g),h(g_0)) < \delta$ implies $||g - g_0||_{L^1(\nu)} < \epsilon_0$ for $g \in Z_0$. Since $\sigma^i\omega = h(S_1^i g_0)$, one gets that $d(\omega,\sigma^i\omega) < \delta$ implies $||g_0 - S_1^i g_0||_{L^(\nu)} < \epsilon_0$. As S_1 is an isometry, $||S_1^j g_0 - S_1^{i+j}g_0||_{L^1(\nu)} < \epsilon_0$ for $j \in \mathbb{Z}_+$. From the choice of ϵ_0 one has $d(\sigma^j\omega,\sigma^{i+j}\omega) < \epsilon$ for $j \in \mathbb{Z}_+$. We conclude that (Ω,σ) is an almost equicontinuous system. □

It is easy to see that (X,T) is strongly scattering iff (X,T) is weakly disjoint from any transitive t.d.s. having an ergodic invariant measure with full support. As an application of Lemma 4.7, one has

Theorem 4.8. *Let (X,T) be a transitive t.d.s.. If (X,T) has no non-trivial almost equicontinuous generic factors, then (X,T) is strongly scattering.*

Remark 4.9. (1) If the converse of Theorem 4.8 holds then there is no non-trivial almost equicontinuous strongly scattering systems, and consequently scattering and strongly scattering are different properties as Akin and Glasner [**5**] has shown the existence of an almost equicontinuous scattering system. If the converse of Theorem 4.8 does not hold, then there exists a non-trivial almost equicontinuous strongly scattering system.

(2) As there is no non-trivial extremely scattering almost equicontinuous t.d.s. [**14**], extremely scattering implies the property of having no non-trivial almost equicontinuous generic factors by Lemma 3.4, which in turn implies strongly scattering by Theorem 4.8. Recall that for an infinite sequence S, a t.d.s. (X,T) is S-equicontinuous, if for any $\epsilon > 0$ there exists $\delta > 0$ such that whenever $x,y \in X$ and $d(x,y) < \delta$ one has $d(T^{s_i}(x),T^{s_i}(y)) < \epsilon$ for each $s_i \in S$. By [**16**] and Lemma 3.4 for any positive upper Banach density set S, a strongly scattering t.d.s. has no non-trivial S-equicontinuous

generic factors. We do not know whether an almost equicontinuous t.d.s. is S-equicontinuous for some S with positive upper Banach density.

(3) A t.d.s. (X,T) is *uniformly rigid* if there is an increasing sequence $\{n_i\}$ of $\mathbb{N}$ such that $T^{n_i} \longrightarrow id$ uniformly. An open question is: if a transitive t.d.s. (X,T) has no non-trivial uniformly rigid generic factors, is (X,T) mildly mixing? Note that in [**16, 13**] the authors showed that a mildly mixing uniformly rigid system is a trivial one. Thus by Lemma 3.4 a mildly mixing system has no non-trivial uniformly rigid generic factors.

DEFINITION 4.10. Let (X,T) be a t.d.s..

(1) We call (X,T) a NAE-transitive system, if (X,T) is transitive and the generating family $\mathcal{F}_X$ of (X,T) is a NAE family;
(2) We call (X,T) a MAE-transitive system, if (X,T) is transitive and the generating family $\mathcal{F}_X$ of (X,T) is a MAE family.

REMARK 4.11. It is well known that a t.d.s. is weakly mixing iff the set of return times $N(U,V)$ is thick (see [**10**, Proposition II.3]). By Propositions 4.4 and 4.6, a weakly mixing system is a NAE-transitive one and a Banach-transitive system is a MAE-transitive one.

PROPOSITION 4.12. *The following statements hold.*

(1) *NAE-transitivity is stable under the natural extensions, almost one to one extensions and generic factors.*
(2) *MAE-transitivity is stable under the natural extensions, almost one to one extensions and generic factors.*

PROOF. (1). Let (X,T) be a NAE-transitive system. Then $\mathcal{F}_X$ is a NAE-family. If $\pi_1 : (\widetilde{X},\sigma_T) \to (X,T)$ is the natural extension of (X,T), $\pi : (Y,S) \to (X,T)$ is an almost one-to-one extension and $\phi : (X,T) \to (Z,R)$ is a generic homomorphism, then $\mathcal{F}_{\widetilde{X}} = \mathcal{F}_X$, $\mathcal{F}_Y = \mathcal{F}_X$ and $\mathcal{F}_Z \subseteq \mathcal{F}_X$ by Lemmas 3.3, 3.5 and 3.6. By the Remark after Definition 4.1, $\mathcal{F}_{\widetilde{X}}$, $\mathcal{F}_Y$ and $\mathcal{F}_Z$ are all NAE-families. Thus $(\widetilde{X},\widetilde{T})$, (Y,S) and (Z,R) are NAE-transitive systems. (2) can be proved similarly. $\square$

Now we show the main results of the section.

THEOREM 4.13. (1) *A NAE-transitive system is strongly scattering.*
(2) *For a MAE-transitive system, weakly scattering is equivalent to strongly scattering.*

PROOF. (1). Since an almost equicontinuous NAE-transitive system is trivial, one has that a NAE-transitive system has no non-trivial almost equicontinuous generic factor by Proposition 4.12 (1). Now by Theorem 4.8, a NAE-transitive system is strongly scattering.

(2). Let (X,T) be a MAE-transitive system. It remains to show that if (X,T) is weakly scattering, then it is strongly scattering. Let (X,T) be weakly scattering. If (X,T) is not strongly scattering, then there exist a t.d.s. (Y,S) and an S-invariant ergodic measure ν on (Y,S) with full support such that $(X\times Y, T\times S)$ is not transitive. By Lemma 4.7 there exists a non-trivial almost equicontinuous system (Ω,σ) such that (Ω,σ) is a generic factor of (X,T). Let $\phi : (X,T) \to (\Omega,\sigma)$ be the corresponding generic homomorphism.

Since (Ω,σ) is a generic factor of (X,T), (Ω,σ) is an MAE-transitive almost equicontinuous system by Proposition 4.12. Hence (Ω,σ) is a non-trivial minimal and equicontinuous system. So there exists $\lambda \in Eig(\Omega,\sigma)\setminus\{1\}$. Thus $\lambda \in Eig(X,T)$ which contradicts the fact that (X,T) has no non-trivial generic eigenvalues. Thus (X,T) is strongly scattering. $\square$

PROPOSITION 4.14. *Let (X,T) be a t.d.s..*

1. *If (X,T) is Banach-transitive, then (X,T) is weakly scattering iff it is strongly scattering.*
2. *If (X,T) is an E-system, then (X,T) is weakly scattering iff it is weakly mixing.*

PROOF. Since a Banach-transitive system is MAE-transitive, (1) follows from Theorem 4.13(2).

(2) As an E-system is Banach transitive (see [**12**]), (2) follows from (1). $\square$

REMARK 4.15. We have the following remarks.

1. For a t.d.s. having a fully supported ergodic measure, Proposition 4.14(2) was proved in [**19**] using the Hilbert-Schmidt operator.
2. Banach-transitivity together with strongly scattering does not imply weakly mixing. To see this note that there is a t.d.s. which is extremely scattering and not weakly mixing [**15**], and an extremely scattering system is Banach-transitive and strongly scattering ([**14**]).

5. Scattering properties in an almost equicontinuous system

Whether weakly scattering, scattering and strongly scattering are the same property is an interesting question. In section 4, we have showed that for Banach transitive systems, the three properties are pairwise equivalent. In the section, we mainly consider the problem for an almost equicontinuous system. First, we give equivalent forms of weakly scattering and scattering respectively. Then we show that the existence of a complete pointed monothetic group which is minimally almost periodic yet does not have the fixed point property, as suggested by Glasner, is equivalent to there exists an almost equicontinuous system which is weakly scattering yet not scattering. We follow some ideas in [**5, 11**].

For topological spaces X and Y, denote by $C(X;Y)$ the set of all continuous maps from X to Y. When X is a compact metric space with a metric d, $C(X;X)$ is metrizable via the sup metric:

$$d(g_1,g_2)=\sup_{x\in X} d(g_1(x),g_2(x)).$$

Under this metric, $C(X;X)$ is a metric topological semigroup as composition is jointly continuous. Let $Homeo(X)$ be the subset of $C(X;X)$ consisting of all homeomorphisms from X to itself. Under the metric d, $Homeo(X)$ forms a complete topological group (cf. [**5**]). For any closed abelian subgroup Λ of $Homeo(X)$, d is an invariant metric on Λ.

For $x\in X$, the evaluation map

$$ev_x : C(X;X)\to X \text{ with } ev_x(f)=f(x)$$

is a uniformly continuous map. If (X, T) is a t.d.s. with a metric d on X, we define a new metric

$$d_T(x, y) = \sup_{i \in \mathbb{Z}_+} d(T^i x, T^i y).$$

The following lemma is a result of Akin, Auslander and Berg (cf. [**3, 5**]).

LEMMA 5.1. *Let (X, T) be an almost equicontinuous system and $x \in Tran_T(X)$. Let Λ_T denote the closure of $\{T^i : i \in \mathbb{Z}_+\}$ in $C(X; X)$. Then Λ_T is an abelian subgroup of $C(X; X)$ in which the cyclic subgroup $\{T^i : i \in \mathbb{Z}\}$ is dense. The continuous evaluation map $ev_x : C(X; X) \to X$ restricts to define an isometry of Λ_T onto $Tran_T(X)$ equipped with the metric d_T, i.e.*

$$d(g_1, g_2) = d_T(g_1(x), g_2(x)) \textit{ for } g_1, g_2 \in \Lambda_T.$$

For an almost equicontinuous system (X, T), Λ_T is a topological group with an invariant metric and has a dense cyclic subgroup generated by T. In general, by *a pointed monothetic group* we mean a pair (G, g) where G is a topological group with a compatible invariant metric and $g \in G$ generates a cyclic subgroup dense in G. A topological group G is called *minimally almost periodic* if the only continuous character on G is the trivial one, i.e. the constant function 1.

THEOREM 5.2. *Let (X, T) be an almost equicontinuous system and Λ_T denote the closure of $\{T^i : i \in \mathbb{Z}_+\}$ in $C(X; X)$. Then (X, T) is weakly scattering iff the pointed monothetic group (Λ_T, T) is minimally almost periodic.*

PROOF. Let $x_0 \in Tran_T(X)$ and γ be a continuous character of Λ_T. Then $f = \gamma \circ ev_{x_0}^{-1} : Tran_T(X) \to \mathbb{T}$ is a non-zero continuous generic eigenfunction associated with $\gamma(T)$. If (X, T) is weakly scattering, then $\gamma(T) = 1$. Hence $\gamma(T^n) = 1$ for any $n \in \mathbb{Z}$, which implies that γ is a trivial character. Therefore, (Λ_T, T) is minimally almost periodic.

Conversely, assume (Λ_T, T) is minimally almost periodic. Let $\lambda \in Eig(T)$. Then there exists a non-zero continuous map $f : Tran_T(X) \to \mathbb{C}$ with $f(Tx) = \lambda f(x)$ for $\forall x \in Tran_T(X)$. Put $g(x) = \frac{f(x)}{f(x_0)}, \forall x \in Tran_T(X)$. Then $|g(x)| \equiv 1$ and $g(x_0) = 1$.

Let $\gamma : \Lambda_T \to \mathbb{T}$ with $\gamma = g \circ ev_{x_0}$, then γ is continuous and $\gamma(T \circ S) = \lambda \cdot \gamma(S)$ for any $S \in \Lambda_T$. Now we show that γ is a group homomorphism.

For $S_1, S_2 \in \Lambda_T$, there exist $n_k, m_k \in \mathbb{Z}_+$ with $T^{n_k} \to S_1, T^{m_k} \to S_2$ in $C(X; X)$. Since γ is continuous,

$$\lim_k \lambda^{n_k} = \lim_k \gamma(T^{n_k}) = \gamma(S_1), \lim_k \lambda^{m_k} = \lim_k \gamma(T^{m_k}) = \gamma(S_2).$$

It is easy to see that there exist $s_i, t_i \in \mathbb{N}$ such that $\lim_{i \longrightarrow +\infty} T^{n_{s_i} + m_{t_i}} = S_1 \circ S_2$ and $\lim_{i \longrightarrow +\infty} \lambda^{n_{s_i} + m_{t_i}} = \gamma(S_1) \cdot \gamma(S_2)$. Moreover,

$$\gamma(S_1) \cdot \gamma(S_2) = \lim_{i \longrightarrow +\infty} \lambda^{n_{s_i} + m_{t_i}} = \lim_{i \longrightarrow +\infty} \gamma(T^{n_{s_i} + m_{t_i}}) = \gamma(S_1 \circ S_2).$$

This shows that γ is a group homomorphism. Hence γ is a continuous character on Λ_T. Since (Λ_T, T) is minimally almost periodic, one has $\gamma \equiv 1$, Particularly $\lambda = \gamma(T) = 1$. Therefore, $Eig(T) = \{1\}$. □

Let G be a topological group. When G acts on (jointly continuous) a compact metric space X we say (X, G) is a *G-flow*. For example, for an almost equicontinuous system (X, T), if Λ_T is the closure of $\{T^i : i \in \mathbb{Z}_+\}$ in $C(X; X)$, then

$\Lambda_T \times X \to X$ with $(g,x) \to g(x)$ is a Λ_T action on X, i.e. (X, Λ_T) is a Λ_T-flow. A G-flow (X,G) is a *minimal flow*, if for each $x \in X$, Gx is dense in X. By Zorn's Lemma, each G-flow has a minimal G-subflow. We say that G has *the fixed point property*, if for each G-flow (X,G) there exists a fixed point $x \in X$, i.e. $g(x) = x, \forall g \in G$. It is easy to see that a topological group G has the fixed point property iff the trivial G-flow is the only minimal G-flow.

In [**5**], Akin and Glasner proved the following result, for completeness we provide a proof.

PROPOSITION 5.3. *Let (X,T) be an almost equicontinuous system and Λ_T denote the closure of $\{T^i : i \in \mathbb{Z}_+\}$ in $C(X;X)$. Then (X,T) is scattering iff the pointed monothetic group (Λ_T, T) has the fixed point property.*

PROOF. ($\Rightarrow$) Assume (X,T) is scattering. Let (Y, Λ_T) be a Λ_T-flow. Without loss of generality we assume (Y, Λ_T) is minimal. Since T is the generator of the monothetic group Λ_T, (Y,T) is also minimal. Thus $(X \times Y, T \times T)$ is transitive.

Take $(x_0, y) \in Tran_{T\times T}(X \times Y)$. For any open neighborhood U of y, since $\pi_y : \Lambda_T \to Y$ with $\pi_y(g) = g(y)$ is a continuous map, $\pi_y^{-1}(U)$ is an open subset of Λ_T and $e \in \pi_y^{-1}(U)$ (e is the unit of Λ_T). Since ev_{x_0} is an isometry between (Λ_T, d) and $(Tran_T(X), d_T)$, there exists an open neighborhood V of x_0 such that $ev_{x_0}(\pi_y^{-1}(U)) = Tran_T(X) \cap V$. Since $N(x_0, V) = N(x_0, ev_{x_0}(\pi_y^{-1}(U))) \subset N(y, U)$, $N(x_0, V)y \subset U$.

As (x_0, y) is a transitive point, $\overline{N(x_0, V)y} = Y$. This shows $Y \subset \overline{U}$. Since U is arbitrary, Y is a singleton. This implies that the pointed monothetic group (Λ_T, T) has the fixed point property.

($\Leftarrow$) Assume (Λ_T, T) has the fixed point property. If (X,T) is not scattering, then there exists an invertible minimal system (Y,S) such that $(X \times Y, T \times S)$ is not transitive. Hence there exists a proper subsystem $(X_1, T \times S)$ of $(X \times Y, T \times S)$ with $\mathrm{int}(X_1) \neq \emptyset$. Let $\pi_X : X_1 \to X$ and $\pi_Y : X_1 \to Y$ be the projection maps. Let $\pi_Y^{-1} : Y \to 2^{X_1}$ with $y \mapsto \pi_Y^{-1}(y)$. Then π_Y^{-1} is continuous on a dense G_δ set D_0 of Y by Lemma 2.4. Put $D = \bigcap_{i\in\mathbb{Z}} T^{-i} D_0$, D is also a dense G_δ set of X.

Set $H = \overline{\{\pi_Y^{-1}(y) : y \in D\}} \subset 2^{X_1}$, it is easy to see that $\widetilde{T \times S}(H) = H$ and for any $A \in H$, $\pi_Y(A)$ is a single point set. Write $\pi_Y(A) = \{y_A\}$, then $A = \pi_X(A) \times \{y_A\}$. Let $\hat{\pi}_Y : H \to Y$ with $\hat{\pi}_Y(A) = y_A$ for $A \in H$. By the choose of D, $\hat{\pi}_Y$ is an almost one-to-one map, so $(H, \widetilde{T \times S})$ is a minimal subsystem of $(2^{X_1}, \widetilde{T \times S})$.

Let $\hat{\pi}_X : H \to 2^X$ with $\hat{\pi}_X(A) = \pi_X(A)$ for $A \in H$. Let $H_X = \hat{\pi}_X(H)$, then $\hat{\pi}_X : (H, \widetilde{T \times S}) \to (H_X, \widetilde{T})$ is a factor map. Therefore $(H_X, \widetilde{T})$ is a minimal system.

For a Λ_T-flow (X, Λ_T), it induces naturally a Λ_T-action on 2^X with $(g, A) \to g(A)$ for any $g \in \Lambda_T$ and $A \in 2^X$. Since T is a generator of Λ_T, $\Lambda_T(H_X) = H_X$. Hence (H_X, Λ_T) is a minimal Λ_T-system. Since Λ_T has the fixed point property, H_X is a singleton.

Since $\mathrm{int}(X_1) \neq \emptyset$, there exists $y \in D$ with $\mathrm{int}(\pi_X(\pi_Y^{-1}(y))) \neq \emptyset$. As H_X is trivial and $\hat{\pi}_X(\pi_Y^{-1}(y)) = \pi_X(\pi_Y^{-1}(y)) \in H_X$, one has $T(\pi_X(\pi_Y^{-1}(y))) = \pi_X(\pi_Y^{-1}(y))$. Hence $\pi_X(\pi_Y^{-1}(y))$ is a T-invariant closed subset of X and has non-empty interior. From the transitivity of (X,T), one knows $\pi_X(\pi_Y^{-1}(y)) = X$. This shows $X_1 \supseteq \{y\} \times X$, and hence $X_1 = X \times Y$ which contradicts the fact that $X_1 \neq X \times Y$. $\square$

A topological group G is called *monothetic*, if there exists $g \in G$ such that the cyclic subgroup generating by g is dense in G. It is well known that a monothetic topological group is abelian.

THEOREM 5.4. *The following three statements are equivalent:*

(1) *There exists a monothetic topological group which is minimally almost periodic yet does not have the fixed point property.*
(2) *There exists a complete pointed monothetic group (G, g) which is minimally almost periodic yet does not have the fixed point property.*
(3) *There exists an almost equicontinuous system which is weakly scattering yet does not scattering.*

PROOF. (1) $\Rightarrow$ (2) Let G be a monothetic topological group which is minimally almost periodic yet not have the fixed pointed property. Then there exists a non-trivial minimal G-flow (X, G), that is, there exists a continuous map $\pi : G \times X \to X$ such that $\pi(e, x) = x$ and $\pi(g_1 g_2, x) = \pi(g_1, \pi(g_2, x)) \forall\ x \in X$ where e is the unit of G and $g_1, g_2 \in G$.

Let $\Phi : G \to Homeo(X)$ with $\Phi(g) = \pi(g, \cdot)$, then Φ is a continuous map. Denote Λ_G the closure of $\Phi(G)$ in $Homeo(X)$, then Λ_G is a closed abelian subgroup of $Homeo(X)$. Hence Λ_G is a complete topological group with a compatible invariant metric d. Now as G is monothetic, there exists $g \in G$ such that the cyclic subgroup generated by g is dense in G. Then $(\Lambda_G, \Phi(g))$ is a complete pointed monothetic group.

Since the map $\Lambda_G \times X \to X$ with $(h, x) \mapsto h(x),\ \forall\ h \in \Lambda_G,\ x \in X$ is jointly continuous and (X, G) is minimal, (X, Λ_G) is a non-trivial minimal Λ_G-flow. This shows that the pointed monothetic group $(\Lambda_G, \Phi(g))$ does not the fixed point property. Let $\chi : \Lambda_G \to \mathbb{T}$ be a continuous character on Λ_G, then $\gamma \doteq \chi \circ \Phi$ is a continuous character on G. Since G is minimally almost periodic, γ is a trivial character. In particular, $\gamma(g) = 1$. Hence $\chi(\Phi(g)) = 1$, and thus χ is a trivial character as the cyclic subgroup generated by $\Phi(g)$ is dense in Λ_G. This shows that $(\Lambda_G, \Phi(g))$ is minimally almost periodic.

(2) $\Rightarrow$ (3) Let (G, g) be a complete pointed monothetic group which is minimally almost periodic yet does not have the fixed point property. Let ρ be a compatible invariant metric of G. Note that $\rho'(x, y) = \min\{\rho(x, y), 1\}, \forall x, y \in G$ is also a compatible invariant metric of G. Without loss of generality we assume that $\rho(x, y) \leq 1$ for any $x, y \in G$.

Now we follow the proof of Step 2 in Lemma 4.7. Define a map $h : G \to I^{\mathbb{Z}}$ by

$$h(g')_i = \rho(g^i g', e) = \rho(g', g^{-i}) \text{ for } i \in \mathbb{Z} \text{ and } g' \in G.$$

Let Ω be the closure of $h(G)$ in $I^{\mathbb{Z}}$, then (Ω, σ) is an almost equicontinuous system and $h : G \to \Omega$ is a uniformly continuous embedding of G into $I^{\mathbb{Z}}$ such that

(a). $h(G) \subseteq Tran_\sigma(\Omega)$
(b). $\sigma h(g_1) = h(g g_1)$ for any $g_1 \in G$.
(c). $\rho(g_1, g_2) = d_\sigma(h(g_1), h(g_2))$ for any $g_1, g_2 \in G$.

Let Λ_σ denote the closure in $C(\Omega; \Omega)$ of $\{\sigma^i : i \in \mathbb{Z}_+\}$. By Lemma 5.1, the continuous evaluation map $ev_{h(e)} : C(\Omega; \Omega) \to \Omega$ restricts to define an isometry of Λ_σ onto $Tran_\sigma(\Omega)$ equipped with the metric d_σ.

Since (G, ρ) and $(Tran_\sigma(\Omega), d_\sigma)$ are complete metric spaces, $h(G) = Tran_\sigma(\Omega)$ by (c) above. Moreover, $ev_{h(e)}^{-1} \circ h : (G, g) \to (\Lambda_\sigma, \sigma)$ is an isometry. Write $\phi =$

$ev_{h(e)}^{-1} \circ h$, then $\phi(g^n) = \sigma^n$, $\forall n \in \mathbb{Z}_+$. In fact, ϕ is also a homomorphism between topological groups. To see this let $g_1, g_2 \in G$. Then there exist $m_i, n_i \in \mathbb{Z}_+$ such that $g^{m_i} \to g_1$, $g^{n_i} \to g_2$. Obviously, $g^{m_i+n_i} \to g_1 g_2$. Since ϕ is an isometric map and $\phi(g^n) = \sigma^n$, $n \in \mathbb{Z}_+$, one has $\phi(g_1) = \lim_{i\to\infty} \sigma^{m_i}$, $\phi(g_2) = \lim_{i\to\infty} \sigma^{n_i}$ and $\phi(g_1 g_2) = \lim_{i\to\infty} \sigma^{m_i+n_i}$. Hence $\phi(g_1 g_2) = \phi(g_1)\phi(g_2)$.

Since (G, g) is minimally almost periodic yet does not have the fixed point property and $\phi : (G, g) \to (\Lambda_\sigma, \sigma)$ is an isometric homomorphism of topological groups, so the pointed monothetic group (Λ_σ, σ) has the same property. This shows that (Ω, σ) is weakly scattering yet does not scattering by Theorem 5.2 and Proposition 5.3.

(3) $\Rightarrow$ (1) If there exists an almost equicontinuous system (X, T) which is weakly scattering yet not scattering, then the monothetic group (Λ_T, T) is minimally almost periodic yet not have the fixed pointed property by Theorem 5.2 and Proposition 5.3. □

Remark 5.5. Glasner [**11**] showed that there exists a non-trivial complete pointed monothetic group having the fixed point property. Hence there exists a non-trivial scattering almost equicontinuous system by the above proof(see also [**5**]). Since a non-trivial almost equicontinuous system is not weakly mixing, this shows that weakly mixing is strictly stronger than scattering (this is a result in [**5**]). In fact, we showed that weakly mixing is strictly stronger than strongly scattering, even extremely scattering (see [**14, 15**]).

By Theorem 5.4, Theorem 2.11 and the remark after Theorem 2.11, we have

Corollary 5.6. *If there exists a monothetic topological group which is minimally almost periodic and does not have the fixed point property, then there exists a Bohr-recurrent nonrecurrent sequence.*

References

[1] M. Ajtai, I. Havas and J. Komlós, *Every group admits a bad topology*, Studies in pure mathematics, 21–34, Birkhäuser, Basel, 1983. MR820207 (86k:22002)

[2] E. Akin, *Recurrence in topological dynamical systems: Furstenberg families and Ellis actions*, Plenum Press, New York, 1997. MR1467479 (2000c:37014)

[3] E. Akin, J. Auslander and K. Berg, *When is a transitive map chaotic*? Convergence in Ergodic Theory and Probability (Columbus, OH, 1993) (Ohio University Math. Res. Inst. Pub., 5). de Gruyter, Berlin, 1996, pp. 25-40. MR1412595 (97i:58106)

[4] E. Akin, J. Auslander and E. Glasner, The topological dynamics of Ellis actions, Memoirs of Amer. Math. Soc., **913**.

[5] E. Akin and E. Glasner, *Residual properties and almost equicontinuity*, J. d'Anal. Math., **84**(2001), 243-286. MR1849204 (2002f:37020)

[6] J. Auslander, *Minimal flows and their extensions*, North-Holland Mathematics Studies, 153. MR956049 (89m:54050)

[7] F. Blanchard, B. Host, A. Maass, *Topological Complexity*, Ergod. Th. and Dynam. Sys., **20** (2000), 641-662. MR1764920 (2002b:37019)

[8] P. Dong, S. Shao and X. Ye, *Product recurrent properties, disjointness and weak disjointness*, to appear in Israel J. of Math., 2012. DOI: 10.1007/s11856-011-0128-z.

[9] T. Downarowicz, *Weakly almost periodic flows and Hidden Eigenvalues*, Contemporary Mathematics, Volume **215**(1998), 101-120. MR1603157 (99c:28044)

[10] H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Math. Systems Theory, **1** (1967), 1-49. MR0213508 (35:4369)

[11] E. Glasner, *On minimal actions of Polsih groups*, Topology Appl., **117**(1998), 259-272. MR1617456 (99c:54057)

[12] E. Glasner and B. Weiss, *Locally equicontinuous dynamical systems*, Colloq. Math., **84-85**(2000), 345-361. MR1784216 (2001j:37021)
[13] E. Glasner and B. Weiss, *On the interplay between measurable and topological dynamics*, Handbook of Dynamical systems, Vol. 1B, Hasselblatt and Katok, eds., Elsevier, Amsterdam, 2006, 597-648. MR2186250 (2006i:37005)
[14] W. Huang and X. Ye, *An explicit scattering, non-weakly mixing example and weak disjointness*, Nonlinearity, **15**(2002), 849-862. MR1901110 (2003b:37016)
[15] W. Huang and X. Ye, *Dynamical systems disjoint from any minimal system*, Trans. Amer. Math. Soc., **357**(2005), 669-694. MR2095626 (2005g:37012)
[16] W. Huang and X. Ye, *Topological complexity, return times and weak disjointness*, Ergod. Th. and Dynam. Sys., **24**(2004), 825-846. MR2062921 (2005c:37011)
[17] Y. Katznelson, *Chromatic numbers of Cayley graphs on $\mathbb{Z}$ and recurrence*, Paul Erdós and his mathematics (Budapest, 1999), Combinatorics, **21**(2001), 211–219. MR1832446 (2002h:05065)
[18] H. B. Keynes and J. B. Robertson, *On ergodicity and mixing in topological transformation groups*, Duke Math. J. **35**(1968), 809–819. MR0234441 (38:2758)
[19] H. B. Keynes and J. B. Robertson, *Eigenvalue theorems in topological transformation groups*, Trans. Amer. Math. Soc., **139**(1969), 359-369. MR0237748 (38:6029)
[20] R. Peleg, *Weak disjointness of transformation groups*, Proc. Amer. Math. Soc., **33**(1972), 165–170. MR0298642 (45:7694)
[21] W. A. Veech, *The equicontinuous structure relation for minimal Abelian transformation groups*, Amer. J. Math., **90**(1968), 723–732. MR0232377 (38:702)
[22] B. Weiss, *A survey of generic dynamics. Descriptive set theory and dynamical systems* (Marseille-Luminy, 1996), 273–291, London Math. Soc. Lecture Note Ser., 277, Cambridge Univ. Press, Cambridge, 2000. MR1774430 (2001j:37008)
[23] B. Weiss, *Single Orbit Dynamics*, AMS Regional Conference Series in Mathematics, N. **95**, 2000. MR1727510 (2000k:37001)

Department of Mathematics, University of Science and Technology of China, Hefei, Anhui, 230026, People's Republic of China
E-mail address: `wenh@mail.ustc.edu.cn`

Department of Mathematics, University of Science and Technology of China, Hefei, Anhui, 230026, People's Republic of China
E-mail address: `yexd@ustc.edu.cn`

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11233

Around King's rank-one theorems: Flows and $\mathbb{Z}^n$-actions

Élise Janvresse, Thierry de la Rue, and Valery Ryzhikov

Dedicated to A. M. Stepin on the occasion of his 70-th birthday.

Abstract. We study the generalizations of Jonathan King's rank-one theorems (Weak-Closure Theorem and rigidity of factors) to the case of rank-one $\mathbb{R}$-actions (flows) and rank-one $\mathbb{Z}^n$-actions. We prove that these results remain valid in the case of rank-one flows. In the case of rank-one $\mathbb{Z}^n$ actions, where counterexamples have already been given, we prove partial Weak-Closure Theorem and partial rigidity of factors.

1. Introduction

Very important examples in ergodic theory have been constructed in the class of rank-one transformations, which is closely connected to the notion of transformations with fast cyclic approximation [**KS66**]: If the rate of approximation is sufficiently fast, then the transformation will be inside the rank-one class. The notion of rank-one transformations has been defined in [**Orn72**], where mixing examples have appeared. Later, Daniel Rudolph used them for a machinery of counterexamples [**Rud79**].

Jonathan King contributed to the theory of rank-one transformations by several deep and interesting facts. His Weak-Closure-Theorem (WCT) [**Kin86**] is now a classical result with applications even out of the range of $\mathbb{Z}$-actions (see for example [**Tik06**]). He also proved the minimal-self-joining (MSJ) property for rank-one mixing automorphisms (see [**Kin88**]), the rigidity of non-trivial factors [**Kin86**], and the weak closure property for all joinings for flat-roof rank-one transformations [**Kin01**].

A natural question is whether the corresponding assertions remain true for flows ($\mathbb{R}$-actions) and for $\mathbb{Z}^n$-actions. We show that for flows the situation is quite similar: The joining proof of the Weak-Closure Theorem given in [**Ryz92**] (see also [**Ryz10**]) can be adapted to the situation of a rank-one $\mathbb{R}$-action (Theorem 5.2). We also give in the same spirit a proof of the rigidity of non-trivial factors of rank-one flows (Theorem 6.2) which, with some simplification, provides a new proof of King's result in the case of $\mathbb{Z}$-actions. We prove a flat-roof flow version as well (Theorem 7.1). Note that a proof of the Weak-Closure Theorem for rank-one flows

---

2010 *Mathematics Subject Classification.* Primary 37A10, 37A15, 37A35.

Key words and phrases. Rank-one actions, weak closure theorem, factors, joinings.

This work is partially supported by the grant NSh 8508.2010.1. The first draft of the paper was written while the third author was visiting the University of Rouen.

had already been published in [**Zei93**]. Unfortunately it relies on the erroneous assumption that if $(T_t)_{t\in\mathbb{R}}$ is a rank-one flow, then there exists a real number t_0 such that T_{t_0} is a rank-one transformation (see beginning of Section 3.2 in [**Zei93**]).

Concerning multidimensional rank-one actions, the situation is quite different. The Weak-Closure Theorem is no more true [**DK02**], and factors may be non-rigid [**DS09**]. Rank-one partially mixing $\mathbb{Z}$-actions have MSJ [**KT91**], however it is proved in [**DS09**] that for $\mathbb{Z}^2$-actions this is generally not true. We remark that it was an answer for $\mathbb{Z}^2$-action to Jean-Paul Thouvenot's question: Whether a mildly mixing rank-one action possesses MSJ, though this interesting problem remains open for $\mathbb{Z}$-actions. Regardless these surprising results, there are some partial versions of WCT: Commuting automorphisms can be partially approximated by elements of the action (Corollary 8.4), and non-trivial factors must be partially rigid (Corollary 8.5). We present these results as consequences of A. Pavlova's theorem (Theorem 8.3, see also [**Ryz08**]) .

2. Preliminaries and notations

Weak convergence of probability measures. We are interested in groups of automorphisms of a Lebesgue space $(X,\mathcal{A},\mu)$, where μ is a continuous probability measure. The properties of these group actions are independent of the choice of the underlying space X, and for practical reasons we will assume that $X=\{0,1\}^{\mathbb{Z}}$, equipped with the product topology and the Borel σ-algebra. This σ-algebra is generated by the cylinder sets, that is sets obtained by fixing a finite number of coordinates. On the set $\mathcal{M}_1(X)$ of Borel probability measures on X, we will consider the topology of weak convergence, which is characterized by

$$\nu_n \xrightarrow[n\to\infty]{w} \nu \quad\iff\quad \text{for all cylinder set } C,\ \nu_n(C)\xrightarrow[n\to\infty]{}\nu(C),$$

and turns $\mathcal{M}_1(X)$ into a compact metrizable space.

We will often consider probability measures on $X\times X$, with the same topology of weak convergence. We will use the following observation: If ν_n and ν in $\mathcal{M}_1(X\times X)$ have their marginals absolutely continuous with respect to our reference measure μ, with bounded density, then the weak convergence of ν_n to ν ensures that for all *measurable* sets A and B in $\mathcal{A}$, $\nu_n(A\times B)\xrightarrow[n\to\infty]{}\nu(A\times B)$.

Self-joinings. Let $T=(T_g)_{g\in G}$ be an action of the Abelian group G by automorphism of the Lebesgue space $(X,\mathcal{A},\mu)$. A *self-joining* of T is any probability measure on $X\times X$ with both marginals equal to μ and invariant by $T\times T=(T_g\times T_g)_{g\in G}$. For any automorphism S commuting with T, we will denote by Δ_S the self-joining concentrated on the graph of S^{-1}, defined by

$$\forall A,B\in\mathcal{A},\ \Delta_S(A\times B):=\mu(A\cap SB).$$

In particular, for any $g\in G$ we will denote by Δ^g the self-joining Δ_{T_g}. In the special case where $S=T_0=\mathrm{Id}$, we will note simply Δ instead of Δ^0 or Δ_{Id}.

If $\mathcal{F}$ is a factor (a sub-σ-algebra invariant under the action (T_g)), we denote by $\mu\otimes_{\mathcal{F}}\mu$ the *relatively independent joining above* $\mathcal{F}$, defined by

$$\mu\otimes_{\mathcal{F}}\mu(A\times B):=\int_X \mathbb{E}_\mu[\mathbb{1}_A|\mathcal{F}]\ \mathbb{E}_\mu[\mathbb{1}_B|\mathcal{F}]\,d\mu.$$

Recall that $\mu\otimes_{\mathcal{F}}\mu$ coincides with Δ on the σ-algebra $\mathcal{F}\otimes\mathcal{F}$.

Flows. A *flow* is a continuous family $(T_t)_{t\in\mathbb{R}}$ of automorphisms of the Lebesgue space $(X,\mathcal{A},\mu)$, with $T_t\circ T_s = T_{t+s}$ for all $t,s\in\mathbb{R}$, and such that $(t,x)\mapsto T_t(x)$ is measurable. We recall that the measurability condition implies that for all measurable set A, $\mu(A \vartriangle T_tA) \xrightarrow[t\to 0]{} 0$.

LEMMA 2.1. *Let $(T_t)_{t\in\mathbb{R}}$ be an ergodic flow on $(X,\mathcal{A},\mu)$. Let Q be a dense subgroup of $\mathbb{R}$, and λ be an invariant probability measure for the action of $(T_t)_{t\in Q}$. Assume further that $\lambda \ll \mu$, with $\frac{d\lambda}{d\mu}$ bounded by some constant C. Then $\lambda=\mu$.*

PROOF. Let $t\in\mathbb{R}$, and let (t_n) be a sequence in Q converging to t. For any measurable set A, we have

$$\lambda\Big(T_tA \vartriangle T_{t_n}A\Big) \le C\mu\Big(T_tA \vartriangle T_{t_n}A\Big) \xrightarrow[n\to\infty]{} 0.$$

Hence $\lambda(T_tA) = \lim_n \lambda(T_{t_n}A) = \lambda(A)$. This proves that λ is T_t-invariant for each $t\in\mathbb{R}$. Since μ is ergodic under the action of $(T_t)_{t\in\mathbb{R}}$, we get $\lambda=\mu$. □

3. Rank-one flows

DEFINITION 3.1. A flow $(T_t)_{t\in\mathbb{R}}$ is of rank one if there exists a sequence (ξ_j) of partitions of the form

$$\xi_j = \left\{E_j,\ T_{s_j}E_j,\ T_{s_j}^2E_j,\ \dots, T_{s_j}^{h_j-1}E_j, X\setminus \bigsqcup_{i=0}^{h_j-1} T_{s_j}^iE_j\right\}$$

such that ξ_j converges to the partition onto points (that is, for every measurable set A and every j, we can find a ξ_j-measurable set A_j in such a way that $\mu(A \vartriangle A_j) \xrightarrow[j\to\infty]{} 0$), s_j/s_{j+1} are integers, $s_j\to 0$ and $s_jh_j\to\infty$.

Several authors have generalized the notion of a rank-one transformation to an $\mathbb{R}$-action using continuous Rokhlin towers (see *e.g.* [**Pri01**]). One can show that the above definition includes all earlier definitions of rank-one flows with continuous Rokhlin towers. The above definition without the requirement that s_j/s_{j+1} be integers was given by the third author in [**Ryz92**].

LEMMA 3.2. *Let $(T_t)_{t\in\mathbb{R}}$ be a rank-one flow. Then the sequences (s_j) and (h_j) in the definition can be chosen so that*

$$s_j^2h_j \xrightarrow[j\to\infty]{} \infty.$$

PROOF. Let (s_j) and (h_j) be given as in the definition. Recall that $h_js_j\to\infty$. For each j, let $n_j>j$ be a large enough integer such that $s_js_{n_j}h_{n_j}>j$. Define $\ell_j := s_j/s_{n_j}\in\mathbb{Z}_+$. We consider the new partition

$$\tilde{\xi}_j := \left\{\tilde{E}_j, T_{s_j}\tilde{E}_j,\cdots, T_{s_j}^{\tilde{h}_j-1}\tilde{E}_j, X\setminus \bigsqcup_{i=0}^{\tilde{h}_j-1} T_{s_j}^i\tilde{E}_j\right\}$$

where

$$\tilde{E}_j := \bigsqcup_{i=0}^{\ell_j-1} T_{s_{n_j}}^i E_{n_j}$$

and $\tilde{h}_j := [h_{n_j}/\ell_j]$. One can easily check that $\tilde{\xi}_j$ still converges to the partition onto points. Moreover we have $s_j^2\tilde{h}_j = s_j^2[h_{n_j}s_{n_j}/s_j]\to\infty$. □

Lemma 3.3 (Choice Lemma for flows, abstract setting). *Let $(T_t)_{t\in\mathbb{R}}$ be an arbitrary flow, and let ν be an ergodic invariant measure under the action of $(T_t)_{t\in\mathbb{R}}$. Let a family of measures (ν_j^k) satisfy the conditions:*

- *There exist sequences (d_j) and (s_j) of positive numbers with $d_j \xrightarrow[j\to\infty]{} 0$, s_j/s_{j+1} is an integer for all j, and $s_j \xrightarrow[j\to\infty]{} 0$, such that for all measurable set A and all k,j*

$$\left|\nu_j^k(T_{s_j}A) - \nu_j^k(A)\right| < s_j\, d_j; \tag{3.1}$$

- *There exists a family of positive numbers (a_j^k) with $\sum_k a_j^k = 1$ for all j, such that*

$$\sum_k a_j^k \nu_j^k \xrightarrow[j\to\infty]{w} \nu. \tag{3.2}$$

Then there is a sequence (k_j) such that $\nu_j^{k_j} \xrightarrow[j\to\infty]{w} \nu$.

Proof. Given a cylinder set B, an integer $j \geq 1$ and $\varepsilon > 0$, we consider the sets K_j of all integers k such that

$$\nu(B) - \nu_j^k(B) > \varepsilon.$$

Suppose that the (sub)sequence K_j satisfies the condition

$$\sum_{k\in K_j} a_j^k \geq a > 0.$$

Let λ be a limit point for the sequence of measures $(\sum_{k\in K_j} a_j^k)^{-1} \sum_{k\in K_j} a_j^k \nu_j^k$. Then $\lambda \neq \nu$ since $\lambda(B) \leq \nu(B) - \varepsilon$, but by (3.2), we have $\lambda \ll \nu$, and $d\lambda/d\nu \leq 1/a$. Moreover, the measure λ is invariant by T_{s_p} for all p. Indeed, for $j \geq p$, since s_p/s_j is an integer, we get from (3.1) that

$$\left|\nu_j^k(T_{s_p}A) - \nu_j^k(A)\right| < s_p\, d_j \xrightarrow[j\to\infty]{} 0.$$

By Lemma 2.1, it follows that $\lambda = \nu$. The contradiction shows that

$$\sum_{k\in K_j} a_j^k \to 0.$$

Thus, for all large enough j, most of the measures ν_j^k satisfy

$$|\nu_j^k(B) - \nu(B)| < \varepsilon.$$

Let $\{B_1, B_2, \dots\}$ be the countable family of all cylinder sets. Using the diagonal method we find a sequence k_j such that for each n

$$|\nu_j^{k_j}(B_n) - \nu(B_n)| \xrightarrow[j\to\infty]{} 0,$$

i.e. $\nu_j^{k_j} \xrightarrow[j\to\infty]{w} \nu$. □

Columns and fat diagonals in $X \times X$. Assume that $(T_t)_{t\in\mathbb{R}}$ is a rank-one flow defined on X, with a sequence (ξ_j) of partitions as in Definition 3.1. For all j and $|k| < h_j - 1$, we define the sets $C_j^k \in X \times X$, called *columns*:

$$C_j^k := \bigsqcup_{\substack{0\le r,\ell\le h_j-1\\ r-\ell=k}} T_{s_j}^r E_j \times T_{s_j}^\ell E_j.$$

Given $0 < \delta < 1$, we consider the set

$$D_j^\delta := \bigsqcup_{k=-[\delta h_j]}^{[\delta h_j]} C_j^k.$$

(See Figure 1.)

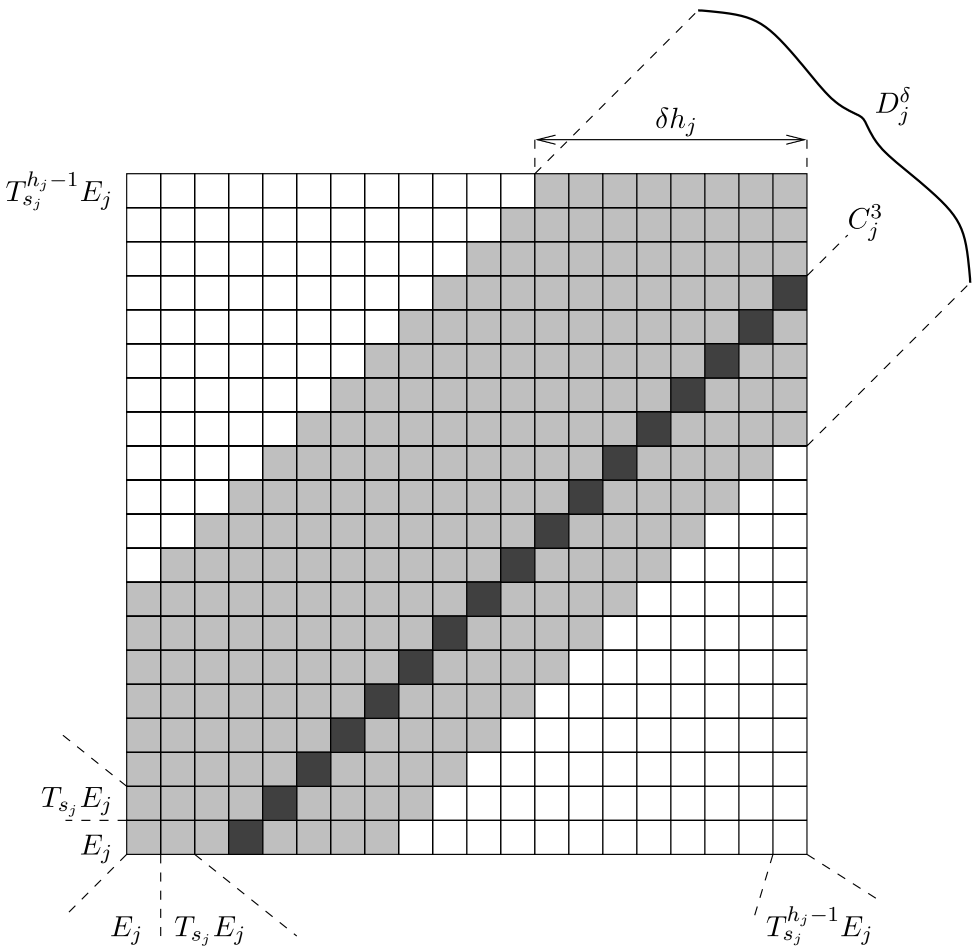

FIGURE 1. Columns and fat diagonals in $X \times X$

4. Approximation theorem

Recall from Section 2 that, given a flow $(T_t)_{t\in\mathbb{R}}$, Δ^t stands for the self-joining supported by the graph of T_{-t}.

LEMMA 4.1. *Let ν be an ergodic joining of the rank-one flow $(T_t)_{t\in\mathbb{R}}$. Let $0<\delta<1$ be such that*

$$\ell_\delta := \lim_j \nu(D_j^\delta) > 0. \tag{4.1}$$

Then there exists a sequence (k_j) with $-\delta h_j \le k_j \le \delta h_j$ such that

$$\Delta^{k_j s_j}(\,\cdot\,|C_j^{k_j}) \xrightarrow[j\to\infty]{w} \nu.$$

PROOF. Our strategy is the following: First we prove that the joining ν can be approximated by sums of parts of off-diagonal measures, then applying the Choice Lemma we find a sequence of parts tending to ν.

By definition of D_j^δ, we have

$$\nu\Big(D_j^\delta \,\vartriangle\, (T_{s_j}\times T_{s_j})D_j^\delta\Big) \le \frac{C}{h_j}.$$

It follows that for any fixed p, the sets D_j^δ are asymptotically $T_{s_p}\times T_{s_p}$-invariant: Indeed, since $T_{s_p} = T_{s_j}^{s_p/s_j}$ where s_p/s_j is an integer when $j\ge p$, we get

$$\nu\Big(D_j^\delta \,\vartriangle\, (T_{s_p}\times T_{s_p})D_j^\delta\Big) \le \frac{s_p}{s_j}\frac{C}{h_j} \xrightarrow[j\to\infty]{} 0$$

(recall that $s_j h_j \to \infty$).

Let λ be a limit measure of $\nu(\,\cdot\mid D_j^\delta)$. Then λ is $T_{s_p}\times T_{s_p}$-invariant for each p, by (4.1), λ is absolutely continuous with respect to ν, and $\frac{d\lambda}{d\nu}\le\frac{1}{\ell_\delta}<\infty$. By Lemma 2.1, it follows that $\lambda=\nu$. Hence we have

$$\nu(\,\cdot\mid D_j^\delta) \xrightarrow[j\to\infty]{w} \nu. \tag{4.2}$$

We now prove that

$$\sum_{k=-[\delta h_j]}^{[\delta h_j]} \nu(C_j^k|D_j^\delta)\Delta^{ks_j}(\,\cdot\,|C_j^k) \xrightarrow[j\to\infty]{w} \nu. \tag{4.3}$$

For arbitrary measurable sets A, B we can find ξ_j-measurable sets A_j, B_j such that

$$\varepsilon_j := \mu(A \vartriangle A_j) + \mu(B \vartriangle B_j) \to 0.$$

We have

$$\sum_k \nu(C_j^k|D_j^\delta)\Delta^{ks_j}(A\times B|C_j^k) - \nu(A\times B) = M_1+M_2+M_3+M_4,$$

where

$$
\begin{aligned}
M_1 &:= \sum_k \nu(C_j^k|D_j^\delta)\left(\Delta^{ks_j}(A\times B|C_j^k) - \Delta^{ks_j}(A_j\times B_j|C_j^k)\right),\\
M_2 &:= \sum_k \nu(C_j^k|D_j^\delta)\Delta^{ks_j}(A_j\times B_j|C_j^k) - \nu(A_j\times B_j|D_j^\delta),\\
M_3 &:= \nu(A_j\times B_j|D_j^\delta) - \nu(A\times B|D_j^\delta),\\
M_4 &:= \nu(A\times B|D_j^\delta) - \nu(A\times B).
\end{aligned}
$$

The density of the projections of the measure $\Delta^{ks_j}(\,\cdot\,|C_j^k)$ with respect to μ is bounded by $(1-\delta)^{-1}$. Hence $M_1 \le \varepsilon_j/(1-\delta)$.

Since A_j, B_j are ξ_j-measurable,

$$\nu(A_j\times B_j|C_j^k) = \Delta^{ks_j}(A_j\times B_j|C_j^k),$$

and we get $M_2 = 0$.

The absolute value of the third term M_3 can be bounded above as follows

$$|M_3| \le \nu(D_j^\delta)^{-1}\nu\Big((A_j\times B_j)\,\triangle\,(A\times B)\Big) \le \frac{\varepsilon_j}{\nu(D_j^\delta)} \to 0.$$

The last term M_4 goes to zero as $j\to\infty$ by (4.2), and this ends the proof of (4.3).

To apply the Choice Lemma for the measures $\nu_j^k = \Delta^{ks_j}(\,\cdot\,|C_j^k)$ and $a_j^k = \nu(C_j^k|D_j^\delta)$, it remains to check the first hypothesis of the lemma. By construction of the columns C_j^k, we have for any measurable subset $A\in X\times X$ and all $k\in\{-[\delta h_j],\dots,[\delta h_j]\}$,

$$\left|\Delta^{ks_j}(T_{s_j}\times T_{s_j}A|C_j^k) - \Delta^{ks_j}(A|C_j^k)\right| < \frac{C}{h_j} \tag{4.4}$$

where C is a constant. We get the desired result by setting $d_j := \dfrac{C}{s_j h_j}$.

The Choice Lemma then gives a sequence (k_j) with $-\delta h_j \le k_j \le \delta h_j$ such that $\Delta^{k_j s_j}(\,\cdot\,|C_j^{k_j}) \xrightarrow[j\to\infty]{w} \nu$. ∎

THEOREM 4.2. *Let a flow $T=(T_t)_{t\in\mathbb{R}}$ be of rank-one and ν be an ergodic self-joining of $(T_t)_{t\in\mathbb{R}}$. Then there is a sequence (k_j) such that $\Delta^{k_j s_j} \xrightarrow[j\to\infty]{w} \frac12\nu + \frac12\nu'$ for some self-joining ν': For all measurable sets A, B*

$$\mu(A\cap T_{s_j}^{k_j}B) \to \frac12\nu(A\times B) + \frac12\nu'(A\times B).$$

PROOF. For any $1/2<\delta<1$, we have

$$\lim_{j\to\infty}\nu(D_j^\delta) > 1-2(1-\delta) = 2\delta-1>0.$$

Hence we can apply Lemma 4.1 for any $1/2<\delta<1$. By a diagonal argument, we get the existence of (k_j) and $(\delta_j)\searrow \frac12$ with $-\delta_j h_j \le k_j \le \delta_j h_j$ such that

$$\Delta^{k_j s_j}\left(\,\cdot\,|C_j^{k_j}\right) \xrightarrow[j\to\infty]{w} \nu.$$

Let us decompose $\Delta^{k_j s_j}$ as

$$\Delta^{k_j s_j} = \Delta^{k_j s_j}\left(\,\cdot\,|C_j^{k_j}\right)\Delta^{k_j s_j}(C_j^{k_j}) + \Delta^{k_j s_j}\left(\,\cdot\,|X\times X\setminus C_j^{k_j}\right)\left(1-\Delta^{k_j s_j}(C_j^k)\right).$$

Since $\liminf_{j\to\infty} \Delta^{k_j s_j}(C_j^{k_j}) \geq 1/2$, we get the existence of some self-joining ν' such that

$$\Delta^{k_j s_j} \xrightarrow[j\to\infty]{w} \frac{1}{2}\nu + \frac{1}{2}\nu'.$$

□

Corollary 4.3. *A mixing rank-one flow has minimal self-joinings of order two.*

Proof. Let ν be an ergodic self-joining of a mixing rank-one flow $(T_t)_{t\in\mathbb{R}}$. Let (k_j) be the sequence given by Theorem 4.2. If $|k_j s_j| \to \infty$, since T is mixing we have

$$\Delta^{k_j s_j} \xrightarrow[j\to\infty]{w} \mu \times \mu,$$

hence $\mu\times\mu = \frac{1}{2}\nu + \frac{1}{2}\nu'$ for some self-joining ν'. The ergodicity of $\mu\times\mu$ then implies that $\mu\times\mu = \nu$. Otherwise, along some subsequence we have $k_j s_j \to s$ for some real number s. Then $\Delta^s = \frac{1}{2}\nu + \frac{1}{2}\nu'$ for some self-joining ν', and again the ergodicity of Δ^s yields $\nu = \Delta^s$. Thus T has minimal self-joinings of order two.. □

5. Weak closure theorem for rank-one flows

Lemma 5.1 (Weak Closure Lemma). *If the automorphism S commutes with the rank-one flow $(T_t)_{t\in\mathbb{R}}$, then there exist $1/2 \leq d \leq 1$, a sequence (k_j) of integers and a sequence of measurable sets (Y_j) such that, for all measurable sets A, B*

$$\mu(A \cap T_{s_j}^{k_j} B \cap Y_j) \to d\,\mu(A \cap SB),$$

where Y_j has the form

$$Y_j^{d,-} := \bigsqcup_{0\leq i<dh_j} T_{s_j}^i E_j \quad \text{or} \quad Y_j^{d,+} := \bigsqcup_{(1-d)h_j<i\leq h_j} T_{s_j}^i E_j.$$

Proof. This lemma is a consequence of the proof of Theorem 4.2, when the joining ν is equal to Δ_S. Given a sequence $(\delta_j) \searrow \frac{1}{2}$, the proof provides a sequence (k_j) where $-\delta_j h_j \leq k_j \leq \delta_j h_j$, such that $\Delta^{k_j s_j}(\,\cdot\,|C_j^{k_j}) \xrightarrow[j\to\infty]{w} \Delta_S$, and $\Delta^{k_j s_j}(C_j^{k_j})$ converges to some number $d \geq 1/2$. Let $Y_j^{k_j}$ be the projection on the first coordinate of $C_j^{k_j}$, that is

$$Y_j^{k_j} = \begin{cases} \bigsqcup_{i=k_j}^{h_j} T_{s_j}^i E_j & \text{if } k_j \geq 0 \\ \bigsqcup_{i=0}^{h_j+k_j} T_{s_j}^i E_j & \text{if } k_j < 0. \end{cases}$$

We then have $\Delta^{k_j s_j}(\,\cdot\,|C_j^{k_j}) = \Delta^{k_j s_j}(\,\cdot\,|Y_j^{k_j} \times X)$, and $\mu(Y_j^{k_j}) = \Delta^{k_j s_j}(C_j^{k_j}) \to d$. This yields, for all measurable sets A, B,

$$\mu(A \cap T_{s_j}^{k_j} B \cap Y_j^{k_j}) \to d\,\mu(A \cap SB).$$

If there exist infinitely many j's such that $k_j \geq 0$, then along this subsequence, we have

$$\mu\left(Y_j^{k_j} \vartriangle Y_j^{d,+}\right) \xrightarrow[j\to\infty]{} 0,$$

since $(h_j - k_j)/h_j \to d$. A similar result holds along the subsequence of j's such that $k_j < 0$, with $Y_j^{d,+}$ replaced by $Y_j^{d,-}$. □

THEOREM 5.2 (Weak Closure Theorem for rank-one flows). *If the automorphism S commutes with the rank-one flow $(T_t)_{t\in\mathbb{R}}$, then there exists a sequence of integers (k_j) such that $\Delta^{k_j s_j} \to \Delta_S$: For all measurable sets A, B,*

$$\mu(A \cap T_{s_j}^{k_j} B) \to \mu(A \cap SB).$$

PROOF. We fix T and consider the set of real numbers d for which the conclusion in the statement of Lemma 5.1 holds. It is easy to show by a diagonal argument that this set is closed. Hence we consider its maximal element, which we still denote by d. (If $d = 1$, the theorem is proved.)

So we start from the following statement: We have a sequence of sets $\{Y_j\}$, of the form given in Lemma 5.1, such that for all measurable A, B

$$\mu(A \cap T_{s_j}^{k_j} B \cap Y_j) \to d\mu(A \cap SB).$$

Then a similar statement holds when Y_j is replaced by SY_j: Indeed, since S commutes with T and μ is invariant by S, we have

$$\begin{aligned}\mu(A \cap T_{s_j}^{k_j} B \cap SY_j) &= \mu(S^{-1}A \cap T_{s_j}^{k_j} S^{-1}B \cap Y_j) \\ &\xrightarrow[j\to\infty]{} d\,\mu(S^{-1}A \cap SS^{-1}B) = d\,\mu(A \cap SB).\end{aligned}$$

Let λ be a limit point for the sequence of probability measures $\{\nu_j\}$ defined on $X \times X$ by

$$\nu_j(A \times B) := \frac{1}{\mu(Y_j \cup SY_j)} \mu\left(A \cap T_{s_j}^{k_j} B \cap (Y_j \cup SY_j)\right).$$

Then $\lambda \leq 2\,\Delta_S$. Moreover, the measure λ is invariant by $T_{s_p} \times T_{s_p}$ for all p. Indeed, for $j \geq p$, we have

$$\mu(T_{s_p} Y_j \vartriangle Y_j) = \mu(T_{s_j}^{s_p/s_j} Y_j \vartriangle Y_j)$$

which is of order $\frac{s_p}{s_j h_j}$, hence vanishes as $j \to \infty$. Since Δ_S is an ergodic measure for the flow $\{T_t \times T_t\}$, we can apply Lemma 2.1, which gives $\lambda = \Delta_S$. We obtain

$$\mu\left(A \cap T_{s_j}^{k_j} B \cap (Y_j \cup SY_j)\right) \to u\,\mu(A \cap SB),$$

where $u := \lim_j \mu(Y_j \cup SY_j)$ (if the limit does not exist, then we consider some subsequence of $\{j\}$).

Our aim is to show that $u = 1$, which will end the proof of the theorem. Let us introduce

$$W_j := \left(\bigsqcup_{0\leq i\leq h_j} T_{s_j}^i E_j \right) \setminus Y_j.$$

Assume that $u < 1$, then (denoting by Y^c the complementary of $Y \subset X$)

$$\lim_j \Delta_S(W_j \times W_j) = \lim_j \mu(W_j \cap SW_j) = \lim_j \mu(Y_j^c \cap SY_j^c) = 1 - u > 0.$$

Let us consider the case where Y_j has the form $Y_j^{d,-} = \bigsqcup_{0\leq i<dh_j} T_{s_j}^i E_j$. Then $W_j = \bigsqcup_{dh_j\leq i\leq h_j} T_{s_j}^i E_j$, and we define for any $\delta' < 1 - d$

$$W_j(\delta') := \bigsqcup_{(1-\delta')h_j<i\leq h_j} T_{s_j}^i E_j \subset W_j.$$

In the same way, if Y_j has the form $Y_j^{d,+} = \bigsqcup_{(1-d)h_j<i\leq h_j} T_{s_j}^i E_j$, we set for $\delta' < 1-d$

$$W_j(\delta') := \bigsqcup_{0<i<\delta' h_j} T_{s_j}^i E_j \subset W_j.$$

In both cases, note that

$$\Delta_S\Big((W_j \times W_j) \setminus (W_j(\delta') \times W_j(\delta'))\Big) \leq 2(1-d-\delta').$$

Thus, for δ' close enough to $1-d$, we get

$$\limsup_j \Delta_S\Big(W_j(\delta') \times W_j(\delta')\Big) \geq 1-u-2(1-d-\delta') > 0.$$

Since $W_j(\delta') \times W_j(\delta') \subset D_j^{\delta'}$, this ensures that

$$\limsup \Delta_S(D_j^{\delta'}) > 0.$$

Lemma 4.1 then provides a sequence (k'_j) with $-\delta' h_j \leq k'_j \leq \delta' h_j$, such that

$$\Delta^{k'_j s_j}(\,\cdot\,|C_j^{k'_j}) \xrightarrow[j\to\infty]{w} \Delta_S,$$

and the projections $Y_j^{k'_j}$ of $C_j^{k'_j}$ on the first coordinate satisfy

$$\lim_j \mu(Y_j^{k'_j}) \geq 1-\delta' > d,$$

which contradicts the maximality of d. Hence $u = 1$. □

6. Rigidity of factors of rank-one flows

LEMMA 6.1. *Let $\mathcal{F}$ be a non-trivial factor of a rank-one flow $(T_t)_{t\in\mathbb{R}}$. Then there exist $1/2 \leq d \leq 1$, a sequence of integers (k_j) with $|k_j s_j| \not\to 0$ and a sequence of measurable sets (Y_j) such that, for all measurable sets $A, B \in \mathcal{F}$*

$$\mu(A \cap T_{s_j}^{k_j} B \cap Y_j) \to d\,\mu(A \cap B),$$

where Y_j has the form

$$Y_j^{d,-} := \bigsqcup_{0\leq i<dh_j} T_{s_j}^i E_j \quad \text{or} \quad Y_j^{d,+} := \bigsqcup_{(1-d)h_j<i\leq h_j} T_{s_j}^i E_j.$$

PROOF. We start with the relatively independent joining above the factor $\mathcal{F}$ (see Section 2). Since $\mathcal{F}$ is a non-trivial factor, $\mu \otimes_{\mathcal{F}} \mu \neq \Delta$, hence we can consider an ergodic component ν such that $\nu(\{(x,x), x \in X\}) = 0$. Observe however that for any sets $A, B \in \mathcal{F}$, we have $\nu(A \times B) = \mu(A \cap B)$.

We repeat the proof of Lemma 5.1 with ν in place of Δ_S. This provides sequences (k_j) and (Y_j) and a real number $1/2 \leq d \leq 1$, such that for all measurable sets A, B

$$\mu(A \cap T_{s_j}^{k_j} B \cap Y_j) \to d\,\nu(A \times B).$$

If we had $k_j s_j \to 0$, then the left-hand side would converge to $d\,\mu(A \cap B)$, which would give $\nu(A \times B) = \mu(A \cap B)$ for all $A, B \in \mathcal{A}$, and this would contradict the hypothesis that ν gives measure 0 to the diagonal. □

THEOREM 6.2. *Let $\mathcal{F}$ be a non-trivial factor of a rank-one flow $(T_t)_{t\in\mathbb{R}}$. Then there exists a sequence of integers (k_j) with $|k_j s_j| \to \infty$ such that, for all measurable sets $A, B \in \mathcal{F}$*

$$\mu(A \cap T_{s_j}^{k_j} B) \to \mu(A \cap B).$$

PROOF. Again we fix some ergodic component ν such that $\nu(\{(x,x), x \in X\}) = 0$. We consider the maximal number d for which the statement of Lemma 6.1 is true. We thus have a sequence of sets $\{Y_j\}$, of the form given in Lemma 6.1, with $\mu(Y_j) \to d$, and a sequence (k_j) of integers with the required condition, such that for all $A, B \in \mathcal{F}$

$$\frac{1}{\mu(Y_j)} \mu(A \cap T_{s_j}^{k_j} B \cap Y_j) \to \mu(A \cap B).$$

Now we rewrite this in the form

$$\forall A, B \in \mathcal{F}, \quad \frac{1}{\mu(Y_j)} \mathbb{E}_\mu \left[\mathbb{1}_A \mathbb{1}_{T_{s_j}^{k_j} B} \mathbb{1}_{Y_j} \right] \to \mu(A \cap B). \tag{6.1}$$

In the above equation, one can replace $\mathbb{1}_{Y_j}$ by $\phi_j(x) := \mathbb{E}_\nu[\mathbb{1}_{Y_j}(x')|x]$: Indeed, since ν coincides with Δ on $\mathcal{F} \otimes \mathcal{F}$, we have $\mathbb{1}_A(x') = \mathbb{1}_A(x)$ and $\mathbb{1}_{T_{s_j}^{k_j} B}(x') = \mathbb{1}_{T_{s_j}^{k_j} B}(x)$ ν-a.s. Hence,

$$\mathbb{E}_\mu \left[\mathbb{1}_A \mathbb{1}_{T_{s_j}^{k_j} B} \mathbb{1}_{Y_j} \right] = \mathbb{E}_\nu \left[\mathbb{1}_A(x) \mathbb{1}_{T_{s_j}^{k_j} B}(x) \mathbb{1}_{Y_j}(x') \right] = \mathbb{E}_\mu \left[\mathbb{1}_A(x) \mathbb{1}_{T_{s_j}^{k_j} B}(x) \phi_j(x) \right].$$

We note that

$$\mathbb{E}_\mu \left[|\phi_j - \phi_j \circ T_{s_j}| \right] \le \mu(Y_j \vartriangle T_{s_j} Y_j) = O\left(\frac{1}{h_j}\right). \tag{6.2}$$

For any $\varepsilon > 0$, let

$$U_j^\varepsilon := \{x : \phi_j(x) > \varepsilon\}.$$

We would like to prove that (6.1) remains valid with $\mathbb{1}_{Y_j}$ replaced by $\mathbb{1}_{U_j^\varepsilon}$ for ε small enough. To this end, we need almost-invariance of U_j^ε under T_{s_j}, which does not seem to be guaranteed for arbitrary ε. Therefore, we use the following technical argument to find a sequence (ε_j) for which the desired result holds.

Fix $\varepsilon > 0$ small enough so that $\mu(U_j^\varepsilon) > \mu(Y_j)/2$ for all large j. By Lemma 3.2, we can assume that $s_j^2 h_j \to \infty$. Let $\delta_j = o(s_j)$ such that $(\delta_j h_j)^{-1} = o(s_j)$. We divide the interval $[\varepsilon/2, \varepsilon]$ into $\varepsilon/(4\delta_j)$ disjoint subintervals of length $2\delta_j$. One of these subintervals, called I_j, satisfy

$$\mu(\{x : \phi_j(x) \in I_j\}) \le \frac{4\delta_j}{\varepsilon}. \tag{6.3}$$

Let us call ε_j the center of the interval I_j. Observe that

$$\mu\left(U_j^{\varepsilon_j} \vartriangle T_{s_j} U_j^{\varepsilon_j}\right) \le \mu(\{x : |\phi_j(x) - \varepsilon_j| < \delta_j\}) + \mu\left(\{x : |\phi_j(x) - \phi_j(T_{s_j}(x))| \ge \delta_j\}\right).$$

By (6.3) and (6.2), we get that

$$\mu\left(U_j^{\varepsilon_j} \vartriangle T_{s_j} U_j^{\varepsilon_j}\right) = O\left(\delta_j + \frac{1}{\delta_j h_j}\right) = o(s_j). \tag{6.4}$$

Taking a subsequence if necessary, we can assume that the sequence of probability measures λ_j, defined by

$$\forall A, B \in \mathcal{A}, \qquad \lambda_j(A \times B) := \frac{1}{\mu(U_j^{\varepsilon_j})} \mathbb{E}_\mu \left[\mathbb{1}_A \mathbb{1}_{T_{s_j}^{k_j} B} \mathbb{1}_{U_j^{\varepsilon_j}} \right],$$

converges to some probability measure λ, which is invariant by $T_{s_p} \times T_{s_p}$ for all p by (6.4). Recall that $\mu(U_j^{\varepsilon_j}) > \mu(Y_j)/2$ and that $\mathbb{1}_{U_j^{\varepsilon_j}} \leq \phi_j/\varepsilon_j$. Then, since $\varepsilon_j > \varepsilon/2$, we have $\lambda|_{\mathcal{F}\otimes\mathcal{F}} \leq \frac{4}{\varepsilon}\Delta|_{\mathcal{F}\otimes\mathcal{F}}$. Since $\Delta|_{\mathcal{F}\otimes\mathcal{F}}$ is an ergodic measure for the flow $\{T_t \times T_t\}|_{\mathcal{F}\otimes\mathcal{F}}$, we can apply Lemma 2.1, which gives $\lambda|_{\mathcal{F}\otimes\mathcal{F}} = \Delta|_{\mathcal{F}\otimes\mathcal{F}}$. This means that (6.1) remains valid with $\mathbb{1}_{Y_j}$ replaced by $\mathbb{1}_{U_j^{\varepsilon_j}}$.

The analogue of (6.1) is also valid when we replace $\mathbb{1}_{Y_j}$ by $\mathbb{1}_{Y_j \cup U_j^{\varepsilon_j}}$: Indeed, we also have the almost-invariance property

$$\mu\left((Y_j \cup U_j^{\varepsilon_j}) \vartriangle T_{s_j}(Y_j \cup U_j^{\varepsilon_j})\right) = o(s_j)$$

and $\mathbb{1}_{Y_j \cup U_j^{\varepsilon_j}} \leq \mathbb{1}_{Y_j} + \mathbb{1}_{U_j^{\varepsilon_j}}$. We conclude by a similar argument.

Since ε can be taken arbitrarily small, we can now use a diagonal argument to show that (6.1) remains valid with $\mathbb{1}_{Y_j}$ replaced by $\mathbb{1}_{Y_j \cup U_j^{\varepsilon_j}}$ where the sequence (ε_j) now satisfies $\varepsilon_j \to 0$. Hence, taking a subsequence if necessary to ensure that $\mu(Y_j \cup U_j^{\varepsilon_j})$ converges to some number u, we get

$$\forall A, B \in \mathcal{F}, \quad \mathbb{E}_\mu\left[\mathbb{1}_A \mathbb{1}_{T_{s_j}^{k_j} B} \mathbb{1}_{Y_j \cup U_j^{\varepsilon_j}}\right] \to u\mu(A \cap B).$$

It now remains to prove that $u = 1$, which we do by repeating the end of the proof of Theorem 5.2. Assume that $u < 1$. Let us introduce

$$W_j := \left(\bigsqcup_{0 \leq i \leq h_j} T_{s_j}^i E_j\right) \setminus Y_j.$$

We have

$$\lim_j \nu(W_j \times W_j) = \lim_j \nu(Y_j^c \times Y_j^c) = \lim_j \mathbb{E}_\mu\left[\mathbb{1}_{Y_j^c}(1 - \phi_j)\right].$$

Observe that $(1 - \phi_j) \geq \mathbb{1}_{(U_j^{\varepsilon_j})^c} - \varepsilon_j$. Hence

$$\lim_j \nu(W_j \times W_j) \geq \lim_j \mathbb{E}_\mu\left[\mathbb{1}_{Y_j^c}\mathbb{1}_{(U_j^{\varepsilon_j})^c}\right] = 1 - u > 0.$$

Let us consider the case where Y_j has the form $Y_j^{d,-} = \bigsqcup_{0 \leq i < dh_j} T_{s_j}^i E_j$. Then $W_j = \bigsqcup_{dh_j \leq i \leq h_j} T_{s_j}^i E_j$, and we define for any $\delta' < 1 - d$

$$W_j(\delta') := \bigsqcup_{(1-\delta')h_j < i \leq h_j} T_{s_j}^i E_j \subset W_j.$$

In the same way, if Y_j has the form $Y_j^{d,+} = \bigsqcup_{(1-d)h_j < i \leq h_j} T_{s_j}^i E_j$, we set for $\delta' < 1-d$

$$W_j(\delta') := \bigsqcup_{0 < i < \delta' h_j} T_{s_j}^i E_j \subset W_j.$$

In both cases, note that

$$\nu\Big((W_j \times W_j) \setminus (W_j(\delta') \times W_j(\delta'))\Big) \leq 2(1 - d - \delta').$$

thus, for δ' close enough to $1 - d$, we get

$$\limsup_j \nu\Big(W_j(\delta') \times W_j(\delta')\Big) \geq 1 - u - 2(1 - d - \delta') > 0.$$

Since $W_j(\delta') \times W_j(\delta') \subset D_j^{\delta'}$, this ensures that

$$\limsup \nu(D_j^{\delta'}) > 0.$$

Lemma 4.1 then provides a sequence (k'_j) with $-\delta' h_j \leq k'_j \leq \delta' h_j$, such that

$$\Delta^{k'_j s_j}(\,\cdot\,|C_j^{k'_j}) \xrightarrow[j\to\infty]{w} \nu.$$

In particular, $\Delta^{k'_j s_j}(\,\cdot\,|C_j^{k'_j})|_{\mathcal{F}\otimes\mathcal{F}} \xrightarrow[j\to\infty]{w} \Delta|_{\mathcal{F}\otimes\mathcal{F}}$. Since the projections $Y_j^{k'_j}$ of $C_j^{k'_j}$ on the first coordinate satisfy

$$\lim_j \mu(Y_j^{k'_j}) \geq 1 - \delta' > d,$$

this contradicts the maximality of d. Hence $u = 1$. $\square$

7. King's theorem for flat-roof rank-one flow

We consider a rank-one flow $(T_t)_{t\in\mathbb{R}}$. We say that $(T_t)_{t\in\mathbb{R}}$ has *flat roof* if we can choose the sequence $\xi_j = \{E_j, T_{s_j}E_j, \ldots, T_{s_j}^{h_j-1}E_j, X \setminus \bigsqcup_{k=0}^{h_j-1} T_{s_j}^k E_j\}$ in the definition such that

$$\frac{\mu\left(T_{s_j}^{h_j}E_j \vartriangle E_j\right)}{\mu(E_j)} \xrightarrow[j\to\infty]{} 0.$$

THEOREM 7.1. *Let $(T_t)_{t\in\mathbb{R}}$ be a flat-roof rank-one flow, and ν be an ergodic self-joining of $(T_t)_{t\in\mathbb{R}}$. Then there exists a sequence (k_j) such that $\Delta^{k_j s_j} \xrightarrow[j\to\infty]{w} \nu$.*

PROOF. Let us defined, for $0 \leq k \leq h_j - 1$

$$a_k^j := \nu\left(T_{s_j}^k E_j \times E_j\right) \quad \text{and} \quad b_k^j := \nu\left(E_j \times T_{s_j}^{h_j-k}E_j\right).$$

We claim that the flat-roof property implies

$$h_j \sum_{k=1}^{h_j-1} |a_k^j - b_k^j| \xrightarrow[j\to\infty]{} 0. \tag{7.1}$$

Indeed, by invariance $a_k^j = \nu\left(T_{s_j}^{h_j}E_j \times T_{s_j}^{h_j-k}E_j\right)$. Hence

$$|a_k^j - b_k^j| \leq \nu\left((T_{s_j}^{h_j}E_j \vartriangle E_j) \times T_{s_j}^{h_j-k}E_j\right),$$

and

$$\sum_{k=1}^{h_j-1} |a_k^j - b_k^j| \leq \nu\left((T_{s_j}^{h_j}E_j \vartriangle E_j) \times X\right) = \mu\left((T_{s_j}^{h_j}E_j \vartriangle E_j)\right).$$

The claim follows, since $\mu(E_j) \sim 1/h_j$.

We gather the columns C_j^k in pairs, defining for $1 \leq k \leq h_j - 1$, $G_j^k := C_j^k \sqcup C_j^{k-h_j}$. (See Figure 2.) We also set $G_j^0 := C_j^0$. Note that $\nu(G_j^k) = (h_j - k)a_k^j + kb_k^j$. Observe also that

$$\nu\left(\bigsqcup_{k=0}^{h_j-1} G_j^k\right) = \nu\left(\bigsqcup_{k=0}^{h_j-1} T_{s_j}^k E_j \times \bigsqcup_{k=0}^{h_j-1} T_{s_j}^k E_j\right) \xrightarrow[j\to\infty]{} 1.$$

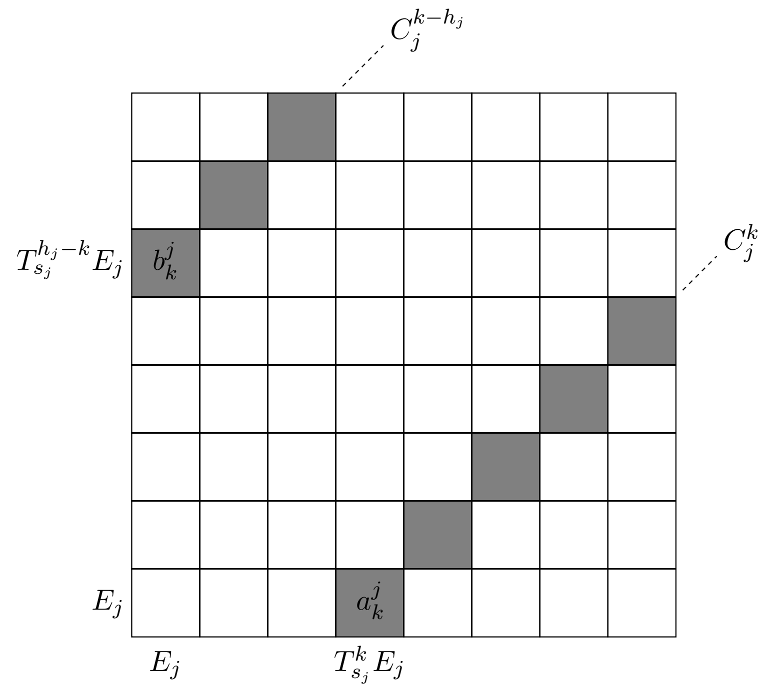

FIGURE 2. The union of C_j^k and $C_j^{k-h_j}$ is denoted by G_j^k.

Hence,

$$\sum_{k=0}^{h_j-1} \nu(G_j^k)\,\nu(\,\cdot\,|G_j^k) \xrightarrow[j\to\infty]{w} \nu. \tag{7.2}$$

We claim that, using the flat-roof property, we can in the above equation replace $\nu(\,\cdot\,|G_j^k)$ by Δ^{ks_j}. Let A and B be ξ_j-measurable sets, which are unions of $T_{s_j}^i E_j$ $(0 \le i \le h_j-1)$. We denote by r_k (respectively ℓ_k) the number of elementary cells of the form $T_{s_j}^{i_1}E_j \times T_{s_j}^{i_2}E_j$ which are contained in $A\times B$ and which belong to the column C_j^k (respectively $C_j^{k-h_j}$). We have

$$\nu(A\times B|G_j^k)\nu(G_j^k) = \ell_k b_k^j + r_k a_k^j. \tag{7.3}$$

Moreover, we will show that the flat-roof property ensures the existence of a sequence (ε_j) with $\varepsilon_j \xrightarrow[j\to\infty]{} 0$ such that

$$\left|\Delta^{ks_j}(A\times B) - \frac{\ell_k + r_k}{h_j}\right| \le \varepsilon_j. \tag{7.4}$$

Indeed, let us cut A into $A_1 := A\cap\bigsqcup_{0\le i\le k-1} T_{s_j}^i E_j$ and $A_2 := A\cap\bigsqcup_{k\le i\le h_j-1} T_{s_j}^i E_j$. We have

$$\Delta^{ks_j}(A_2\times B) = r_k\mu(E_j),$$

and

$$\Delta^{ks_j}(A_1\times B) = \ell_k\Delta^{ks_j}(E_j\times T_{s_j}^{h_j-k}E_j) + \Delta^{ks_j}\left((A_1\times B)\setminus C_j^{k-h_j}\right).$$

Recalling that $\Delta^{ks_j}(E_j \times T_{s_j}^{h_j-k} E_j) = \mu(E_j \cap T_{s_j}^{h_j} E_j)$, we get

$$\text{(7.5)}\quad \Delta^{ks_j}(A\times B) = (r_k+\ell_k)\mu(E_j) - \ell_k\mu(E_j \setminus T_{s_j}^{h_j} E_j) + \Delta^{ks_j}\left((A_1 \times B) \setminus C_j^{k-h_j}\right).$$

The second term of the right-hand side is bounded by $h_j\mu(E_j \Delta T_{s_j}^{h_j} E_j)$, which goes to 0 by the flat-roof property. To treat the last term, we consider the particular case $A = B = \bigsqcup_{0\le i\le h_j-1} T_{s_j}^i E_j$, for which this last term is maximized. We have then

$$1 - \Delta^{ks_j}(A \times B) \le 2\mu\left(X \setminus \bigsqcup_{0\le i\le h_j-1} T_{s_j}^i E_j\right) \xrightarrow[j\to\infty]{} 0.$$

On the other hand, (7.5) gives

$$\Delta^{ks_j}\left((A_1 \times B) \setminus C_j^{k-h_j}\right) = \Delta^{ks_j}(A \times B) - h_j\mu(E_j) + k\mu(E_j \setminus T_{s_j}^{h_j} E_j).$$

Since $h_j\mu(E_j) \to 1$, and $k\mu(E_j \setminus T_{s_j}^{h_j} E_j) \le h_j\mu(E_j \Delta T_{s_j}^{h_j} E_j) \to 0$, we get that the last term of (7.5) goes to 0 uniformly with respect to k, A and B. It follows that

$$\left|\Delta^{ks_j}(A \times B) - (\ell_k + r_k)\mu(E_j)\right| \xrightarrow[j\to\infty]{} 0,$$

uniformly with respect to k, A and B. This concludes the proof of (7.4).

Equations (7.4) and (7.3) give

$$\begin{aligned}
&\sum_{k=0}^{h_j-1} \left|\nu(A \times B|G_j^k) - \Delta^{ks_j}(A \times B))\right|\nu(G_j^k)\\
&\le \sum_{k=0}^{h_j-1} |a_k^j - b_k^j| \left|\ell_k - \frac{k}{h_j}(\ell_k + r_k)\right| + \varepsilon_j\\
&\le h_j \sum_{k=0}^{h_j-1} |a_k^j - b_k^j| + \varepsilon_j
\end{aligned}$$

which goes to 0 as $j \to \infty$ by (7.1).

Recalling (7.2), we obtain

$$\sum_{k=0}^{h_j-1} \nu(G_j^k)\Delta^{ks_j} \xrightarrow[j\to\infty]{w} \nu.$$

It remains to apply the Choice Lemma to conclude the proof of the theorem. □

8. $\mathbb{Z}^n$-Rank-one action

We consider now an action of $\mathbb{Z}^n$ ($n \ge 1$). For $k \in \mathbb{Z}^n$, we denote by $k(1), \dots, k(n)$ its coordinates.

Definition 8.1. A $\mathbb{Z}^n$-action $\{T_k\}_{k\in\mathbb{Z}^n}$ is of rank one if there exists a sequence (ξ_j) of partitions converging to the partition onto points, where ξ_j is of the form

$$\xi_j = \left\{ (T_k E_j)_{k\in R_j}, X \setminus \bigsqcup_k T_k E_j \right\},$$

and R_j is a rectangular set of indices:

$$R_j = \{0, \dots, h_j(1) - 1\} \times \cdots \times \{0, \dots, h_j(n) - 1\}.$$

Note that the above definition corresponds to so-called $\mathcal{R}$*-rank one actions* defined in [**RŞ11**] with the additional condition that the shapes in the sequence $\mathcal{R}$ be rectangles. The sequence (ξ_j) in the above definition being fixed, we define as for the rank-one flows the notions of columns and fat diagonals: For any $k \in \mathbb{Z}^n$, we set

$$C_j^k := \bigsqcup_{\substack{r,\ell\in R_j\\ r-\ell=k}} T_r E_j \times T_\ell E_j,$$

and given $0<\delta<1$,

$$D_j^\delta := \bigsqcup_{k:\ \prod_i (h_j(i)-|k(i)|)\geq (1-\delta)\prod_i h_j(i)} C_j^k.$$

LEMMA 8.2. *For any self-joining ν of the rank-one action $\{T_k\}_{k\in\mathbb{Z}^n}$, for any $\delta > 1-1/2^n$, we have*

$$\liminf_{j\to\infty} \nu(D_j^\delta) > 0.$$

PROOF. We can find $\varepsilon > 0$, small enough such that

$$\left(\frac{1}{2}-\varepsilon\right)^n > 1-\delta.$$

Let $r\in\mathbb{Z}^n$ be such that

$$\forall i,\ \left(\frac{1}{2}-\varepsilon\right) h_j(i) < r(i) < \left(\frac{1}{2}+\varepsilon\right) h_j(i).$$

Then, for any $\ell\in R_j$, we have for all i: $|r(i)-\ell(i)| < \left(\frac{1}{2}+\varepsilon\right) h_j(i)$. Hence

$$\prod_i \Big(h_j(i) - |r(i)-\ell(i)|\Big) > (1-\delta)\prod_i h_j(i),$$

which means that for any $\ell\in R_j$, the column $C_j^{r-\ell}$ is contained in D_j^δ. It follows that

$$\left(\bigsqcup_{r:\ \forall i,\ |r(i)-h_j(i)/2|<\varepsilon h_j(i)} T_r E_j\right) \times \left(\bigsqcup_{\ell\in R_j} T_\ell E_j\right) \subset D_j^\delta.$$

We then get

$$\liminf_{j\to\infty} \nu(D_j^\delta) \geq \liminf_{j\to\infty} \mu\left(\bigsqcup_{r:\ \forall i,\ |r(i)-h_j(i)/2|<\varepsilon h_j(i)} T_r E_j\right) = (2\varepsilon)^n.$$

□

We can now state the analogue of Theorem 4.2 for $\mathbb{Z}^n$-rank-one action, which was first proved by A.A. Pavlova in [**Pav08**].

THEOREM 8.3. *Let ν be an ergodic self-joining of the $\mathbb{Z}^n$-rank-one action $\{T_k\}_{k\in\mathbb{Z}^n}$. Then we can find a sequence (k_j) in $\mathbb{Z}^n$ and some self-joining ν' such that $\Delta^{k_j} \xrightarrow[j\to\infty]{w} \frac{1}{2^n}\nu + \left(1-\frac{1}{2^n}\right)\nu'$: For all measurable sets A, B*

$$\mu(A\cap T_{k_j}B) \to \frac{1}{2^n}\nu(A\times B) + \left(1-\frac{1}{2^n}\right)\nu'(A\times B).$$

PROOF. The proof follows the same lines as for Theorem 4.2. First note that Lemma 4.1 can be easily adapted to the $\mathbb{Z}^n$-situation. Hence, by Lemma 8.2, using a diagonal argument, we get the existence of (k_j) and $(\delta_j) \searrow 1-\frac{1}{2^n}$ with $C_j^{k_j} \subset D_j^{\delta_j}$ such that

$$\Delta^{k_j}\left(\cdot\,|C_j^{k_j}\right) \xrightarrow[j\to\infty]{w} \nu.$$

To conclude, it remains to prove that $\liminf \Delta^{k_j}(C_j^{k_j}) \geq 1/2^n$. To this aim, we count the number of pairs (r,ℓ) such that $T_r E_j \times T_\ell E_j \subset C_j^{k_j}$. We can easily check that these are exactly the pairs (r,ℓ) such that, for all $1 \leq i \leq n$, there exists $m(i) \in \{0,\ldots,h_j(i)-1-|k_j(i)|\}$ with

$$\big(r(i),\ell(i)\big) = \begin{cases} \big(k_j(i)+m(i),m(i)\big) & \text{if } k_j(i) \geq 0 \\ \big(m(i),-k_j(i)+m(i)\big) & \text{otherwise.}\end{cases}$$

Hence $\Delta^{k_j}(C_j^{k_j}) = \prod_i \Big(h_j(i)-1-|k_j(i)|\Big)\mu(E_j)$. Using the fact that $C_j^{k_j} \subset D_j^{\delta_j}$, we get the desired result. □

When $n \geq 2$, it is known that the Weak Closure Theorem fails (counterexamples have been given in [**DK02, DS09**]). However, as a consequence of Theorem 8.3, we get the following:

COROLLARY 8.4 (Partial Weak Closure Theorem for $\mathbb{Z}^n$-rank-one action). *Let S be an automorphism commuting with the $\mathbb{Z}^n$-rank-one action $\{T_k\}_{k\in\mathbb{Z}^n}$. Then we can find a sequence (k_j) in $\mathbb{Z}^n$ and some self-joining ν' such that*

$$\Delta^{k_j} \xrightarrow[j\to\infty]{} \frac{1}{2^n}\Delta_S + \left(1-\frac{1}{2^n}\right)\nu'.$$

Moreover, if $S \notin \{T_k \;\; k \in \mathbb{Z}^n\}$, then $\{T_k\}_{k\in\mathbb{Z}^n}$ is partially rigid: There exists a sequence (k'_ℓ) in $\mathbb{Z}^n$ with $|k'_\ell| \to \infty$ such that for all measurable sets A and B

$$\liminf_{\ell\to\infty} \mu\left(A \cap T_{k'_\ell} B\right) \geq \frac{1}{2^{2n}}\mu(A\cap B).$$

PROOF. The first part is a direct application of Theorem 8.3 with $\nu = \Delta_S$. If moreover $S \notin \{T_k \;\; k \in \mathbb{Z}^n\}$, then the sequence (k_j) of the theorem must satisfy $|k_j| \to \infty$. Let us enumerate the cylinder sets as $\{A_0, A_1, \ldots, A_\ell, \ldots\}$. Let (ε_ℓ) be a sequence of positive numbers decreasing to zero. For any ℓ, we can find a large enough integer $j_1(\ell)$ such that for all cylinder sets $A, B \in \{A_0, A_1, \ldots, A_\ell\}$,

$$\mu\left(T_{k_{j_1(\ell)}} A \cap SB\right) \geq \left(\frac{1}{2^n}-\varepsilon_\ell\right)\mu(SA\cap SB) = \left(\frac{1}{2^n}-\varepsilon_\ell\right)\mu(A\cap B).$$

Then, we can find a large enough integer $j_2(\ell)$ with $|j_2(\ell)| > 2|j_1(\ell)|$ such that for all cylinder sets $A, B \in \{A_0, A_1, \ldots, A_\ell\}$,

$$\mu\left(T_{k_{j_1(\ell)}} A \cap T_{k_{j_2(\ell)}} B\right) \geq \left(\frac{1}{2^n}-\varepsilon_\ell\right)\mu(T_{k_{j_1(\ell)}} A\cap SB).$$

It follows that for all $\ell \geq 0$ and all cylinder sets $A, B \in \{A_0, A_1, \ldots, A_\ell\}$,

$$\mu\left(A \cap T_{k_{j_2(\ell)}-k_{j_1(\ell)}} B\right) \geq \left(\frac{1}{2^n}-\varepsilon_\ell\right)^2 \mu(A\cap B).$$

This proves the result announced in the corollary when A and B are cylinder sets with $k'_\ell := k_{j_2(\ell)} - k_{j_1(\ell)}$, and this extends in a standard way to all measurable sets. □

The counterexample given in [**DS09**] also shows that the rigidity of factors is no more valid when $n \geq 2$. Theorem 8.3 only ensures the partial rigidity of factors of $\mathbb{Z}^n$-rank-one actions.

Corollary 8.5 (Partial rigidity of factors of $\mathbb{Z}^n$-rank-one action). *Let $\mathcal{F}$ be a non-trivial factor of the $\mathbb{Z}^n$-rank-one action $\{T_k\}_{k\in\mathbb{Z}^n}$. Then there exists a sequence (k_j) in $\mathbb{Z}^n$ with $|k_j| \to \infty$ such that, for all measurable sets $A, B \in \mathcal{F}$*

$$\liminf \mu(A \cap T_{k_j} B) \geq \frac{1}{2^n} \mu(A \cap B).$$

Proof. This is a direct application of Theorem 8.3 where ν is an ergodic component of the relatively independent joining above the factor $\mathcal{F}$. □

References

[DK02] T. Downarowicz and J. Kwiatkowski, *Weak closure theorem fails for $\mathbb{Z}^2$-actions*, Studia Math. **153** (2002), no. 2, 115–125. MR1948920 (2003h:37006)

[DS09] Tomasz Downarowicz and Jacek Serafin, *Phenomena in rank-one $\mathbb{Z}^2$-actions*, Studia Math. **192** (2009), no. 3, 281–294. MR2504842 (2011b:37005)

[Kin86] Jonathan L. King, *The commutant is the weak closure of the powers, for rank-1 transformations*, Ergodic Theory Dynam. Systems **6** (1986), no. 3, 363–384. MR863200 (88a:28021)

[Kin88] ———, *Joining-rank and the structure of finite rank mixing transformations*, J. Analyse Math. **51** (1988), 182–227. MR963154 (89k:28009)

[Kin01] ———, *Flat stacks, joining-closure and genericity*, Preprint, 2001.

[KS66] A. B. Katok and A. M. Stepin, *Approximation of ergodic dynamical systems by periodic transformations*, Dokl. Akad. Nauk SSSR **171** (1966), 1268–1271. MR0207961 (34:7774a)

[KT91] Jonathan L. King and Jean-Paul Thouvenot, *A canonical structure theorem for finite joining-rank maps*, J. Analyse Math. **56** (1991), 211–230. MR1243104 (94i:28013)

[Orn72] Donald S. Ornstein, *On the root problem in ergodic theory*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory (Berkeley, Calif.), Univ. California Press, 1972, pp. 347–356. MR0399415 (53:3259)

[Pav08] A.A. Pavlova, *Self-joinings and weak closure of rank one actions*, Thesis, Moscow State University, 2008.

[Pri01] A. A. Prikhod′ko, *Stochastic constructions of mixing systems of positive local rank*, Mat. Zametki **69** (2001), no. 2, 316–319. MR1830233 (2002c:37004)

[RŞ11] E. Arthur Robinson, Jr. and Ayşe A. Şahin, *Rank-one $\mathbb{Z}^d$ actions and directional entropy*, Ergodic Theory Dynam. Systems **31** (2011), no. 1, 285–299. MR2755932

[Rud79] Daniel J. Rudolph, *An example of a measure preserving map with minimal self-joinings, and applications*, J. Analyse Math. **35** (1979), 97–122. MR555301 (81e:28011)

[Ryz92] Valery Ryzhikov, *Mixing, rank and minimal self-joining of actions with invariant measure*, Mat. Sb. **183** (1992), no. 3, 133–160. MR1180921 (93m:28020)

[Ryz08] ———, *Rank, rigidity of factors, and weak closure of measure-preserving $\mathbb{Z}^n$-actions*, Moscow Univ. Math. Bull. **63** (2008), no. 4, 135–137.

[Ryz10] ———, *Self-joinings of rank-one actions and applications*, Semin. et Congr. de la SMF **20** (2010), 193–206.

[Tik06] S. V. Tikhonov, *Embeddings of lattice actions in flows with multidimensional time*, Mat. Sb. **197** (2006), no. 1, 97–132. MR2230134 (2007i:22028)

[Zei93] Paul Zeitz, *The centralizer of a rank-one flow*, Israel J. Math. **84** (1993), no. 1-2, 129–145. MR1244663 (95c:28017)

Laboratoire de Mathématiques Raphaël Salem, Université de Rouen, CNRS, Avenue de l'Université, F76801 Saint Étienne du Rouvray

E-mail address: `Elise.Janvresse@univ-rouen.fr`

Laboratoire de Mathématiques Raphaël Salem, Université de Rouen, CNRS, Avenue de l'Université, F76801 Saint Étienne du Rouvray

E-mail address: `Thierry.de-la-Rue@univ-rouen.fr`

Moscow State University, Faculty of Mechanics and Mathematics, Leninskie Gory, Moscow, 119991 Russia

E-mail address: `vryzh@mail.ru`

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11255

Random walks on random horospheric products

Vadim A. Kaimanovich and Florian Sobieczky

Abstract. By developing the entropy theory of random walks on equivalence relations and analyzing the asymptotic geometry of horospheric products we describe the Poisson boundary for random walks on random horospheric products of trees.

Introduction

Horospheric products of trees were first introduced in the work of Diestel and Leader [**DL01**] in an attempt to answer a question of Woess [**Woe91**] on existence of vertex-transitive graphs not quasi-isometric to Cayley graphs. Although the fact that the Diestel–Leader graphs indeed provide such an example was only recently proved by Eskin, Fisher and Whyte [**EFW07**], in the meantime the construction of Diestel and Leader attracted a lot of attention because of its numerous interesting features (see [**Woe05, BNW08**] and the references therein). For instance, as it was observed by Woess, the horospheric product of two homogeneous trees of the same degree $p+1$ is isomorphic to the Cayley graph of the lamplighter group (the wreath product $\mathbb{Z}\wr\mathbb{Z}_p$) with respect to an appropriate generating set. This observation was the starting point for Bartholdi and Woess [**BW05a**] who showed that, along with lamplighter groups, horospheric products of homogeneous trees (not necessarily of the same degree!) provide one of very few examples of infinite graphs, for which all spectral invariants can be exhibited in an absolutely explicit form (and, in addition, the spectrum happens to be pure point).

The construction of horospheric products is very natural from a geometrical viewpoint. Namely, by choosing a point γ on the boundary ∂T of an infinite tree T one converts it into the genealogical tree generated by the "mythological progenitor" γ. The *Busemann cocycle* β_γ on T can be interpreted as the signed "generations gap" in the genealogical tree: its level sets are the "generations" in T as seen from γ. Given another pointed at infinity tree (T',γ'), the horospheric product (or, rather, products) of (T,γ) and (T',γ') are then the level sets of the aggregate cocycle $\beta_\gamma+\beta_{\gamma'}$ on $T\times T'$, and the Busemann cocycles determine a natural *"height cocycle"* on individual horospheric products. It is important for what follows that

2010 *Mathematics Subject Classification.* Primary 60J50; Secondary 05C81, 37A20.

Key words and phrases. Random walk, equivalence relation, Poisson boundary, asymptotic entropy, horospheric product.

The research of the first author was undertaken, in part, thanks to funding from the Canada Research Chairs program. Support from ERC grants 208091 GADA and 257110 RaWG is also gratefully acknowledged.

each horospheric product is endowed with two boundaries ("lower" and "upper") isomorphic to the punctured boundaries $\partial T \setminus \{\gamma\}$ and $\partial T' \setminus \{\gamma'\}$ of the trees T and T'.

In the previous paper [**KS10**] we considered the problem of *stochastic homogenization* for horospheric products. The approach that we used there was an implementation of the ideas from [**Kai03**]: to consider random graphs as leafwise graphs of an appropriate *graphed equivalence relation*; stochastic homogenization means that there is a probability measure invariant with respect to this relation. Here we are continuing to apply the ideas from [**Kai03**] to random horospheric products by looking at random walks on them, or, in view of the aforementioned reduction, at random walks along classes of graphed equivalence relations, leafwise graphs of which are horospheric products.

The problem we address is that of the boundary behaviour of such leafwise random walks, more precisely, the problem of identification of their *Poisson boundaries*. In the case of isotropic random walks on horospheric products of homogeneous trees this problem (or, actually, even the more general problem of describing the Martin boundary) was solved by Brofferio and Woess [**Woe05, BW05b, BW06**]. However, their approach (as is almost always the case with the Martin boundary) heavily depends on explicit estimates of the Green kernel only possible for highly symmetrical Markov chains. Following [**Kai03**], instead of this we use the *entropy theory*. Originally developed for dealing with the Poisson boundary of random walks on groups (see [**KV83, Kai00**] and the references therein), it is actually applicable in all situations when there is an appropriate probability path space endowed with a measure preserving time shift, in particular for random walks on equivalence relations in the presence of a global stationary measure. In this setup the entropy theory was already outlined by the first author in [**Kai98, Kai03**]; here we give a more detailed exposition.

In the group case the entropy theory produces not only the entropy criterion of *boundary triviality*, but also very efficient geometrical conditions for *identification of the Poisson boundary* ("ray" and "strip" approximations). Both these conditions readily carry over to random walks along classes of graphed equivalence relations as well. In order to apply them to horospheric products we establish the necessary geometrical ingredients. Namely, we completely characterize *geodesics* in horospheric products and give necessary and sufficient conditions for a sequence of points to be *regular*, i.e., to follow a geodesic with a sublinear deviation.

As a consequence we establish our **main result** (Theorem 2.22), according to which the Poisson boundary of a random walk on an equivalence relation graphed by horospheric products in the presence of a global stationary probability measure is completely determined by the *height drift* h (the expectation of the height cocycle): if $h = 0$, then a.e. leafwise Poisson boundary is trivial, whereas if $h \neq 0$ then a.e. leafwise Poisson boundary coincides with the corresponding (lower or upper, depending on the sign of h) boundary of the underlying horospheric product endowed with the corresponding limit (hitting) distribution. This description is in perfect keeping with the situation for horospheric products of homogeneous trees [**Woe05**] or for lamplighter groups [**Kai91**] (whose Cayley graphs for an appropriate choice of generators are horospheric products of homogeneous trees of the same degree [**Woe05**]).

The main result implies that for *reversible* random walks on random horospheric products the leafwise Poisson boundaries are almost surely trivial, because in this situation the height drift (being the expectation of an additive cocycle) vanishes. This is the case for simple random walks on stochastically homogeneous horospheric products considered in [**KS10**], in particular, for horospheric products of augmented Galton–Watson trees with the same offspring expectation.

On the other hand, although already lamplighter groups and horospheric products of homogeneous trees readily provide examples of random walks on horospheric products with *non-zero height drift*, it would be interesting to have more "probabilistically natural" examples of this kind.

It is worth mentioning in this respect that the *homesick simple random walk* with an integer parameter d on a pointed at infinity tree T can be interpreted as the projection of the usual simple random walk on the horospheric product of T and the homogeneous tree of degree $d+1$ (usually "homesickness" is defined with respect to a reference point inside the graph, e.g., see [**LPP96a**], but this definition in an obvious way adapts to pointed at infinity trees as well). For usual simple random walks on random Galton–Watson trees existence of a linear rate of escape was established in [**LPP95**] by using an explicit stationary measure on the space of trees. In the homesick case, although a linear rate of escape still exists [**LPP96b**], no such construction is known.

Yet another link between horospheric products and homesick random walks worth further investigation is provided by a rather unexpected behavior of the rate of escape of homesick random walks on the lamplighter group exhibited in [**LPP96b**] (although homesickness in [**LPP96b**] is defined with respect to the standard generating set rather than the one whose Cayley graph is a horospheric product).

Let us finally mention that our results (with rather straightforward modifications) carry over to horospheric products with more than two multipliers which were introduced in [**KW02**, p. 356] and further studied in [**BNW08**].

The paper has the following structure. In Section **1** we study the *asymptotic geometry of individual horospheric products*. After reminding the necessary definitions concerning trees (Section **1.A**) and their horospheric products (Section **1.B**), in Section **1.C** we reprove Bertacchi's formula [**Ber01**] for the distance in a horospheric product (Proposition 1.5). Our argument is somewhat different and provides an explicit description of geodesic segments in horospheric products, on the base of which we further describe geodesic rays and bilateral geodesics (Proposition 1.8 and Proposition 1.9, respectively). In Section **1.D** we give criteria for a sequence of points in a horospheric product to be regular (Theorem 1.11). Finally, in Section **1.E** we discuss boundaries of horospheric products.

Section **2** contains the probabilistic part of our arguments. We begin by reminding the basic definitions concerning graphed equivalence relations and random graphs (Section **2.A**). In Section **2.B** we discuss Markov chains along classes of an equivalence relation endowed with a quasi-invariant measure; the exposition here is based on [**Kai98**]. We express the action of the corresponding Markov operator on measures in terms of the leafwise transition probabilities and the Radon–Nikodym cocycle of the equivalence relation (Proposition 2.2) and give a necessary and sufficient condition for stationarity of a measure on the state space (Corollary 2.10). In particular, an invariant measure of a graphed equivalence relation becomes stationary for the leafwise simple random walk after multiplication by the density equal to the vertex degree function (Corollary 2.11).

In Section **2.C** we develop the *entropy theory for random walks on equivalence relation*. The exposition here follows the outlines given in [**Kai98, Kai03**] and is completely parallel to the entropy theory for random walks in random environment on groups [**Kai90**] (which, in turn, was inspired by the case of the usual random walks on groups [**KV83**]). First we prove that the leafwise tail and Poisson boundaries coincide $\mathbf{P}_x$ – mod 0 for a.e. initial point x (Theorem 2.13), after which we define the *asymptotic entropy* $\mathfrak{h}$ and prove that the leafiwse tail ($\equiv$ Poisson) boundaries are a.e. trivial if and only if $\mathfrak{h} = 0$ (Theorem 2.17). By passing to an appropriate *boundary extension* of the original equivalence relation [**Kai05**], Theorem 2.17 is also applicable to the problem of description of non-trivial Poisson boundaries of leafwise Markov chains. Indeed, a quotient of the Poisson boundary is *maximal* (i.e., coincides with the whole Poisson boundary) if and only if for almost all conditional chains determined by the points of this quotient the Poisson boundary is trivial. Thus, the criterion from Theorem 2.17 allows one to carry over the *ray* and the *strip* criteria used for identification of the Poisson boundary in the group case [**Kai00**] to the setup of random walks along classes of graphed equivalence relations.

Finally, in Section **2.D** we formulate and prove the main result of the present paper: the aforementioned description of Poisson boundaries of random walks along random horospheric products (Theorem 2.22).

1. Asymptotic geometry of horospheric products

1.A. Trees. We begin by recalling that a *tree* is a connected graph without cycles. Any two vertices x, y in a tree T can be joined with a unique segment $[x, y]$ which is *geodesic* with respect to the standard graph distance d. Throughout the paper we will only be considering trees "without leaves", i.e., such that the degree of any vertex is at least 2.

Any locally finite tree T has a natural *compactification* $\overline{T} = T \sqcup \partial T$ obtained in the following way: a sequence of vertices x_n which goes to infinity in T converges in this compactification if and only if for a certain ($\equiv$ any) reference point $o \in T$ the geodesic segments $[o, x_n]$ converge pointwise. Thus, for any reference point $o \in T$ the *boundary* ∂T can be identified with the space of geodesic rays issued from o (and endowed with the topology of pointwise convergence). There are many other equivalent descriptions of the boundary ∂T (and of the compactification $\overline{T}$), in particular, as the *space of ends* of T and as the *hyperbolic boundary* of T.

A tree T with a distinguished boundary point $\gamma \in \partial T$ is called *pointed at infinity* ($\equiv$ remotely rooted; in the terminology of Cartier [**Car72**] the point γ is called a "mythological progenitor"). We shall use the notation $\partial_{\odot} T = \partial T \setminus \{\gamma\}$ for the *punctured boundary* of a pointed at infinity tree (T, γ). A triple $T_o^\gamma = (T, o, \gamma)$ with $o \in T$ and $\gamma \in \partial T$ is a *rooted tree pointed at infinity*.

Any two geodesic rays converging to the same boundary point eventually meet, so that any boundary point $\gamma \in \partial T$ determines the associated additive $\mathbb{Z}$-valued *Busemann cocycle* on T. It is defined as

$$\beta_\gamma(x, y) = d(y, z) - d(x, z) \ , \tag{1.1}$$

where $z = x \curlywedge_\gamma y$ is the *confluence* of the geodesic rays $[x, \gamma)$ and $[y, \gamma)$, see Figure 1.

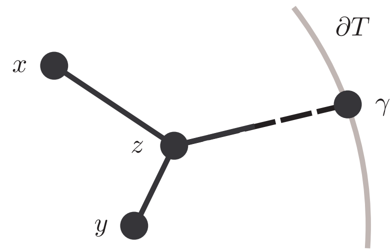

FIGURE 1

Obviously,

$$|\beta_\gamma(x, y)| \le d(x, y) \qquad \forall\, x, y \in T,\ \gamma \in \partial T\ .$$

The Busemann cocycle can also be defined as

$$\beta_\gamma(x, y) = \lim_{z \to \gamma} \big[d(y, z) - d(x, z) \big]\ ,$$

so that it is a "regularization" of the formal expression $d(y, \gamma) - d(x, \gamma)$. In the presence of a reference point $o \in T$ one can also talk about the *Busemann function*

$$b_\gamma(x) = \beta_\gamma(o, x)\ .$$

The level sets

$$H_k = \{x \in T : b_\gamma(x) = k\}$$

of the Busemann function ($\equiv$ of the Busemann cocycle) are called *horospheres* centered at the boundary point γ, see Figure 2.

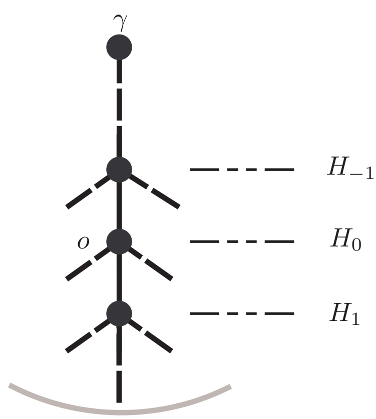

FIGURE 2

1.B. Horospheric products.

Definition 1.2. Let $T = (T, o, \gamma)$ and $T' = (T', o', \gamma')$ be two rooted trees pointed at infinity, and let $b = \beta_\gamma(o, \cdot)$, $b' = \beta_{\gamma'}(o', \cdot)$ be the corresponding Busemann functions. The *horospheric product* $T \upharpoonleft\!\downharpoonright T'$ is the graph with the vertex set

$$\{(x, x') \in T \times T' : b(x) + b'(x') = 0\}$$

and the edge set

$$\big\{\big((x, x'), (y, y')\big) : (x, y) \text{ and } (x', y') \text{ are edges in } T, T', \text{ respectively}\big\} .$$

Geometrically one can think about the horospheric products in the following way [**KW02**]. Draw the tree T' upside down next to T so that the respective horospheres $H_k(T)$ and $H_{-k}(T')$ are at the same level. Connect the two origins o, o' with an elastic spring. It can move along each of the two trees, may expand or contract, but must always remain horizontal. The vertex set of $T \upharpoonleft\!\downharpoonright T'$ consists then of all admissible positions of the spring. From a position (x, x') with $b(x)+b'(x') = 0$ the spring may move downwards to one of the "sons" of x and at the same time to the "father" of x', or upwards in an analogous way. Such a move corresponds to going to a neighbour (y, y') of (x, x'), see Figure 3.

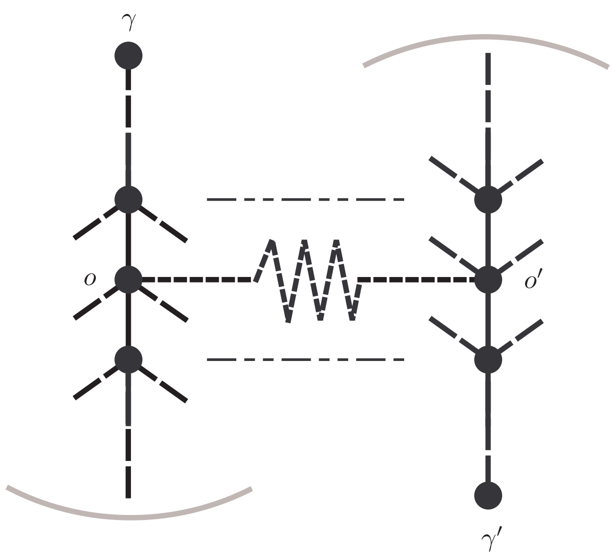

Figure 3

See [**Woe05, BNW08, KS10**] and the references therein for the historical background and recent works on horospheric products of trees (aka *Diestel–Leader graphs* or *horocyclic products*).

We shall use capital letters for denoting points of the horospheric product $T \upharpoonleft\!\downharpoonright T'$ (so that $X = (x, x')$ with $x \in T, x' \in T'$, etc.). In particular, we denote by $O = (o, o')$ the reference point in $T \upharpoonleft\!\downharpoonright T'$. The graph $T \upharpoonleft\!\downharpoonright T'$ is endowed with the *height cocycle*

$$\mathcal{B}(X, Y) = \beta_\gamma(x, y) = -\beta_{\gamma'}(x', y') . \tag{1.3}$$

For simplicity below we shall use the "height function" on $T \mathbin{\uparrow\!\downarrow} T'$

$$\overline{X} = -\mathcal{B}(O, X)$$

defined in accordance with Figure 3 (so that the "higher" is the level, the bigger is the value $\overline{X}$). In the same way, we put $\overline{x} = -b(x)$ and $\overline{x'} = b'(x')$ for any $x \in X, x' \in X'$, so that

$$\overline{X} = \overline{x} = \overline{x'} \qquad \forall\, X = (x, x') \in T \mathbin{\uparrow\!\downarrow} T' \;.$$

1.C. Geodesic segments and rays. Before establishing an explicit formula for the graph metric on the horospheric product $T \mathbin{\uparrow\!\downarrow} T'$ let us first notice that the sheer existence of the natural projections of $T \mathbin{\uparrow\!\downarrow} T'$ onto T, T' and $\mathbb{Z}$ (the latter by the height function) implies the obvious inequalities

$$d(x,y),\; d(x',y'),\; |\mathcal{B}(X,Y)| \le d(X,Y) \tag{1.4}$$

for all pairs of points $X = (x, x'), Y = (y, y') \in T \mathbin{\uparrow\!\downarrow} T'$.

Formula (1.6) below for the graph metric in $T \mathbin{\uparrow\!\downarrow} T'$ was first established by Bertacchi [**Ber01**, Proposition 3.1] (although Bertacchi considered horospheric products of homogeneous trees only, her arguments are actually valid in the general case as well). We shall give here a somewhat different argument, which, in particular, allows us to obtain an explicit description of all geodesics in $T \mathbin{\uparrow\!\downarrow} T'$.

PROPOSITION 1.5. *The graph distance in the horospheric product $T \mathbin{\uparrow\!\downarrow} T'$ is*

$$d(X,Y) = d(x,y) + d(x',y') - |\mathcal{B}(X,Y)| \;. \tag{1.6}$$

for all $X = (x, x')$, $Y = (y, y') \in T \mathbin{\uparrow\!\downarrow} T'$.

PROOF. Let Φ be a path joining the points X and Y. Then its projection φ to T (resp., its projection φ' to T') joins x and y (resp., x' and y'). Since T and T' are trees, φ and φ' should pass through all edges of the geodesics $[x, y]$ and $[x', y']$, respectively. Let $x \curlywedge y = x \curlywedge_\gamma y$ and $x' \curlywedge y' = x' \curlywedge_{\gamma'} y'$ be the confluences of the geodesic rays $[x, \gamma), [y, \gamma)$ and $[x', \gamma'), [y', \gamma')$, respectively. The geodesic $[x, y]$ in T consists of the ascending part $[x, x \curlywedge y]$ (along which the height increases) and the descending part $[x \curlywedge y, y]$ (along which the height decreases). In the same way the geodesic $[x', y']$ in T' consists of the descending part $[x', x' \curlywedge y']$ and the ascending part $[x' \curlywedge y', y']$, cf. Figure 1.

Thus,

(1.7) the projection ϕ of Φ to $\mathbb{Z}$ (by the height function) joins the points $\overline{X}, \overline{Y} \in \mathbb{Z}$ and contains all the edges (with the appropriate orientation!) from the oriented segments $\left[\,\overline{X}, \overline{x \curlywedge y}\,\right]$, $\left[\,\overline{x \curlywedge y}, \overline{Y}\,\right]$ and $\left[\,\overline{X}, \overline{x' \curlywedge y'}\,\right]$, $\left[\,\overline{x' \curlywedge y'}, \overline{Y}\,\right]$.

These segments do not overlap (if their orientation is taken into account), except for the oriented segment $\left[\,\overline{X}, \overline{Y}\,\right]$ which appears twice (see Figure 4, where $\overline{X} < \overline{Y}$).

Therefore, the length of Φ satisfies the inequality

$$|\Phi| \ge \left|\left[\,\overline{X}, \overline{x \curlywedge y}\,\right]\right| + \left|\left[\,\overline{x \curlywedge y}, \overline{Y}\,\right]\right| + \left|\left[\,\overline{X}, \overline{x' \curlywedge y'}\,\right]\right| + \left|\left[\,\overline{x' \curlywedge y'}, \overline{Y}\,\right]\right| - \left|\left[\,\overline{X}, \overline{Y}\,\right]\right| ,$$

the right-hand side of which being precisely the right-hand side of equation (1.6), so that we have proved the inequality

$$d(X,Y) \ge d(x,y) + d(x',y') - |\mathcal{B}(X,Y)| \;.$$

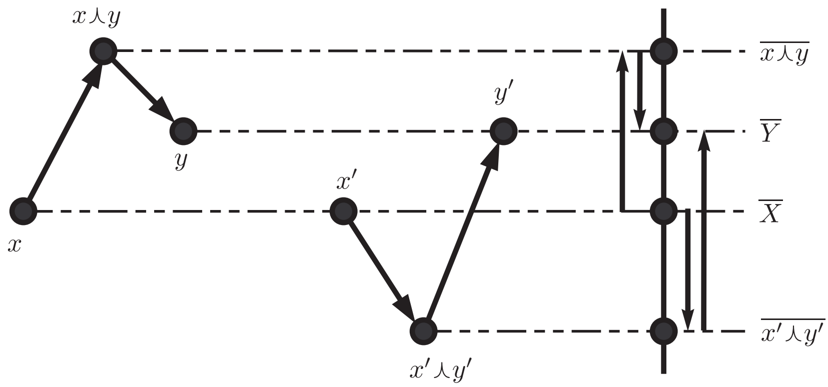

FIGURE 4

Now we shall show that paths of length $d(x,y) + d(x',y') - |\mathcal{B}(X,Y)|$ joining X and Y do exist, and, moreover, we shall explicitly describe all of them. Let us consider three cases.

(i) $\overline{X} < \overline{Y}$. Then there exists a unique path ϕ in $\mathbb{Z}$ of length $d(x,y)+d(x',y') - |\mathcal{B}(X,Y)|$ satisfying condition (1.7). Indeed, there is only one way to make a path joining $\overline{X}$ and $\overline{Y}$ by using (one time each) all the oriented edges contained in the segments from (1.7). This is the path

$$\phi = \left[\,\overline{X}, \overline{x' \curlywedge y'}\,\right] \; \left[\,\overline{x' \curlywedge y'}, \overline{x \curlywedge y}\,\right] \; \left[\,\overline{x \curlywedge y}, \overline{Y}\,\right] \; .$$

In order to lift it to $T \upharpoonleft\!\downharpoonright T'$ one has to choose a point $z \in T$ with $\overline{z} = \overline{x' \curlywedge y'}$ and such that z is a descendant of x (i.e., x lies on the geodesic ray $[z, \gamma)$), and a point $z' \in T'$ with $\overline{z'} = \overline{x \curlywedge y}$ and such that z' is a descendant of y'. Then the resulting path $\Phi = (\varphi, \varphi')$ with the projections

$$\varphi = [x,z]\,[z, x \curlywedge y]\,[x \curlywedge y, y] \; , \qquad \varphi' = [x', x' \curlywedge y']\,[x' \curlywedge y', z']\,[z', y']$$

is a geodesic joining X and Y, and all geodesics between X and Y have this form, see Figure 5.

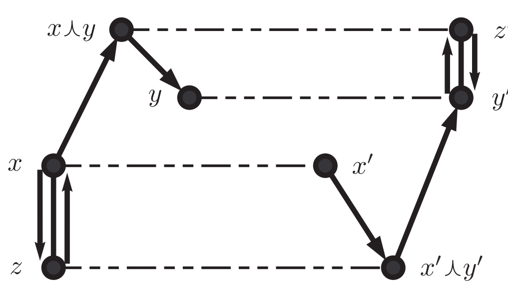

FIGURE 5

(ii) $\overline{X} > \overline{Y}$. *Mutatis mutandis*, the situation is precisely the same as in case (i), see Figure 6.

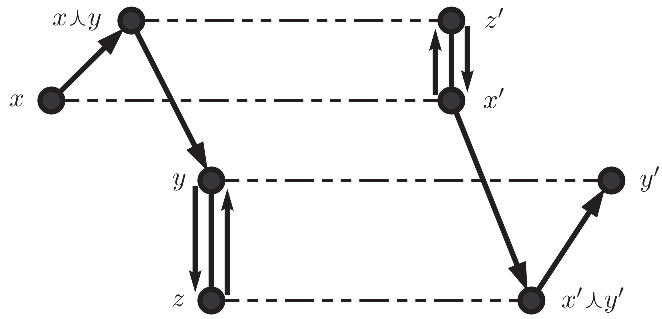

FIGURE 6

(iii) $\overline{X} = \overline{Y}$. In this case, due to absence of the $\left[\,\overline{X}, \overline{Y}\,\right]$ segment, there are two paths in $\mathbb{Z}$ satisfying condition (1.7):

$$\phi_1 = \left[\,\overline{X}, \overline{x' \curlywedge y'}\,\right]\ \left[\,\overline{x' \curlywedge y'}, \overline{x \curlywedge y}\,\right]\ \left[\,\overline{x \curlywedge y}, \overline{Y}\,\right]\ .$$

and

$$\phi_2 = \left[\,\overline{X}, \overline{x \curlywedge y}\,\right]\ \left[\,\overline{x \curlywedge y}, \overline{x' \curlywedge y'}\,\right]\ \left[\,\overline{x' \curlywedge y'}, \overline{Y}\,\right]\ .$$

Correspondingly, there are two types of geodesics joining X and Y, see Figure 7.

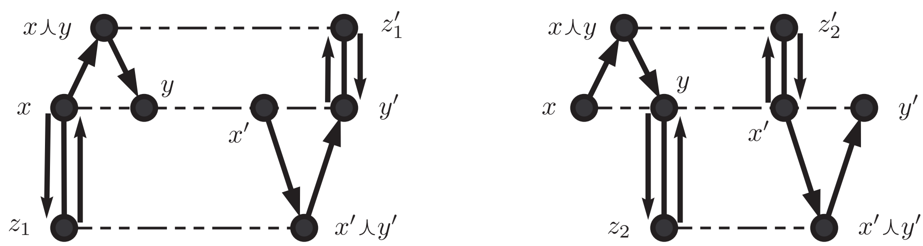

FIGURE 7

$\square$

By letting the length of geodesics go to infinity in the classification obtained in the proof of Proposition 1.5, we obtain the following description of geodesic rays and bilateral geodesics in $T \upharpoonleft\!\downharpoonright T'$:

PROPOSITION 1.8. *Given a point* $X = (x, x') \in T \upharpoonleft\!\downharpoonright T'$, *any pair* $(z, \omega') \in T \times \partial_{\odot} T'$ *with* $\overline{z} = \overline{x' \curlywedge \omega'}$ *determines a geodesic ray* $\Phi = (\varphi, \varphi')$ *in* $T \upharpoonleft\!\downharpoonright T'$ *issued from* (x, x') *with the projections*

$$\varphi = [x, z]\,[z, \gamma)\ , \qquad \varphi' = [x', \omega')\ ;$$

any pair $(\omega, z') \in \partial_\odot T \times T'$ *with* $\overline{z'} = \overline{x \curlywedge \omega}$ *determines a geodesic ray* $\Phi = (\varphi, \varphi')$ *in* $T \upharpoonleft\!\downharpoonright T'$ *issued from* (x, x') *with the projections*

$$\varphi = [x, \gamma) , \qquad \varphi' = [x', z'] \, [z', \omega') ,$$

and all geodesic rays in $T \upharpoonleft\!\downharpoonright T'$ *are of this form, see Figure* 8.

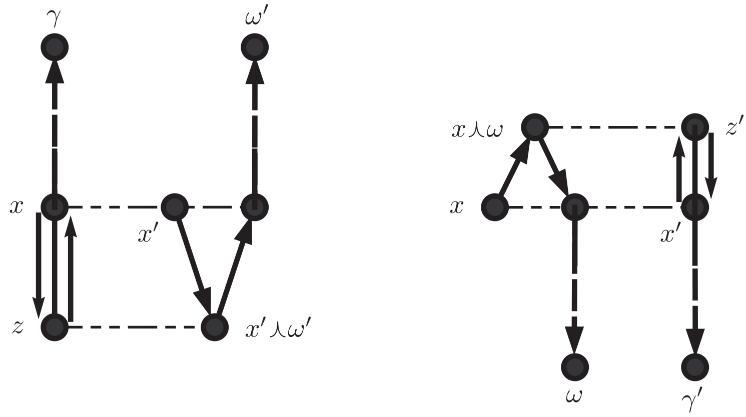

FIGURE 8

PROPOSITION 1.9. *All bilateral geodesics* Φ *in* $T \upharpoonleft\!\downharpoonright T'$ *belong to one of the following 3 classes described in terms of their projections* ϕ *to* $\mathbb{Z}$ *(by the height functions) and* φ, φ' *to* T, T'*, respectively:*

(i) ϕ *coincides with* $\mathbb{Z}$ *(run in either positive or negative direction), and* φ, φ' *are, respectively, the bilateral geodesics* (ω, γ) *and* (γ', ω') *(both run in either the positive or the negative direction) with* $\gamma \in \partial_\odot T$ *and* $\gamma' \in \partial_\odot T'$*;*

(ii) *There is* $h \in \mathbb{Z}$ *such that* ϕ *is the concatenation* $(-\infty, h][h, -\infty)$ *of two copies of the geodesic ray* $[h, -\infty)$ *run in the opposite directions,* φ *is the geodesic* (ω_1, ω_2) *for certain* $\omega_1 \neq \omega_2 \in \partial_\odot T$ *with* $\overline{\omega_1 \curlywedge \omega_2} = h$*, and* $\varphi' = (\gamma', x'][x', \gamma')$ *for a certain* $x' \in T'$ *with* $\overline{x'} = h$*;*

(iii) *The same as (ii) with* T *and* T' *exchanged:* $\phi = (\infty, h][h, \infty)$*,* $\varphi = (\gamma, x][x, \gamma)$ *for* $x \in T$ *with* $\overline{x} = h$*, and* $\varphi' = (\omega_1', \omega_2')$ *for* $\omega_1' \neq \omega_2' \in \partial_\odot T'$ *with* $\overline{\omega_1' \curlywedge \omega_2'} = h$*.*

1.D. Regular sequences.

DEFINITION 1.10. A sequence of points (x_n) in a connected graph X is called *regular* if there exist a geodesic ray Φ (with the natural parameterization) and a real number $\ell \geq 0$ (the *rate of escape*) such that

$$d(x_n, \Phi(\ell n)) = o(n) .$$

If $\ell = 0$, then (x_n) is called a *trivial* regular sequence.

This notion was introduced by Kaimanovich [**Kai89**] by analogy with the notion of Lyapunov regularity for sequences of matrices. Any non-trivial regular sequence in a tree T converges to a boundary point in the compactification $\overline{T} = T \sqcup \partial T$ (e.g., see [**CKW94**]).

THEOREM 1.11. *For a sequence of points $X_n = (x_n, x'_n)$ in the horospheric product of trees $T \updownarrows T'$ the following conditions are equivalent:*

(i) *The sequence (X_n) is regular with the rate of escape $\ell \geq 0$;*
(ii) *$d(X_n, X_{n+1}) = o(n)$ and $\overline{X_n} = hn + o(n)$ for a constant (which we call* height drift*) h with $|h| = \ell$;*
(iii) *The sequences (x_n) and (x'_n) are regular in the trees T and T', respectively, with the same rate of escape ℓ.*

PROOF. (i) $\Longrightarrow$ (ii). Obvious in view of inequalities (1.4) and the description of geodesic rays in $T \updownarrows T'$ from Proposition 1.8.

(ii) $\Longrightarrow$ (iii). By (1.4) condition (ii) for the sequence (X_n) implies that the analogous condition is satisfied for its projections (x_n) and (x'_n) to the trees T and T', respectively, i.e.,

$$d(x_n, x_{n+1}) = o(n) \ , \quad d(x'_n, x'_{n+1}) = o(n) \ , \quad \overline{x_n} = \overline{x'_n} = hn + o(n) \ .$$

Then by [**CKW94**, Proposition 1] the sequences $(x_n), (x'_n)$ are both regular with the rate of escape $|h|$.

(iii) $\Longrightarrow$ (i). If $\ell = 0$, then (X_n) is a trivial regular sequence by formula (1.6). If $\ell > 0$, then both (x_n) and (x'_n) are non-trivial regular sequences. Since $\overline{x_n} = \overline{x'_n}$, one of these sequences converges to the distinguished boundary point of the corresponding tree, whereas the other sequence converges to a "plain" boundary point. For instance, let $\lim x_n = \gamma$ and $\lim x'_n = \omega' \in \partial_\odot T'$ (which corresponds to positivity of the height drift h). Take the geodesic ray Φ in $T \updownarrows T'$ with the projections

$$\varphi = [o, z]\,[z, \gamma) \ , \qquad \varphi' = [o', \omega') \ ,$$

where $\overline{z} = \overline{o' \curlywedge \omega'}$ (cf. Proposition 1.8), then $d(X_n, \Phi(\ell n)) = o(n)$. $\square$

1.E. Boundaries of horospheric products. For the horospheric product $T \updownarrows T'$ there is a natural compactification

$$\overline{T \updownarrows T'} = T \updownarrows T' \cup \partial(T \updownarrows T') \tag{1.12}$$

obtained by embedding $T \updownarrows T'$ into the product $T \times T'$ and further taking the closure in $\overline{T} \times \overline{T'}$, where $\overline{T}$ and $\overline{T'}$ are the canonical compactifications of the trees T and T', respectively. One can easily check (see [**Ber01**, Proposition 3.2] for details) that the boundary of this compactification is

$$\partial(T \updownarrows T') = \left(\{\gamma\} \times \overline{T'}\right) \cup \left(\overline{T} \times \{\gamma'\}\right) \ .$$

Let

$$\partial_\uparrow(T \updownarrows T') = \{\gamma\} \times \partial_\odot T' \subset \partial(T \updownarrows T')$$

and

$$\partial_\downarrow(T \updownarrows T') = \partial_\odot T \times \{\gamma'\} \subset \partial(T \updownarrows T')$$

be, respectively, the *upper* and the *lower* boundaries of the horospheric product $T \updownarrows T'$. Similar pairs of boundaries arise for the dyadic-rational affine group [**Kai91**] or for *treebolic spaces* [**BSCSW11**].

Proposition 1.8 and Theorem 1.11 imply

PROPOSITION 1.13. *A non-trivial regular sequence in $T \updownarrows T'$ converges in the compactification* (1.12) *either to a point from $\partial_\uparrow(T \updownarrows T')$ (if the height drift is positive) or to a point from $\partial_\downarrow(T \updownarrows T')$ (if the height drift is negative).*

REMARK 1.14. It is *not* true (unlike in the tree case) that *any* boundary point is the limit of a certain non-trivial regular sequence. It might be an instructive exercise to look at the *Busemann compactification* of the horospheric product $T \updownarrows T'$ (which should not be difficult in view of the explicit descriptions of geodesics in $T \updownarrows T'$ obtained in Section **1.C**).

Proposition 1.9 describes which pairs of boundary points can be joined with a bilateral geodesic in $T \updownarrows T'$ (which is necessarily unique, as it follows from Proposition 1.9). In particular,

COROLLARY 1.15. *For any pair of boundary points from $\partial_\downarrow(T \updownarrows T') \times \partial_\uparrow(T \updownarrows T')$ there exists a unique bilateral geodesic in $T \updownarrows T'$ joining these points.*

2. Random horospheric products

2.A. Graphed equivalence relations and random graphs. In the present article we shall consider random graphs from the point of view of the theory of graphed measured equivalence relations. Let us remind the basic definitions (see [**FM77, Ada90, Kai97**]).

Let $(\mathcal{X}, \mu)$ be a Lebesgue measure space (below all the properties related to measure spaces will be understood *mod 0*, i.e., up to measure 0 subsets). A *partial transformation* of $(\mathcal{X}, \mu)$ is a measure class preserving bijection between two measurable subsets of $\mathcal{X}$. An equivalence relation $R \subset \mathcal{X} \times \mathcal{X}$ is called *discrete measured* if it is generated by an at most countable family of partial transformations. Then there exists a multiplicative *Radon–Nikodym cocycle* $\Delta = \Delta_\mu : R \to \mathbb{R}_+$ such that for any partial transformation $f : A \to B$ whose graph is contained in R

$$\Delta(x, y) = \frac{d\, f^{-1}\mu}{d\mu}(x) = \frac{d\mu}{d\, f\mu}(y) \ .$$

Alternatively, the Radon–Nikodym cocycle can be defined as the Radon–Nikodym ratio of the left and the right counting measures on R:

$$\Delta(x, y) = \frac{d\check{\mathrm{M}}}{d\mathrm{M}}(x, y) \ ,$$

where the *left counting measure* M on R is the result of integration of the counting measures $\#_x$ on the classes of the equivalence relation (considered as the fibers of the *left projection* $\pi : (x, y) \to x$ from R onto $\mathcal{X}$) against the measure μ on the state space $\mathcal{X}$:

$$d\mathrm{M}(x, y) = d\mu(x) d\#_x(y) = d\mu(x) \ ,$$

and the *right counting measure* $\check{\mathrm{M}}$ is the image of the left one under the *involution*

$$(x, y) \mapsto (y, x) \ .$$

If the Radon–Nikodym cocycle Δ is identically 1, then the measure μ is called *R-invariant* ($\equiv$ the equivalence relation R *preserves* the measure μ).

A (non-oriented) *graph structure* on a discrete measured equivalence relation $(\mathcal{X}, \mu, R)$ is determined by a measurable symmetric subset $K \subset R \setminus \mathsf{diag}$. The result of the restriction of this graph structure to an equivalence class $[x]$ gives the *leafwise graph* denoted by $[x]^K$ (by analogy with the theory of foliations classes of a discrete equivalence relation are often called *leaves*). We shall call $(\mathcal{X}, \mu, R, K)$ a *graphed equivalence relation*. We shall always deal with the graph structures which are *locally finite*, i.e., any vertex has only finitely many neighbours, and denote by deg

the integer valued function which assigns to any point $x \in \mathcal{X}$ the degree (valency) of x in the graph $[x]^K$. We shall also always assume that the graph structure is *leafwise connected*, i.e., a.e. leafwise graph $[x]^K$ is connected. Let us denote by $[x]^K_\bullet = ([x]^K, x)$ the graph $[x]^K$ rooted at the point x. Thus, we have the map $x \mapsto [x]^K_\bullet$ from $\mathcal{X}$ to the space of connected locally finite rooted graphs $\mathcal{G}$ endowed with the usual ball-wise convergence topology. In particular, if μ is a probability measure, then its image under the above map is a probability measure on the space of rooted graphs $\mathcal{G}$, i.e., a *random rooted graph*.

2.B. Random walks on equivalence relations.

DEFINITION 2.1 ([**Kai98**]). A *random walk along equivalence classes* of a discrete measured equivalence relation $(\mathcal{X}, \mu, R)$ is determined by a measurable family of leafwise transition probabilities $\{\pi_x\}_{x\in X}$, so that any π_x is concentrated on the equivalence class of x, and

$$(x, y) \mapsto p(x, y) = \pi_x(y)$$

is a measurable function on $R \subset \mathcal{X} \times \mathcal{X}$. By

$$p^n(x, y) = \pi_x^n(y) , \qquad n \geq 1 ,$$

we shall denote the corresponding n-step transition probabilities which are then also measurable as a function on R.

Since the measure class of μ is preserved by the equivalence relation R, the associated Markov operator P on the space $L^\infty(\mathcal{X}, \mu)$ is well-defined (cf. Proposition 2.2 below). The dual operator then acts on the space of measures λ absolutely continuous with respect to μ (notation: $\lambda \prec \mu$). Following a probabilistic tradition, we shall denote this action by $\lambda \mapsto \lambda P$. The density of the measure λP with respect to μ can be explicitly described in terms of the density of λ and of the Radon–Nikodym cocycle $\Delta = \Delta_\mu$ of the measure μ.

PROPOSITION 2.2 ([**Kai98**]). *For any σ-finite measure $\lambda \prec \mu$*

$$\frac{d\lambda P}{d\mu}(y) = \sum_{x\in[y]} p(x, y)\, \Delta(y, x)\, \frac{d\lambda}{d\mu}(x) . \tag{2.3}$$

PROOF. Let us run the Markov chain determined by the operator P from the initial (time 0) distribution λ. Then the time 1 distribution is, by definition, the measure λP, and the joint distribution of the positions of the chain at times 0 and 1 is the measure

$$d\Pi(x, y) = d\lambda(x)\, p(x, y) , \tag{2.4}$$

which is obviously absolutely continuous with respect to the counting measure M. The corresponding Radon–Nikodym derivative is

$$\frac{d\Pi}{d\mathrm{M}}(x, y) = \frac{d\lambda}{d\mu}(x)\, p(x, y) . \tag{2.5}$$

Since the left and the right counting measures are equivalent,

$$\frac{d\Pi}{d\check{\mathrm{M}}}(x, y) = \frac{d\Pi}{d\mathrm{M}}(x, y)\, \frac{d\mathrm{M}}{d\check{\mathrm{M}}}(x, y) = \frac{d\lambda}{d\mu}(x)\, p(x, y)\, \Delta(y, x) . \tag{2.6}$$

On the other hand, since λP is the result of the right projection of the measure Π to $\mathcal{X}$,

$$\frac{d\Pi}{d\check{\mathrm{M}}}(x,y) = \frac{d\lambda P}{d\mu}(y)\,\check{p}(y,x)\ , \tag{2.7}$$

where $\check{p}(\cdot,\cdot)$ are the corresponding *cotransition probabilities* (cf. formula (2.5)), whence summation of (2.6) over $x \in [y]$ yields the claim. □

REMARK 2.8. If the measure λ is infinite, then the density from formula (2.3) may well be infinite on a set of positive measure μ; however, even in this case the measure λP is absolutely continuous with respect to μ in the sense that any null set of μ is also null with respect to λP.

Comparison of (2.5) and (2.7) leads to the following useful formula relating transition and cotransition probabilities:

$$\frac{d\lambda}{d\mu}(x)p(x,y) = \frac{d\lambda P}{d\mu}(y)\check{p}(y,x)\Delta(x,y)\ ,$$

or, in a somewhat informal way,

$$d\lambda(x)p(x,y) = d\lambda P(y)\check{p}(y,x)\ ,$$

which is what one could expect.

Given a measure $\lambda \prec \mu$, we denote by $\{\lambda_x\}_{x\in\mathcal{X}}$ the family of leafwise measures on the equivalence classes of $\mathcal{X}$ defined as

$$\lambda_x(y) = \frac{d\lambda(y)}{d\mu(x)} = \frac{d\lambda}{d\mu}(y)\Delta(x,y)\ . \tag{2.9}$$

The measures λ_x corresponding to different equivalent points x are obviously all proportional.

COROLLARY 2.10. *A measure $\lambda \prec \mu$ is P-stationary (i.e., $\lambda = \lambda P$) if and only if the leafwise measures λ_x (2.9) are almost surely stationary with respect to the transition probabilities $p(\cdot,\cdot)$.*

If the measure λ is stationary, then, as it follows from a comparison of formulas (2.6) and (2.7), the cotransition probabilities

$$\check{p}(y,x) = p(x,y)\Delta_\lambda(y,x)\ ,$$

where Δ_λ is the Radon–Nikodym cocycle of the measure λ, determine another Markov chain along equivalence classes with the same stationary measure λ. It is called the *time reversal* of the original random walk. If it coincides with the original walk, then the latter is called *reversible*.

If the equivalence relation $(\mathcal{X},\mu,R)$ is endowed with a graph structure K, then the transition probabilities

$$p(x,y) = \begin{cases} 1/\operatorname{\mathsf{deg}} x\ , & (x,y)\in K \\ 0\ , & \text{otherwise}\ . \end{cases}$$

determine the *simple random walk* along the classes of the equivalence relation R.

COROLLARY 2.11. *Let $(\mathcal{X},\mu,R,K)$ be a graphed equivalence relation. If the measure μ is R-invariant, then the measure $\lambda = \mathsf{deg}\cdot\mu$ is stationary with respect to the simple random walk along the classes of R (i.e., $\lambda = \lambda P$, where P is the Markov operator of the simple random walk).*

REMARK 2.12. Actually, R-invariance of the measure μ is precisely equivalent to the combination of the above stationarity condition with *reversibility* of the leafwise simple random walk with respect to the measure $\mathsf{deg} \cdot \mu$, see [**Kai98**, Proposition 2.4.1 and its Corollary].

2.C. Entropy and leafwise Poisson boundaries. Below we shall be interested in describing the *Poisson boundary* of Markov chains along individual classes of an equivalence relation. We remind, without going into details, that the Poisson boundary is responsible for describing the stochastically significant behaviour of a Markov chain at infinity. The main tools used for its identification are the general *0–2 laws* and the *entropy theory* (see [**Kai92**] and the references therein). The latter one is especially expedient when dealing with random walks on groups [**KV83, Kai00**].

As it was on numerous occasions mentioned by the first author (e.g., see [**Kai86, Kai88, Kai90, Kai98**]), the entropy theory is also applicable in all situations when there is an appropriate probability path space endowed with a measure preserving time shift. The most general currently known setup is provided by *random walks on groupoids* [**Kai05**] with a finite stationary measure on the space of objects. A particular case of it consists of Markov chains along classes of an equivalence relation in the presence of a global stationary probability measure [**Kai98**], and, as it was pointed out in [**Kai98**], the entropy theory is perfectly applicable in this situation, providing criteria for triviality and identification of the Poisson boundary of leafwise random walks (also see [**Kai03**] where this theory was used for describing the Poisson boundary on the graphed equivalence relations associated with the fractal limit sets of iterated function systems). In other particular cases (most of which can actually be completely described in terms of random walks on equivalence relations) the entropy theory was along the same lines implemented in [**KW02**] (for Markov chains with a transitive group of symmetries), [**ACFdC11**] (for random walks along orbits of pseudogroups acting on a measure space), [**Bow10**] (for random walks on random Schreier graphs), [**BC10**] (for simple random walks on unimodular and stationary random graphs).

Since [**Kai98**] and [**Kai03**] contain only a brief outline of the entropy theory for random walks along equivalence relations, we shall give more details here (although these arguments are essentially the same as in the case of random walks in random environment [**Kai90**]). For the rest of this section we shall assume that $(\mathcal{X}, \mu, R)$ is a discrete measured equivalence relation endowed with a Markov operator P determined by a measurable family of transition probabilities π_x, and that $\lambda \prec \mu$ is a P-stationary probability measure. Denote by $\mathbf{P}_x$ the probability measure in the space of paths of the associated leafwise Markov chain issued from a point $x \in \mathcal{X}$. The one-dimensional distributions of $\mathbf{P}_x$ are the n-step transition probabilities π_x^n from the point x.

THEOREM 2.13. *For λ-a.e. point $x \in \mathcal{X}$ the tail and the Poisson boundaries of the leafwise Markov chain coincide $\mathbf{P}_x$ – mod 0.*

PROOF. Let

$$\varphi_n(x) = \|\delta_x P^n - \delta_x P^{n+1}\| = \|\pi_x^n - \pi_x^{n+1}\| , \qquad x \in X .$$

Then

$$\begin{aligned}\varphi_{n+1}(x) = \|\delta_x P^{n+1} - \delta_x P^{n+2}\| &= \|(\delta_x P^n - \delta_x P^{n+1})P\| \\ &\le \|\delta_x P^n - \delta_x P^{n+1}\| = \varphi_n(x) ,\end{aligned}$$

so that there exists a limit

$$\varphi(x) = \lim_n \varphi_n(x) .$$

Moreover,

$$\begin{aligned}\varphi_{n+1}(x) &= \|\delta_x P^{n+1} - \delta_x P^{n+2}\| = \|\delta_x P(P^n - P^{n+1})\| \\ &\le \sum_y p(x,y)\|\delta_y P^n - \delta_y P^{n+1}\| = \sum_y p(x,y)\varphi_n(y) ,\end{aligned}$$

whence φ is *subharmonic*:

$$\varphi \le P\varphi .$$

The function φ is clearly measurable. Then

$$\langle \lambda, P\varphi \rangle = \langle \lambda P, \varphi \rangle = \langle \lambda, \varphi \rangle$$

by stationarity of the measure λ, so that in fact φ is harmonic. Therefore, by a classical property of Markov chains with a finite stationary measure φ must be constant along a.e. sample path (e.g., see [**Kai92**]). By the corresponding 0–2 law (see again [**Kai92**]) in this situation φ can take values 0 and 2 only (obviously, in the ergodic case only one of these two options may occur). In the first case the Poisson and the tail boundary coincide for an arbitrary initial distribution, whereas in the second case for any $x \in \mathcal{X}$ the one-dimensional distributions π_x^n are all pairwise singular, so that the Poisson and the tail boundaries coincide $\mathbf{P}_x$ – mod 0. □

Let

$$H_n(x) = H(\pi_x^n)$$

be the *entropies* of n-step transition probabilities, and let

$$H_n = \int H_n(x)\, d\lambda(x) \tag{2.14}$$

be their averages over the space $(\mathcal{X}, \lambda)$. In terms of the shift-invariant measure

$$\mathbf{P}_\lambda = \int \mathbf{P}_x \, d\lambda(x)$$

on the space of sample paths $\boldsymbol{x} = (x_n) \in \mathcal{X}^{\mathbb{Z}_+}$ which corresponds to the stationary initial distribution λ,

$$H_n = -\int \log \pi_{x_0}^n(x_n)\, d\mathbf{P}_\lambda(\boldsymbol{x}) . \tag{2.15}$$

In yet another language, that of measurable partitions and their (conditional) entropies (e.g., see [**Roh67**]),

$$H_n = \mathbf{H}_\lambda(\alpha_n|\alpha_0) = \int \mathbf{H}_x(\alpha_n)\, d\lambda(x) , \tag{2.16}$$

where α_k denotes the k-th coordinate partition in the path space $\mathcal{X}^{\mathbb{Z}_+}$, and $\mathbf{H}_\lambda$ (resp., $\mathbf{H}_x$) denotes the (conditional) entropy with respect to the measure $\mathbf{P}_\lambda$ (resp., $\mathbf{P}_x$).

THEOREM 2.17. *If $H_1 < \infty$, then all the average entropies H_n are also finite, there exists a limit (the* asymptotic entropy*)*

$$\mathfrak{h} = \mathfrak{h}(P, \lambda) = \lim_n \frac{H_n}{n} < \infty \, , \tag{2.18}$$

and $\mathfrak{h} = 0$ if and only if for λ-a.e. point $x \in \mathcal{X}$ the Poisson boundary of the leafwise Markov chain is trivial with respect to the measure $\mathbf{P}_x$.

Let us put

$$\alpha_k^n = \bigvee_{i=k}^{n} \alpha_i \, , \qquad 0 \le k \le n \le \infty \, ,$$

where, as before, α_i are the coordinate partitions in the path space. For proving Theorem 2.17 we shall need the following

LEMMA 2.19. *For any $k \le n$ the conditional entropy of the partition α_1^k with respect to the partition α_n^∞ in the path space $\left(\mathcal{X}^{\mathbb{Z}_+}, \mathbf{P}_\lambda\right)$ is*

$$\mathbf{H}_\lambda \left(\alpha_1^k \,|\, \alpha_0 \vee \alpha_n^\infty\right) = \mathbf{H}_\lambda \left(\alpha_1^k \,|\, \alpha_0 \vee \alpha_n\right) = kH_1 + H_{n-k} - H_n \, . \tag{2.20}$$

PROOF. Formula (2.20) and its proof are completely analogous to the group case considered in [**KV83**]. Indeed, the leftmost identity in (2.20) immediately follows from the Markov property, whereas

$$\mathbf{H}_\lambda \left(\alpha_1^k \,|\, \alpha_0 \vee \alpha_n\right) = \int \mathbf{H}_x \left(\alpha_1^k \,|\, \alpha_n\right) \, d\lambda(x) \, ,$$

cf. formula (2.16). By the definition of conditional entropy, for any $x \in \mathcal{X}$

$$\mathbf{H}_x \left(\alpha_1^k \,|\, \alpha_n\right) = -\int \log \mathbf{P}_x \left(\alpha_1^k(\boldsymbol{x}) \,|\, \alpha_n(\boldsymbol{x})\right) \, d\mathbf{P}_x(\boldsymbol{x}) \, ,$$

where $\xi(\boldsymbol{x})$ denotes the element of a partition ξ which contains a sample path $\boldsymbol{x}$. Now,

$$\begin{aligned} \mathbf{P}_x \left(\alpha_1^k(\boldsymbol{x}) \,|\, \alpha_n(\boldsymbol{x})\right) &= \frac{\mathbf{P}_x \left(\alpha_1^k(\boldsymbol{x}) \cap \alpha_n(\boldsymbol{x})\right)}{\mathbf{P}_x \left(\alpha_n(\boldsymbol{x})\right)} \\ &= \frac{p(x_0, x_1) p(x_1, x_2) \cdots p(x_{k-1}, x_k) p^{n-k}(x_k, x_n)}{p^n(x_0, x_n)} \, , \end{aligned} \tag{2.21}$$

which implies the claim in view of formula (2.15) and shift invariance of the measure $\mathbf{P}_\lambda$. □

PROOF OF THEOREM 2.17. Formula (2.15) in combination with shift invariance of the measure $\mathbf{P}_\lambda$ easily implies subadditivity of the sequence H_n and existence of the limit (2.18). Moreover, the sequence of functions $\varphi_n(\boldsymbol{x}) = -\log \pi_{x_0}^n(x_n)$ satisfies conditions of Kingman's subadditive ergodic theorem, which implies existence of individual limits

$$\lim_n -\frac{1}{n} \log \pi_{x_0}^n(x_n)$$

for $\mathbf{P}_\lambda$-a.e. sample path $\boldsymbol{x} = (x_n)$ as well. If the shift T is ergodic, then these individual limits almost surely coincide with $\mathfrak{h}$. Note that ergodicity of T is equivalent to absence of non-trivial subsets of the state space $\mathcal{X}$ invariant with respect to the operator P (by aforementioned general property of Markov chains with a finite stationary measure), which, in the case when pairs of points $(x, y) \in R$ with $\pi_x(y) > 0$ generate the relation R, is equivalent to ergodicity of R.

Actually, Lemma 2.19 provides a stronger form of existence of the limit (2.18). Namely, since the sequence of partitions $\alpha_0 \vee \alpha_n^\infty$ is decreasing on n, monotonicity properties of conditional entropy (e.g., see [**Roh67**]) imply that $\mathbf{H}_\lambda\left(\alpha_1 \,|\, \alpha_0 \vee \alpha_n^\infty\right)$ increases on n. In view of formula (2.20) it means that not only the limit $\mathfrak{h} = \lim H_n/n$ exists, but also that $\left[H_{n+1} - H_n\right] \searrow \mathfrak{h}$.

As we have already seen on a similar occasion in the proof of Lemma 2.19, the left-hand side of formula (2.20) can be rewritten as

$$\mathbf{H}_\lambda\left(\alpha_1^k \,|\, \alpha_0 \vee \alpha_n^\infty\right) = \int \mathbf{H}_x\left(\alpha_1^k \,|\, \alpha_n^\infty\right)\, d\lambda(x)\ .$$

By continuity of conditional entropy (see again [**Roh67**]), for any $x \in \mathcal{X}$

$$\mathbf{H}_x\left(\alpha_1^k \,|\, \alpha_n^\infty\right) \nearrow \mathbf{H}_x\left(\alpha_1^k \,|\, \alpha^\infty\right) \le \mathbf{H}_x\left(\alpha_1^k\right)\ ,$$

where $\alpha^\infty = \lim_n \alpha_n^\infty$ is the *tail partition*. The right-hand side in the above formula is integrable, and

$$\int \mathbf{H}_x\left(\alpha_1^k\right) = kH_1\ ,$$

cf. formula (2.21). Therefore, after passing in (2.20) to a limit as $n \to \infty$ we conclude that for any $k > 0$

$$\int \mathbf{H}_x\left(\alpha_1^k \,|\, \alpha^\infty\right)\, d\lambda(x) = k(H_1 - \mathfrak{h})$$

and

$$k\mathfrak{h} = \int \left[\mathbf{H}_x(\alpha_1^k) - \mathbf{H}_x\left(\alpha_1^k \,|\, \alpha^\infty\right)\right] d\lambda(x)\ .$$

It means that $\mathfrak{h} = 0$ if and only if for λ-a.e. $x \in \mathcal{X}$ the tail partition α^∞ is $\mathbf{P}_x$-independent of all coordinate partitions α_1^k, the latter condition being equivalent to triviality of the tail partition $\mathbf{P}_x$ – mod 0.

Finally, by Theorem 2.13, for $\mathbf{P}_\lambda$-a.e. $x \in \mathcal{X}$ the tail and the Poisson boundaries coincide $\mathbf{P}_x$ – mod 0, which completes the proof. □

By passing to an appropriate *boundary extension* of the original equivalence relation [**Kai05**], Theorem 2.17 is also applicable to the problem of description of non-trivial Poisson boundaries of leafwise Markov chains. Indeed, a quotient of the Poisson boundary is *maximal* (i.e., coincides with the whole Poisson boundary) if and only if for almost all conditional chains determined by the points of this quotient the Poisson boundary is trivial. Thus, the criterion from Theorem 2.17 allows one to carry over the *ray* and the *strip* criteria used for identification of the Poisson boundary in the group case [**Kai00**] to the setup of random walks along classes of graphed equivalence relations.

2.D. The Poisson boundary of random horospheric products. Below by a *random horospheric product* we shall mean a graphed equivalence relation $(\mathcal{X}, \mu, R, K)$ such that a.e. leafwise graph is a horospheric product. Moreover, we shall assume that the "orientations" (signs) of leafwise height cocycles (1.3) are chosen in a consistent way, i.e., that there exists a global $\mathbb{Z}$-valued *measurable* cocycle $\mathcal{B}$ on R such that its restriction to a.e. leaf is a height cocycle. Therefore, a.e. leafwise graph $[x]^K$, being a horospheric product, is endowed with its lower and upper boundaries $\partial_\downarrow[x]^K$ and $\partial_\uparrow[x]^K$, respectively, and one can easily see that the corresponding boundary bundles over $(\mathcal{X}, \mu, R, K)$ are measurable (cf. [**Kai04**]).

THEOREM 2.22. *Let $(\mathcal{X}, \mu, R, K, \mathcal{B})$ be a random horospheric product with uniformly bounded vertex degrees, and P — the Markov operator of a random walk along classes of the equivalence relation R determined by a measurable family of leafwise transition probabilities $\{\pi_x\}_{x\in\mathcal{X}}$. If $\lambda \prec \mu$ is a P-stationary probability measure such that the transition probabilities $\{\pi_x\}$ have a finite first moment*

$$\int_R d(x,y)\, d\Pi(x,y) \ , \tag{2.23}$$

where d is the leafwise graph distance, and Π is the measure (2.4)*, then the Poisson boundaries of leafwise random walks are determined by the* global height drift

$$h = \int_R \mathcal{B}(x,y)\, d\Pi(x,y) \ .$$

If $h = 0$, then the Poisson boundary is a.s. trivial, whereas when $h > 0$ (resp., $h < 0$) a.e. leafwise Poisson boundary coincides (mod 0) with the upper (resp., lower) leafwise boundary endowed with the corresponding limit distribution (which is well-defined by Proposition 1.13*).*

PROOF. Theorem 1.11 in combination with the standard ergodic arguments (cf. [**Kai00**]) implies that a.e. sample path is regular with the height drift h. If $h = 0$, then regularity implies vanishing of the asymptotic entropy, and therefore triviality of leafwise Poisson boundaries. If $h \neq 0$, then by Proposition 1.13 a.e. sample path converges to the corresponding boundary (the upper, if $h > 0$, and the lower, if $h < 0$) of leafwise horospheric products. The fact that these boundaries are actually maximal (i.e., coincide with the leafwise Poisson boundaries) then follows from the ray criterion (or Corollary 1.15 in combination with the strip criterion) in precisely the same way as in the group case, cf. [**Kai00**]. $\square$

REMARK 2.24. Finiteness of the entropies (2.14) (which is crucial for Theorem 2.17) follows, in the usual way, from finiteness of the first moment (2.23) and uniform boundedness of vertex degrees, cf. [**Der86**, p. 259] or [**Kai00**, Lemma 5.2].

Obviously, if the operator P is reversible with respect to a stationary measure λ (see the discussion at the end of Section **2.B** for the definition), then the integral of any additive cocycle on R with respect to Π vanishes. In particular, in this case the global height drift h vanishes, whence

COROLLARY 2.25. *Under conditions of Theorem* 2.22*, if the operator P is reversible with respect to the measure λ, then the leafwise Poisson boundaries are a.s. trivial.*

COROLLARY 2.26. *Under conditions of Theorem* 2.22*, if $\lambda = \mathsf{deg} \cdot \mu$ is the stationary measure of the leafwise simple random walk corresponding to a finite R-invariant measure μ (see Corollary* 2.11*), then the Poisson boundary of the leafwise simple random walks is a.s. trivial.*

References

[ACFdC11] Fernando Alcalde Cuesta and María P. Fernández de Córdoba, *Nombre de branchement d'un pseudogroupe*, Monatsh. Math. **163** (2011), 389–414. MR2820370

[Ada90] Scot Adams, *Trees and amenable equivalence relations*, Ergodic Theory Dynam. Systems **10** (1990), no. 1, 1–14. MR91d:28041

[BC10] Itai Benjamini and Nicolas Curien, *Ergodic theory on stationary random graphs*, arXiv:1011.2526v1, 2010.

[Ber01] D. Bertacchi, *Random walks on Diestel-Leader graphs*, Abh. Math. Sem. Univ. Hamburg **71** (2001), 205–224. MR1873044 (2002i:60095)

[BNW08] Laurent Bartholdi, Markus Neuhauser, and Wolfgang Woess, *Horocyclic products of trees*, J. Eur. Math. Soc. (JEMS) **10** (2008), no. 3, 771–816. MR2421161

[Bow10] Lewis Bowen, *Random walks on coset spaces with applications to Furstenberg entropy*, arXiv:1008.4933v1, 2010. MR2653969 (2011h:37007)

[BSCSW11] Alexander Bendikov, Laurent Saloff-Coste, Maura Salvatori, and Wolfgang Woess, *The heat semigroup and Brownian motion on strip complexes*, Adv. Math. **226** (2011), no. 1, 992–1055. MR2735780

[BW05a] Laurent Bartholdi and Wolfgang Woess, *Spectral computations on lamplighter groups and Diestel-Leader graphs*, J. Fourier Anal. Appl. **11** (2005), no. 2, 175–202. MR2131635 (2006e:20052)

[BW05b] Sara Brofferio and Wolfgang Woess, *Green kernel estimates and the full Martin boundary for random walks on lamplighter groups and Diestel-Leader graphs*, Ann. Inst. H. Poincaré Probab. Statist. **41** (2005), no. 6, 1101–1123. MR2172211 (2007a:60045)

[BW06] ______, *Positive harmonic functions for semi-isotropic random walks on trees, lamplighter groups, and DL-graphs*, Potential Anal. **24** (2006), no. 3, 245–265. MR2217953 (2007b:31010)

[Car72] P. Cartier, *Fonctions harmoniques sur un arbre*, Symposia Mathematica, Vol. IX (Convegno di Calcolo delle Probabilità, INDAM, Rome, 1971), Academic Press, London, 1972, pp. 203–270. MR0353467 (50 #5950)

[CKW94] D. I. Cartwright, V. A. Kaimanovich, and W. Woess, *Random walks on the affine group of local fields and of homogeneous trees*, Ann. Inst. Fourier (Grenoble) **44** (1994), no. 4, 1243–1288. MR1306556 (96f:60121)

[Der86] Y. Derriennic, *Entropie, théorèmes limite et marches aléatoires*, Probability measures on groups, VIII (Oberwolfach, 1985), Lecture Notes in Math., vol. 1210, Springer, Berlin, 1986, pp. 241–284. MR879010 (89f:60076b)

[DL01] Reinhard Diestel and Imre Leader, *A conjecture concerning a limit of non-Cayley graphs*, J. Algebraic Combin. **14** (2001), no. 1, 17–25. MR1856226 (2002h:05082)

[EFW07] Alex Eskin, David Fisher, and Kevin Whyte, *Quasi-isometries and rigidity of solvable groups*, Pure Appl. Math. Q. **3** (2007), no. 4, part 1, 927–947. MR2402598 (2009b:20074)

[FM77] Jacob Feldman and Calvin C. Moore, *Ergodic equivalence relations, cohomology, and von Neumann algebras. I*, Trans. Amer. Math. Soc. **234** (1977), no. 2, 289–324. MR58:28261a

[Kai86] Vadim A. Kaimanovich, *Brownian motion and harmonic functions on covering manifolds. An entropic approach*, Soviet Math. Dokl. **33** (1986), no. 3, 812–816, English tranlation. MR88k:58163

[Kai88] ______, *Brownian motion on foliations: entropy, invariant measures, mixing*, Funktsional. Anal. i Prilozhen. **22** (1988), no. 4, 82–83. MR977003 (91b:58124)

[Kai89] ______, *Lyapunov exponents, symmetric spaces and a multiplicative ergodic theorem for semisimple Lie groups*, J. Soviet Math. **47** (1989), 2387–2398. MR89m:22006

[Kai90] ______, *Boundary and entropy of random walks in random environment*, Probability theory and mathematical statistics, Vol. I (Vilnius, 1989), "Mokslas", Vilnius, 1990, pp. 573–579. MR1153846 (93k:60174)

[Kai91] ______, *Poisson boundaries of random walks on discrete solvable groups*, Probability measures on groups, X (Oberwolfach, 1990), Plenum, New York, 1991, pp. 205–238. MR1178986 (94m:60014)

[Kai92] ______, *Measure-theoretic boundaries of Markov chains,* 0-2 *laws and entropy*, Harmonic analysis and discrete potential theory (Frascati, 1991), Plenum, New York, 1992, pp. 145–180. MR94h:60099

[Kai97] ______, *Amenability, hyperfiniteness, and isoperimetric inequalities*, C. R. Acad. Sci. Paris Sér. I Math. **325** (1997), no. 9, 999–1004. MR98j:28014

[Kai98] ______, *Hausdorff dimension of the harmonic measure on trees*, Ergodic Theory Dynam. Systems **18** (1998), no. 3, 631–660. MR99g:60123

[Kai00] ______, *The Poisson formula for groups with hyperbolic properties*, Ann. of Math. (2) **152** (2000), no. 3, 659–692. MR1815698 (2002d:60064)

[Kai03] ______, *Random walks on Sierpiński graphs: hyperbolicity and stochastic homogenization*, Fractals in Graz 2001, Trends Math., Birkhäuser, Basel, 2003, pp. 145–183. MR2091703 (2005h:28022)

[Kai04] ______, *Boundary amenability of hyperbolic spaces*, Discrete geometric analysis, Contemp. Math., vol. 347, Amer. Math. Soc., Providence, RI, 2004, pp. 83–111. MR2077032 (2005j:20051)

[Kai05] ______, *Amenability and the Liouville property*, Israel J. Math. **149** (2005), 45–85, Probability in mathematics. MR2191210 (2007c:43001)

[KS10] Vadim A. Kaimanovich and Florian Sobieczky, *Stochastic homogenization of horospheric tree products*, Probabilistic approach to geometry, Adv. Stud. Pure Math., vol. 57, Math. Soc. Japan, Tokyo, 2010, pp. 199–229. MR2648261 (2011e:37013)

[KV83] V. A. Kaimanovich and A. M. Vershik, *Random walks on discrete groups: boundary and entropy*, Ann. Probab. **11** (1983), no. 3, 457–490. MR85d:60024

[KW02] Vadim A. Kaimanovich and Wolfgang Woess, *Boundary and entropy of space homogeneous Markov chains*, Ann. Probab. **30** (2002), no. 1, 323–363. MR2003d:60152

[LPP95] Russell Lyons, Robin Pemantle, and Yuval Peres, *Ergodic theory on Galton-Watson trees: speed of random walk and dimension of harmonic measure*, Ergodic Theory Dynam. Systems **15** (1995), no. 3, 593–619. MR1336708 (96e:60125)

[LPP96a] ______, *Biased random walks on Galton-Watson trees*, Probab. Theory Related Fields **106** (1996), no. 2, 249–264. MR1410689 (97h:60094)

[LPP96b] ______, *Random walks on the lamplighter group*, Ann. Probab. **24** (1996), no. 4, 1993–2006. MR1415237 (97j:60014)

[Roh67] V. A. Rohlin, *Lectures on the entropy theory of transformations with invariant measure*, Uspehi Mat. Nauk **22** (1967), no. 5 (137), 3–56. MR0217258 (36 #349)

[Woe91] Wolfgang Woess, *Topological groups and infinite graphs*, Discrete Math. **95** (1991), no. 1-3, 373–384, Directions in infinite graph theory and combinatorics (Cambridge, 1989). MR1141949 (93i:22004)

[Woe05] ______, *Lamplighters, Diestel-Leader graphs, random walks, and harmonic functions*, Combin. Probab. Comput. **14** (2005), no. 3, 415–433. MR2138121 (2006d:60021)

Department of Mathematics and Statistics, University of Ottawa, 585 King Edward, Ottawa ON, K1N 6N5, Canada

E-mail address: `vkaimano@uottawa.ca, vadim.kaimanovich@gmail.com`

Department of Mathematics, University of Colorado at Boulder, Campus Box 395

E-mail address: `florian.sobieczky@colorado.edu`

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11229

Statistics of gaps in the sequence $\{\sqrt{n}\}$

Ya. G. Sinai

Dedicated to A. M. Stepin on the occasion of his 70th birthday

Abstract. The problem of statistics of spacings in the sequence $\{\sqrt{n}\}$ was solved completely by N. Elkies and C. McMullen. In this paper I propose a different approach to the problem which is based on our previous result with C. Ulcigrai.

Dedication

1. Formulation of the Problem

Let $0 < \alpha < 1$ and $1 \leq n \leq N$ where n, N are integers. Consider the set $\Omega_\alpha(N) = (\{n^\alpha\}, \quad 1 \leq n \leq N) \subset [0,1]$. Since $|\Omega_\alpha(N)| = N$ it is natural to expect that typical distances between neighboring points in $\Omega_\alpha(N)$ are $O(\frac{1}{N})$. M. Boshernitzan was the first to propose the study of statistical properties of the normalized distances between neighbors and their limit as $N \to \infty$ (see [**Bo**]). His numerical results showed that for $\alpha \neq \frac{1}{2}$ the corresponding distributions were close to exponential while for $\alpha = \frac{1}{2}$ it was strikingly different.

The case $\alpha = \frac{1}{2}$ was solved completely by N. Elkies and C. McMullen (see [EM]). Their approach was based upon the observation that the problem could be reduced to some questions about 2–dimensional lattices and the subsequent integration in the space of lattices. It gave also an explicit expression for the limiting distribution. This distribution didn't appear before in problems of probability theory. However, it is likely that it is an infinite–divisible distribution.

An additional interest comes from the fact that for each n it is difficult to find the neighboring point of $\{n^\alpha\}$. In this sense the whole problem is close to problems of quantum chaos in which the order of eigen-values of quantum systems has nothing common with their analytic expressions.

In this paper we propose another approach to the main problem considered by [**EM**] by which we show the existence of the limiting distribution but don't give its explicit form. However, many properties of this distribution can be studied. There is some hope that our approach can be used in the case of other α. We recommend to the reader a very interesting paper [A,Ma,Mo] on the connections of number theory and ergodic theory.

The author thanks NSF for the financial support, grant DMS 0901235.

2010 *Mathematics Subject Classification.* Primary 37A45.
Research supported by NSF grant DMS-0901235.

Useful discussion with M. Boshernitzan and I. Vinogradov are highly appreciated.

2. Preliminary construction

As in [**EM**], it is more convenient to consider $1 \le n \le N^2$. In this case $1 \le \sqrt{n} \le N$ and it is natural to expect that the distances between neighbors are $O\left(\frac{1}{N^2}\right)$. We shall study the distance from each point to its right neighbor.

Let $k \ge 1$ be an integer and $0 \le x \le 1$. Set

$$
\begin{aligned}
n_{k,x} &= \left[(k+x)^2\right] = (k+x)^2 - \left\{(k+x)^2\right\} \\
&= (k+x)^2 \left(1 - \frac{\left\{2kx + x^2\right\}}{(k+x)^2}\right) \qquad (1)
\end{aligned}
$$

In what follows, N is assumed to be a fixed large integer. Let

$$
\begin{aligned}
&t_q = t_{q,N} = \frac{q}{2N}\ , \qquad \text{for integers } q \ge 0\ , \\
&K = K_N = \{1, 2, \dots,\ N-1\}, \\
&X = X_N = \{t_p | 0 \le p \le 2N-1\} \qquad\qquad \text{and} \\
&Y = Y_N = X + K = \left\{t_q | 2N \le q \le 2N^2 - 1\right\}.
\end{aligned}
$$

We routinely assume that $x \in X,\ k \in K$ and $y \in Y$. Clearly, given $y = k + x$, then k, x are uniquely determined by $y:\ k = [y]$ and $x = \{y\}$.

Let us show that each integer $n,\ 1 \le n < N^2$, has a (not necessarily unique) representation in the form

$$
n = n_{k,x} = \left[(k+x)^2\right], \qquad (2.a)
$$

with some $k \in K$ and $x \in X$. In view of definition of Y, it suffices to show that

$$
n = \left[(t_q)^2\right] \qquad (2.b)
$$

for some $q,\ 2N \le q \le 2N^2 - 1$. (Both presentations of n, (2.a) and (2.b), do not need to be unique).

The existence of such q follows immediately from the relations

$$
\left[(t_{2N})^2\right] = 1. \quad \left[(t_{2N^2-1})^2\right] = N^2 - 1,
$$

and from the inequalities

$$
0 < t_q^2 - t_{q-1}^2 = \frac{2q-1}{4N^2} \le \frac{2 \cdot 2N^2 - 3}{4N^2} < 1, \qquad \text{for } q \le 2N^2 - 1.
$$

One derives from (1) that

$$
\begin{aligned}
\sqrt{n_{k,x}} &= (k+x)\left(1 - \frac{\left\{2kx + x^2\right\}}{2(k+x)^2} + \delta_1\right) \\
&= k + x - \frac{\left\{2kx + x^2\right\}}{2(k+x)} + \delta_2.
\end{aligned}
$$

Here and below δ's with indices mean values which are sufficiently small as k increases (in the sense to be made later). We write

$$
\left\{\sqrt{n_{k,x}}\right\} = x - \frac{2kx + x^2}{2(k+x)} + \delta_2 \qquad (3)
$$

The last formula shows that the points in the rhs of (3) corresponding to some x and different k are to the left from x and are within the distance $O\left(\frac{1}{k}\right)$ from x.

3. The analysis of the differences $\sqrt{n_{k_1,x}} - \sqrt{n_{k_2,x}}$

Denote $y = k + x$. For any pair $y_1 = k_1 + x_1,\ y_2 = k_2 + x_2$ from (3).

$$\begin{aligned}\left\{\sqrt{n_{k_1,x_1}}\right\} - \left\{\sqrt{n_{k_2,x_2}}\right\} &= x_1 - x_2 + \frac{\left\{2k_2x_2 + x_2^2\right\}}{2y_2}\\ &- \frac{\left\{2k_1x_1 + x_1^2\right\}}{2y_1} + \delta_3\end{aligned} \tag{4}$$

Consider the simple case $x_1 = x_2 = x$. Then the rhs of (4) equals to

$$\begin{aligned}&\frac{1}{2y_1}\left(-\left\{2k_1x + x^2\right\} + \frac{y_1}{y_2}\left\{2k_2x + x^2\right\}\right) + \delta_4\\ &= \frac{1}{2y_1}\left(\left(\frac{k_1 - k_2}{k_2 + x_2} + 1\right)\left\{2k_2x + x^2\right\} - \left\{2k_1x + x^2\right\}\right)\\ &+ \delta_4 =\\ &= \frac{1}{2y_1}\left(\left\{2k_2x + x^2\right\} - \left\{2k_1x + x^2\right\}\right) + \frac{k_1 - k_2}{2y_1(k_2 + x)}\cdot\\ &\cdot\left\{2k_2x + x^2\right\}\Bigg) + \delta_4\end{aligned} \tag{5}$$

Observe that $\left\{2k_2x + x^2\right\} - \left\{2k_1x + x^2\right\} = \{2(k_2 - k_1)x\}$, assuming that $\left\{2k_2x + x^2\right\}$ is the right neighbor of $\left\{2k_1x + x^2\right\}$.

Now (5) can be written as

$$\frac{1}{2y}(\{2x(k_2 - k_1)\} - \frac{k_2 - k_1}{k_2 + x}\left\{2k_2x + x^2\right\} \tag{6}$$

For our proof it is necessary to find $k_2 - k_1$ such that

$$\{2x\,(k_2 - k_1)\} - \frac{k_2 - k_1}{k_2 + x}\left\{2k_2x + x^2\right\} = O\left(\frac{1}{N}\right)$$

In this case

$$\begin{aligned}&\{2x(k_2 - k_1)\} - \frac{k_2 - k_1}{k_2 + x}\left\{2k_2x + x^2\right\} =\\ &= \left\{\left(2x - \frac{\left\{2k_2x + x^2\right\}}{k_2 + x}\right)(k_2 - k_1)\right\} = O\left(\frac{1}{N}\right).\end{aligned}$$

We recall the theorem from [SU] which gives the existence of the limiting distribution for $\frac{k_2 - k_1}{N}$ which implies the existence of the needed limiting distribution in our case when $x_1 = x_2$.

4. The case of arbitrary x_1, x_2.

For any pair $y_1 = k_1 + x_1,\ \ y_2 = k_2 + x_2$ from (3)

$$\begin{aligned}\left\{\sqrt{n_{k_1,x_1}}\right\} - \left\{\sqrt{n_{k_2,x_2}}\right\} &= x_1 - x_2\\ &+ \frac{\left\{2k_2x_2 + x_2^2\right\}}{2y_2} - \frac{\left\{2k_1x_1 + x_1^2\right\}}{2y_1}\end{aligned} \tag{7}$$

and the rhs of (7) equals to

$$\begin{aligned}
&\frac{1}{2y_1}\left(2y_1(x_1-x_2)-\left\{2k_1x_1+x_1^2\right\}+\frac{y_1}{y_2}\left\{2k_2x_2+x_2^2\right\}\right)\\
&+\delta_5=\\
&=\frac{1}{2y_1}(2k_1(x_1-x_2)-\left\{2k_1x_1+x_1^2\right\}+\frac{k_1}{k_2+x_2}\left\{2k_2x_2+x_2^2\right\}\\
&+2x_1(x_1-x_2)+\frac{x_1}{k_2+x_2}\left\{2k_2x_2+x_2^2\right\})+\delta_5 \qquad (8)
\end{aligned}$$

Consider x_1, x_2, k_2 as three independent parameters and assume that

$$c_1\ N \le k_1,\ k_2 \le c_2\ N \qquad (9)$$

where c_1, c_2 are some positive constants. Then

$$\frac{\left\{2k_2x_2+x_2^2\right\}}{2y_2} \le \frac{1}{2k_2} = \frac{1}{2c_1\ N}$$

The inequalities (9) can be used because in the case of weak convergence of distributions one can restrict himself by the values of k_1, k_2 for which (9) holds for arbitrary c_1, c_2. Also (3) shows that for all x

$$x-\frac{1}{2c_1\ N} \le \left\{\sqrt{n_{k_1,\ x}}\right\} \le x$$

If we want to consider pairs y_1, y_2 for which

$$\left\{\sqrt{n_{k_1,x_1}}\right\}-\left\{\sqrt{n_{k_2,\ x_2}}\right\}=O\left(\frac{1}{N^2}\right)$$

we may assume that in $x_1-x_2=\frac{p_1}{2N}-\frac{p_2}{2N}$ the difference p_1-p_2 satisfies the inequalities

$$0 \le p_1-p_1 \le \frac{1}{2c_1}.$$

Return back to (5). We can write

$$\begin{aligned}
&2y_1\left(\left\{\sqrt{n_{k_1,\ x_1}}\right\}-\left\{\sqrt{n_{k_2,\ x_2}}\right\}\right)\\
&=(2k_1(x_1-x_2)+\frac{k_1}{k_2+x_2}\left\{2k_2x_2+x_2^2\right\}\\
&+2x_1(x_1-x_2)+\frac{x_1}{k_2+x_2}\left\{2k_2x_2+x_2^2\right\}\\
&-\left\{2k_1x_1+x_1^2\right\}+\delta_5)\\
&=(2(x_1-x_2)(1+2x_1)\\
&+\frac{k_1+x_1}{k_2+x_2}\left\{2k_2x_2+x_2^2\right\}-\left\{2k_1x_1+x_1^2\right\}+\delta_5)\\
&=(2(x_1-x_2)(1+2x_1)+\left\{2k_2x_2+x_2^2\right\}\\
&-\left\{2k_1x_1+x_1^2\right\}+\frac{k_1-k_2}{k_2+x_2}\left\{2k_2x_2+x_2^2\right\}+\delta_5) \qquad (10)
\end{aligned}$$

As before,

$$\begin{aligned}&\left\{2k_2x_2+x_2^2\right\}-\left\{2k_1x_1+x_1^2\right\}\\&=\left\{2x_2\left(k_2-k_1\right)+2k_1x_1+x_1^2+2k_1\left(x_2-x_1\right)+\left(x_2-x_1\right)\left(x_2+x_1\right)\right\}\\&-\left\{2k_1x_1+x_1^2\right\}.\end{aligned}$$

If we introduce the rotation R_1 of the unit circle by the angle $2x_2$ then the last expression can be written as

$$\begin{aligned}&R_1^{(k_2-k_1)}(\left\{2k_1x_1+x_1^2+2k_1(x_2-x_1)+(x_2-x_1)(x_2+x_1)\right\}\\&-\left\{2k_1x_1+x_1^2\right\}\\&=R_1^{(k_2-k_1)}(2k_1(x_2-x_1)+(x_2+x_1)(x_2+x_1))\end{aligned}$$

Returning back to (10) we introduce the rotation by the angle

$$2x_1-\frac{\left\{2k_2x_2+x_2^2\right\}}{k_2+x_2}.$$

If x_1, x_2, k_1 are independent parameters then $2k_1(x_2-x_1)+(x_2-x_1)(x_2+x_1)$ has a limiting distribution and again the theorem from [**SU**] gives the result.

5. The remainders

It is easy to see that $|\delta_1|\le\frac{count}{|k|^3}$, $|\delta_2|\le\frac{count}{k^2}$, $|\delta_3|\le\frac{count}{k^2}|\delta_4$, $|\le\frac{count}{k^2}$, $|\delta_5|\le\frac{count}{k^2}$. These estimates are enough for our purposes.

References

[A,Ma,Mo] A. Arbieto, C. Matheus, C. Moreira, A remarkable effectiveness of ergodic theory in number theory, *Ensaios Matemáticos*, 2009, vol. 17, 1–104. MR2582821 (2011a:11015)

[Bo] M. D. Boshernitzan, Uniform Distribution and Hardy Fields, *J. d'Analyse Math.* **62**(1994), 240-250. MR1269206 (95e:11085)

[EM] N. D. Elkies, C. McMullen, Gaps in $\sqrt{n}$ mod 1 and erzodic theory, *Duke Math. J.*, **123**(2004), N1, 95-139. MR2060024 (2005f:11143)

[SU] Ya. G. Sinai and C. Ulcigrai, Renewal-type limit theorem for the Gauss map and continued fractions. *Ergodic Theory and Dynamical Systems*, **28**(2008), N2, 643-655. MR2408397 (2009i:37019)

Mathematics Department of Princeton University, Princeton, New Jersey 08544, and Landau Institute of Theoretical Physics, Russian Academy of Sciences, Moscow, Russia.

E-mail address: sinai@math.princeton.edu

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11254

Invariant distributions for interval exchange transformations

William A. Veech

Dedicated to Anatoly Stepin, on the Occasion of his Seventieth Birthday

Abstract. Our purpose is to point out that the coboundary equation (8.7) of [Veech, *Amer. J. Math.* **106** (1984), pp. 1331–1359] leads naturally to spaces of invariant distributions for interval exchange maps.We have been motivated by A. Bufetov's insightful characterization of fibers relative to Forni's Lyapunov splitting of the Forni cocycle ([Forni], [Bufetov, April 2008], [Bufetov, February 2009], [Bufetov, *Electron Res. Announc. Math. Sci.* **17** (2010)], [Bufetov, March 2010]). For a generic point in a stratum Bufetov's characterization is in terms of additive cocycles for the vertical foliation that are Hölder on vertical leaves and holonomy invariant with respect to the horizontal foliation. Such a cocycle induces on a vertical separatrix a distribution that is invariant under the interval exchange maps determined by horizontal returns to finite subsegments.

1. Introduction

Let $T = T_{(\lambda,\pi)}$, $\lambda \in \Lambda_m \overset{def}{=} (\mathbb{R}^+)^m, \pi \in \mathfrak{S}_m$, be an interval exchange on $I^\lambda = [0, |\lambda|)$. Assume throughout that T satisfies Keane's i.d.o.c. ([**Ke75**]). Use I_j^λ to denote basic intervals for T, i.e.,

$$I_j^\lambda = [\beta_{j-1}, \beta_j),\ \beta_k = \beta_k(\lambda) = \sum_{1\le i\le k} \lambda_i,\ 0 \le k \le m. \tag{1.1}$$

We recall that $T_{(\lambda,\pi)}$ may be defined in terms of an alternating matrix L^π:

$$\begin{aligned} Tx \ &= \ T_{(\lambda,\pi)}x = x + (L^\pi\lambda)_i\,, x \in I_i^\lambda \\ L_{ij}^\pi &= \begin{cases} 1, \pi j < \pi i \\ 0, \pi j \ge \pi i \end{cases} = -L_{ji}^\pi, i < j. \end{aligned} \tag{1.2}$$

([**R79**], [**V78**]).

Introduce a space $\mathcal{Y}_{|\lambda|}$ of functions on $[0, |\lambda|)$ as

$$\mathcal{Y}_{|\lambda|} = \left\{F \in BV([0, |\lambda|)) \mid F(x^+) = F(x), 0 \le x < |\lambda|\right\}. \tag{1.3}$$

(Notice that the signed measure dF has no point mass at 0. Notice also that if $F \in \mathcal{Y}_{|\lambda|}$, $F(|\lambda|^-) \overset{def}{=} \lim_{x\nearrow|\lambda|} F(x)$ exists and is finite.) Since T is piecewise isometric

---

2010 *Mathematics Subject Classification.* Primary 32G15, 30F60, 37D40, 37F30.

and right continuous, $\mathcal{T}F \stackrel{def}{=} F \circ T$ satisfies

$$\mathcal{T}\mathcal{Y}_{|\lambda|} = \mathcal{Y}_{|\lambda|}.$$

$\mathcal{T}$ is continuous (but not isometric) relative to the norm

$$\|F\|_{\mathcal{Y}_{|\lambda|}} = \|F\|_{\infty} + \|F\|_{BV}\,, \tag{1.4}$$

and therefore it makes sense to speak of invariant elements of the dual $\mathcal{Y}^{*}_{|\lambda|}$, elements such that

$$\Lambda\left(\mathcal{T}F\right) = \Lambda\left(F\right),\ F \in \mathcal{Y}_{|\lambda|}. \tag{1.5}$$

Define the weak derivative of $\varphi \in \mathbf{C}_0\left([0,|\lambda|]\right)$ as a linear functional, Λ_φ, on $\mathcal{Y}_\lambda$:

$$\Lambda_\varphi\left(F\right) \stackrel{def}{=} -\int\limits_{[0,|\lambda|)} \varphi dF\ + \varphi\left(|\lambda|\right) F(|\lambda|^{-}),\ F \in \mathcal{Y}_\lambda, \varphi \in \mathbf{C}_0\left([0,|\lambda|]\right). \tag{1.6}$$

Let $\mathbf{C}_0\left([0,|\lambda|]\right)$ be the Banach space of continuous functions φ on $[0,|\lambda|]$ such that $\varphi(0) = 0$. Define $\mathcal{K}(T) \subset \mathbf{C}_0\left([0,|\lambda|]\right)$ by

$$\mathcal{K}(T) = \left\{\varphi \in \mathbf{C}_0\left([0,|\lambda|]\right) \mid \Lambda_\varphi \circ \mathcal{T} = \Lambda_\varphi\right\}. \tag{1.7}$$

We shall prove

Theorem 1.1. *Let $T = T_{(\lambda,\pi)}, \lambda \in \Lambda_m \stackrel{def}{=} \left(\mathbb{R}^{+}\right)^m, \pi \in \mathfrak{S}_m$, be an interval exchange on $I^\lambda = [0,|\lambda|)$, and assume that T satisfies Keane's i.d.o.c. The dimension of $\mathcal{K}(T)$ satisfies the inequality*

$$Dim\left(\mathcal{K}(T)\right) \leq \frac{1}{2} Rank\left(L^\pi\right). \tag{1.8}$$

Remark 1.2. The inequality corresponding to (1.8) for finite invariant measures is due to Katok [**Ka73**], in the context of flows on surfaces, and to the author [**V78**], in the context of interval exchange transformations. It is a consequence of Forni's Theorem, [**F02**], that for Lebesgue-almost all λ equality holds in (1.8). According to [**M82**], [**V82**], it is also true for Lebesgue-almost all λ that $T_{(\lambda,\pi)}$ is uniquely ergodic, implying that a codimension one subspace of $\mathcal{K}(T)$ consists of functions $\varphi \notin \mathcal{Y}_{|\lambda|}$. The related issue of generic weak mixing was treated first, for 3-interval exchanges, in the famous paper of Katok and Stepin, [**KS67**]. That their result is best possible is a consequence of the existence of solutions to a coboundary equation for $\mathbb{Z}_2-$skew products ([**V69**]). See also [**Sa75**] and [**S81**]. A rather complete theory of 3-interval exchanges is given in a series of papers by Ferenczi, Holton and Zamboni, of which one, [**FHZ04**], focuses elegantly on spectral properties. The Katok-Stepin genericity result was extended to an infinite collection of permutations in [**V84-I**] and, finally, to all irreducible (non-rotation) permutations in [**AF07**].

We close this introduction with an indication of the organization of the paper. Sections 2-4 are devoted to the proof of Theorem 1.1. The proof turns on the observation that $\mathcal{K}(T)$ coincides with the set of $\varphi \in \mathbf{C}_0\left([0,|\lambda|]\right)$ that satisfy the coboundary equation, (8.7) of [**V84-I**],

$$\varphi \circ T - \varphi = \sum_{i=1}^{m} z_i \chi_{I_i^\lambda} \tag{1.9}$$

for some $z = z^{\varphi} \in \mathbb{R}^m$. The connection between $\mathcal{K}(T)$ and the coboundary equation will be made in terms of an intermediate space of T-invariant additive set functions on the algebra, $\mathcal{F}_{|\lambda|}$, of finite unions of left closed-right open intervals in $[0, |\lambda|)$. The set functions have the form $\nu^{\varphi}([a, b)) = \varphi(b) - \varphi(a)$, $\varphi \in \mathbf{C}_0([0, |\lambda|])$. In this respect the upper bound (1.8) is more directly related to the classical upper bounds in [**Ka73**] and [**V78**]. As indicated in Remark 1.2, the upper bound is achieved for a.e. λ.

In Section 5 we recall properties of the zippered rectangles flow from [**V82**], [**V84-II, V84-III**], and describe how Lyapunov regularity of the Rauzy cocycle follows from a corresponding statement for the zippered rectangles flow. Since the invariant measure for Rauzy induction is infinite, the multiplicative ergodic theorem is not directly applicable to the Rauzy cocycle. The "speeded-up" version of Rauzy induction introduced by Zorich ([**Z96**]) does admit a finite invariant measure, and the multiplicative ergodic theorem does apply directly to Zorich's cocycle.

Section 6 discusses the multiplicative analog of (1.9) that was introduced in [**V84-I**] in the context of the problem of genericity of weak mixing for non-rotation interval exchange transformations. The approach was successful for an infinite set of irreducible permutations. A much deeper study of this cocycle was made by Avila and Forni, who gave a complete solution to the genericity problem ([**AF07**]).

As we have indicated, the present study has been motivated by work of Bufetov cited in the abstract. In what follows we shall indicate in more detail the connection between Bufetov's work and the works by Kontsevich-Zorich ([**Ko97**]), Forni ([**F02**]) and Avila-Viana ([**AV07**]) that have motivated it. The description in the next paragraph of the Kontsevich-Zorich cocycle is taken directly from [**V08**]. Sections 7-8 are devoted to the description of a dictionary between the Kontsevich-Zorich cocycle, the Rauzy cocycle and the zippered rectangles flow.

Begin with the unit cotangent bundle, Q_p^1, of Teichmüller space and the trivial bundle

$$Q_p^1 \times H^1(M_p, \mathbb{R}).$$

Endow each fiber with the Hodge inner product and norm. The geodesic flow on Q_p^1, here denoted $q \to tq$, is extended to the product to be trivial (but not isometric) in the second coordinate. The mapping class group acts on the product, commuting with the $\mathbb{R}$-action and preserving the fiberwise Hilbert norm. The quotient is an $\mathbb{R}$-action on a Hilbert bundle, $\mathbf{B}$, over the moduli space of unit norm quadratic differentials. This is the *Kontsevich-Zorich cocycle*. The Kontsevich-Zorich conjecture *contains* the statement that over each component of each stratum (of squares) $\mathbf{B}$ splits as a measurable sum of invariant, degree g stable-unstable bundles, to wit that 0 does not lie in its Lyapunov spectrum. Forni's proof of such a splitting involved introducing a "Forni cocycle", a measurably isomorphic bundle, defined by a splitting, $\mathbf{F} = H + V$ where over a.e. $q \sim (X, \omega^2)$, H (resp. V) is the space of invariant currents of dimension one and order one for the horizontal (resp. vertical) flow h_t^+ (h_t^-) associated to the holomorphic 1-form ω. Forni's map carries $\mathbf{F}$ a.e. isomorphically onto $\mathbf{B}$, with H and V going onto the stable and unstable bundles, respectively. With some further discussion Forni associates to the fibers of, say, V a g-dimensional space of invariant distributions for the vertical flow and obtains a theorem which implies the following: For a.e. element of the stratum, if a test function (weak L^2 derivatives) f is in the kernel of all of the invariant distributions, then for all x not lying on a horizontal separatrix

$$\int_0^T f(h_t^+ x)dt \;=\; o(T^\varepsilon), \quad \text{all } \varepsilon > 0. \tag{1.10}$$

One is left with a question: What are Forni's distributions? Bufetov's elegant answer involves a construction from the vertical flow associated to a generic $\xi = (X, \omega)$ of an isomorphic suspension flow over a Vershik-Markov compactum ([**Ver82**]). When the dust has settled, he finds associated to a.e. $\xi = (X, \omega)$ a $g-$dimensional linear space, $\mathcal{N}_\xi$, of continuous, additive, horizontal-holonomy invariant cocycles $\Phi(x,t)$ for the vertical flow. The cocycle enjoys the additional property that each $\Phi(x,t)$ is Hölder in t with exponent determined by where Φ sits in terms of a gradation of $\mathcal{N}_\xi$ determined by the Lyapunov spectrum. The gradation is available thanks to the work of Forni and Avila-Viana ([**F02**], [**AV07**]). Bufetov's construction of $\Phi \in \mathcal{N}_\xi$ is quite explicit, reminiscent of a Neumann expansion. He finally constructs a linear map $f \to \Phi_f$ from $Lip(X) \to \mathcal{N}_\xi$ such that the cocycle $\Phi_f(x,T)$ approximates the cocycle $\Psi_f(x,T) = \int_0^T f(h_t^+ x)dt$, in the sense that

$$\left| \int_0^T f(h_t^+ x)dt \;-\; \Phi_f(x,T) \right| \le C_\varepsilon \left(1 + T^\varepsilon\right) \|f\|_{Lip}, \quad \text{all } \varepsilon, T > 0. \tag{1.11}$$

For any fixed x, T $f \to \Phi_f(x,T)$ is, *mirabile dictu,* a Forni distribution.

2. The spaces $\mathcal{L}(T)$ and $\mathcal{Z}(T)$

Notations and assumptions are as in the first paragraphs of Section 1. Let $\mathcal{F}_{|\lambda|}$ be the algebra of finite unions of intervals $[a,b) \subseteq [0,|\lambda|)$. If $\varphi \in \mathbf{C}_0\left([0,|\lambda|]\right)$, define the additive set function ν^φ on $\mathcal{F}_{|\lambda|}$ by assigning

$$\nu^\varphi\left([a,b)\right) = \varphi(b) - \varphi(a) \tag{2.1}$$

to the generators. Since $T^{-1}\mathcal{F}_{|\lambda|} = \mathcal{F}_{|\lambda|}$, it makes sense to define $T\upsilon$ and the notion of invariant element. Define

$$\mathcal{L}(T) = \left\{\varphi \in \mathbf{C}_0\left([0,|\lambda|]\right) \mid T\nu^\varphi = \nu^\varphi\right\}. \tag{2.2}$$

We shall approach the study of $\mathcal{L}(T)$ by first analyzing a more general class of functions.

Let $\mathcal{D} \subseteq I^\lambda = [0,|\lambda|)$ be a nonempty set that contains $\beta_k(\lambda) = \sum_{1 \le i \le k} \lambda_i$, $0 \le k < m$ and satisfies

$$T^{-1}\mathcal{D} \subseteq \mathcal{D}, \; T = T_{(\lambda,\pi)}. \tag{2.3}$$

Define $\mathcal{F}\left(\mathcal{D}\right)$ to be the algebra of finite unions of intervals $[a,b)$, $a, b \in \mathcal{D} \cup \{|\lambda|\}$. Evidently,

$$T^{-1}\mathcal{F}\left(\mathcal{D}\right) \subseteq \mathcal{F}\left(\mathcal{D}\right). \tag{2.4}$$

If $\varphi \in \mathbb{R}^{\mathcal{D} \cup \{|\lambda|\}}$, define an additive set function ν^φ on $\mathcal{F}\left(\mathcal{D}\right)$ by assigning

$$\nu^\varphi\left([a,b)\right) = \varphi(b) - \varphi(a), a, b \in \mathcal{D} \cup \{|\lambda|\}$$

to the generators. Since $T^{-1}\mathcal{F}(\mathcal{D}) \subseteq \mathcal{F}(\mathcal{D})$, it makes sense to define $T\upsilon$ and the notion of invariant element. Define

$$\mathcal{L}_{\mathcal{D}}(T) = \left\{\varphi \in \mathbb{R}^{\mathcal{D}\cup\{|\lambda|\}} |\ \varphi(0) = 0 \text{ and } T\nu^{\varphi} = \nu^{\varphi}\right\}. \tag{2.5}$$

It is clear that there is a natural one-to-one correspondence between $\mathcal{L}_{\mathcal{D}}(T)$ and the set of real-valued, additive, invariant set functions on $\mathcal{F}(\mathcal{D})$.

REMARK 2.1. When (2.3) is strengthened by

$$T^{\pm 1}\mathcal{D} \subseteq \mathcal{D} \tag{2.6}$$

i.e., by $T^{-1}\mathcal{D} = \mathcal{D}$, then

$$\mathcal{L}_{\mathcal{D}}(T) = \mathcal{L}_{\mathcal{D}}(T^{-1}) \tag{2.7}$$

and

$$T^{\pm 1}\nu^{\varphi} = \nu^{\varphi},\ \varphi \in \mathcal{L}_{\mathcal{D}}(T). \tag{2.8}$$

NOTATION 2.2. let $\mathcal{D}(T)$ be the set

$$\mathcal{D}(T) = \bigcup_{-\infty<n<\infty} T^{-n}\left\{\beta_k(\lambda)|\ 1 \leq k < m\right\}. \tag{2.9}$$

In what follows we shall identify the set $\mathcal{L}_{\mathcal{D}(T)}(T)$ with the help of Rauzy induction.

What follows is a reminder of how Rauzy induction is defined. See [**R79**], and, for notation, [**V81**], [**V82**]. Denote by $\mathfrak{S}_n^0$, $n > 1$, the set of irreducible permutations on $\{1, 2, \cdots, n\}$. Represent a permutation $\pi \in \mathfrak{S}_n^0$ as a function, i.e.,

$$\pi = (\pi(1), \pi(2), \cdots, \pi(n)). \tag{2.10}$$

There are two Rauzy operations a and b: If $j = \pi^{-1}(n)$, then $a\pi$ is represented by

$$a\pi = (\pi(1), \pi(2), \cdots, \pi(j-1), n, \pi(n), \pi(j+1), \cdots, \pi(n-1)). \tag{2.11}$$

$b\pi$ may be represented in terms of (2.11) as

$$b\pi = \left(a\pi^{-1}\right)^{-1}.$$

A direct definition of $b\pi$ is

$$(b\pi)(i) = \left\{\begin{array}{c} \pi(i),\ \pi(i) \leq \pi(n) \\ \pi(i)+1,\ \pi(n) < \pi(i) < n \\ \pi(n)+1,\ i = \pi^{-1}(n) \end{array}\right. . \tag{2.12}$$

The operations a and b determine, for each $\pi \in \mathfrak{S}_m^0$, a pair of visitation matrices, $A(\pi, a)$ and $A(\pi, b)$. While $a\pi$ and $b\pi$ depend upon the representation (2.10) for π, $A(\pi, a)$ and $A(\pi, b)$ depend only upon the value $\pi^{-1}(m)$. More precisely, we recall that

$$A(\pi, a)_{ij} = \left\{\begin{array}{c} \delta_{ij} + \delta_{i,\pi^{-1}(m)}\delta_{j,\pi^{-1}(m)+1},\ (i,j) \in \left[1, \cdots, \pi^{-1}(m)\right] \times [1, \cdots, m] \\ \delta_{i+1,j},\ (i,j) \in \left[\pi^{-1}(m)+1, \cdots, m-1\right] \times [1, \cdots, m] \\ \delta_{\pi^{-1}(m)+1,j},\ i = m, j \in [1, \cdots, m] \end{array}\right.$$

(2.13)

$$A(\pi, b)_{ij} = \left\{\begin{array}{c} \delta_{ij},\ (i,j) \in [1, \cdots, m-1] \times [1, \cdots, m] \\ \delta_{mj} + \delta_{\pi^{-1}m,j},\ i = m, j \in [1, \cdots, m] \end{array}\right.$$

Use R to denote Rauzy induction, which is defined in terms of an a.e. dichotomy

$$R(\lambda,\pi) = \left(c\pi, A(c,\pi)^{-1}\lambda\right)$$
$$(2.14)\qquad c = c(\lambda,\pi) = \begin{cases} a, , \lambda_m < \lambda_{\pi^{-1}m} \\ b, \lambda_m > \lambda_{\pi^{-1}m} \end{cases}$$

Denote the n^{th} iteration of Rauzy induction by

$$(2.15)\qquad R^n(\lambda,\pi) = \left(\pi_n, \left(A^{(n)}(\lambda,\pi)\right)^{-1}\lambda\right) \stackrel{def}{=} \left(\pi_n, \lambda^{(n)}\right), \ n > 0.$$

The cone

$$S(\lambda,\pi) = \bigcap_{n=1}^{\infty}\left(A^{(n)}(\lambda,\pi)\left(\mathbb{R}^+\right)^m\right)$$

can be identified with the cone of finite invariant measures for $T_{(\lambda,\pi)}$ ([**V78**]).

The original exchange $T_{(\lambda,\pi)}$ may be viewed as having been created by stacking over the exchange $T_{\left(\pi_n,\lambda^{(n)}\right)}$. The interval $I_j^{\lambda^{(n)}}$ is the base of a column holding $a_j^{(n)}$ identical copies of the base, where

$$(2.16)\qquad a_j^{(n)} = \sum_{i=1}^{m} A_{ij}^{(n)}.$$

Let $\mathcal{F}_n(T)$ denote the algebra generated by the elements of the stacks for $R^n(\lambda,\pi)$. ($\mathcal{F}_0(T)$ is taken to be generated by the atoms I_i^{λ}, $1 \le i \le m$.) Clearly,

$$(2.17)\qquad \mathcal{F}_n(T) \subset \mathcal{F}_{n+1}(T), \ n \ge 0.$$

Let $1 \le k < m$ and $n > 0$. There exist integers $p(k,n) \ge 0$ and $l(k,n) \in [1,m)$ such that the division point $\beta_k(\lambda^{(n)})$ has the form

$$\beta_k(\lambda^{(n)}) = T^{-p(k,n)}\beta_{l(k,n)}(\lambda)$$
$$(2.18)\qquad \lim_{n\to\infty} p(k,n) = +\infty, 1 \le k < m.$$

There also exists $q(k,n) \ge 0$ such that the return map on $I^{\lambda^{(n)}}$ satisfies

$$T_{\left(\pi_n,\lambda^{(n)}\right)}\beta_k(\lambda^{(n)}) = T^{q(k,n)}\beta_{l(k,n)}(\lambda)$$
$$(2.19)\qquad \lim_{\substack{n\to\infty \\ l(k,n)\ne\pi^{-1}1}} q(k,n) = +\infty, \ 1 \le k < m$$

Since it is also true that

$$(2.20)\qquad T_{\left(\pi_n,\lambda^{(n)}\right)}0 = T^{1+q_n}\beta_{\pi^{-1}1}(\lambda), \ q_n = q(1,n) \to +\infty,$$

we can say that $\bigcup_{n=0}^{\infty}\mathcal{F}_n(T)$ contains every interval $[a,b)$ with endpoints in $\mathcal{D}(T)$. That is,

Lemma 2.3. *With notations as above,*

$$(2.21)\qquad \mathcal{F}(\mathcal{D}(T)) = \bigcup_{n=0}^{\infty}\mathcal{F}_n(T).$$

In particular,

$$\mathcal{L}_{\mathcal{D}(T)}(T) = \mathcal{L}_{\mathcal{D}(T)}(T^{-1}). \tag{2.22}$$

LEMMA 2.4. *Let $n \geq 0$. There exists $N > n$ such that*

$$I_1^{\lambda^{(n)}} = I^{\lambda^{(N)}}. \tag{2.23}$$

PROOF. It is a feature of Rauzy induction that it involves either (a) the deletion of the m^{th} interval and the creation in the $\left(\pi^{-1}m\right)^{th}$ interval of a new division point $\beta_{\pi^{-1}m}(\lambda) - (\beta_m(\lambda) - \beta_{m-1}(\lambda))$ or (b) the creation of a new right endpoint $\beta_m(\lambda) - (\beta_{\pi^{-1}m}(\lambda) - \beta_{\pi^{-1}m-1}(\lambda))$, interior to the m^{th} interval ([**R79**]). Since the i.d.o.c. assumption implies intervals upon which induction is performed decrease to 0, every division point must eventually be deleted. Said another way, every division point is eventually the right endpoint of an interval on which induction is performed. Apply this observation to $T_{\left(\lambda^{(n)},\pi_n\right)}$ and the division point $\beta_1\left(\lambda^{(n)}\right)$. $\square$

LEMMA 2.5. *Let $\mathcal{F}_n'(T)$ consist of elements of $\mathcal{F}_n(T)$ that do not intersect the base together with the full base $I^{\lambda^{(n)}}$. Given $n \geq 0$, there exists $N > n$ such that*

$$\mathcal{F}_n(T) \subset \mathcal{F}_N'(T). \tag{2.24}$$

PROOF. If $1 < j \leq m$, and if $N > n$ is large enough that $I^{\lambda^{(N)}} \subset I_1^{\lambda^{(n)}}$, then $I_j^{\lambda^{(n)}} \in \mathcal{F}_N'(T)$. The case $j = 1$ is a consequence of the previous lemma. $\square$

CONSTRUCTION. Let there be given a vector w in $\mathbb{R}^m$. Use w and Rauzy induction to determine a sequence of vectors

$$w^{(n)} = \left(A^{(n)}\left(\lambda,\pi\right)\right)^{-1} w,\ n > 0. \tag{2.25}$$

Then use the j^{th} coordinate of $w^{(n)}$ as the value assigned to each element of the stack above $I_j^{\lambda^{(n)}}$. This determines a set function ν_n on $\mathcal{F}_n(T)$. Clearly,

$$\nu_n(T^{-1}B) = \nu_n(B),\ B \in \mathcal{F}_n'(T). \tag{2.26}$$

We shall now compare ν_n with ν_{n+1}.

LEMMA 2.6. *For all $n \geq 0$*

$$\nu_{n+1}|_{\mathcal{F}_n(T)} = \nu_n. \tag{2.27}$$

PROOF. View the first two vectors, w and $w^{(1)}$, as assigning "masses" to basic intervals. In the notation of the proof of Lemma 2.4 one finds in case (a) that $w^{(1)}$ arises from w by dropping the last coordinate and replacing the $\left(\pi^{-1}m\right)^{th}$ coordinate by two coordinates with successive masses $w_{\pi^{-1}m} - w_m$ and w_m. In case (b) the last coordinate of w is simply replaced by $w_m - w_{\pi^{-1}m}$. If one next considers the stacking process as it moves from $R^n\left(\lambda,\pi\right)$ to $R^{n+1}\left(\lambda,\pi\right)$, one finds that either (a) each interval in the $\left(\pi_n^{-1}m\right)^{th}$ stack at time n is divided into two intervals that are assigned masses $w^{(n)}_{\pi_n^{-1}m} - w^{(n)}_m$ and $w^{(n)}_m$ or (b) each interval in the m^{th} stack at time n is divided into two intervals that are assigned masses $w^{(n)}_m - w^{(n)}_{\pi^{-1}m}$ and $w^{(n)}_{\pi^{-1}m}$. It is now clear that (2.27) is true. $\square$

PROPOSITION 2.7. *With all notations as above, the set functions $\nu_n, n \geq 0$, coalesce to define an additive and invariant set function, ν^w, on $\mathcal{F}\left(\mathcal{D}(T)\right)$.*

PROOF. Lemmas 2.3 and 2.6 imply the existence of an additive set function ν^w on $\mathcal{F}(\mathcal{D}(T))$ such that $\nu^w|_{\mathcal{F}_n(T)} = \nu_n, n \geq 0$. By (2.26) and Lemma 2.5 ν^w is invariant. □

With notations as above, we associate to ν_n a step function φ_n on $[0, |\lambda|]$. Define $B_n(x)$ to be the union of elements of $\mathcal{F}_n(T)$ that are contained in the interval $[0, x)$, and then set

$$\varphi_n(x) = \nu_n\left(B_n(x)\right),\ \varphi_n(0) = 0.$$

Evidently, φ_n and φ_{n+1} agree at any x such that $[0, x) \in \mathcal{F}_n(T)$. Therefore, the pointwise limit

$$\varphi^w(x) \overset{def}{=} \lim_{n\to\infty} \varphi_n(x),\ x \in \mathcal{D}(T)$$

exists. Define

$$\nu^w\left([a, b)\right) = \varphi^w(b) - \varphi^w(a),\ a, b \in \mathcal{D}(T), \tag{2.28}$$

and observe that $\varphi^w \in \mathcal{L}_{\mathcal{D}(T)}(T)$.

NOTATION 2.8. Let $\mathcal{I}_{\mathcal{D}(T)}(T)$ denote the set of finite invariant additive set functions ν on $\mathcal{F}(\mathcal{D}(T))$.

PROPOSITION 2.9. *With notations as above, (2.28) defines a linear isomorphism between $\mathcal{L}_{\mathcal{D}(T)}(T)$ and $\mathcal{I}_{\mathcal{D}(T)}(T)$. The natural map $\nu \to w^\nu$,*

$$w_i^\nu = \nu\left(I_i^\lambda\right), 1 \leq i \leq m. \tag{2.29}$$

defines a linear isomorphism of $\mathcal{I}_{\mathcal{D}(T)}(T)$ onto $\mathbb{R}^m$.

PROOF. The first statement is clear from the discussion above. The map $\nu \to w^\nu$ is onto by Proposition 2.7. To prove that it is injective, fix $\nu \in \mathcal{I}_{\mathcal{D}(T)}(T)$. With respect to Rauzy induction, the interval $I_j^{\lambda^{(n)}}$ is the base of a column holding $\sum_{i=1}^m A_{ij}^{(n)} = a_j^{(n)}$ identical copies of the base. Invariance of ν implies each copy of $I_j^{\lambda^{(n)}}$ is assigned the same ν-measure. Coupled with additivity this implies

$$\sum_{j=1}^m A_{ij}^{(n)} \nu\left(I_j^{\lambda^{(n)}}\right) = \nu\left(I_i^\lambda\right) = w_i^\upsilon. \tag{2.30}$$

In particular,

$$\nu\left(I_j^{\lambda^{(n)}}\right) = \left(\left(A^{(n)}\left(\lambda, \pi\right)\right)^{-1} w^\nu\right)_j,\ 1 \leq j \leq m,\ n > 0 \tag{2.31}$$

Therefore, ν is determined by w. □

PROPOSITION 2.10. *If $\varphi \in \mathcal{L}_{\mathcal{D}(T)}(T)$, and if $\nu^\varphi \in \mathcal{I}_{\mathcal{D}(T)}(T)$ is the associated set function, then $w(\nu^\varphi) \in \mathbb{R}^m$ and φ satisfy*

$$\varphi\left(Tx\right) - \varphi\left(x\right) = z_i^\varphi, x \in \mathcal{D}(T) \cap I_i^\lambda$$

(2.32) *where*

$$z_i^\varphi = \left(L^\pi w(\nu^\varphi)\right)_i,\ 1 \leq i \leq m.$$

PROOF. It is a consequence of invariance that

$$\varphi(Tx) - \varphi(T\beta_{i-1}) = \varphi(x) - \varphi(\beta_{i-1})$$

(2.33) and

$$\varphi(Tx) - \varphi(x) = \varphi(T\beta_{i-1}) - \varphi(\beta_{i-1}),\ 1 \le i \le m,\ x \in \mathcal{D}(T) \cap I_i^\lambda.$$

By definition

$$\varphi(\beta_{i-1}) = \nu^\varphi([0, \beta_{i-1})) = \sum_{j<i} w(\nu^\varphi)_j. \tag{2.34}$$

Since $T\nu^\varphi = \nu^\varphi$, it is also true that

$$\begin{aligned} \varphi(T\beta_{i-1}) &= \nu^\varphi([0, T\beta_{i-1})) \\ &= \nu^\varphi\left(T^{-1}[0, T\beta_{i-1})\right) = \sum_{\pi k < \pi i} w(\nu^\varphi)_k. \end{aligned} \tag{2.35}$$

It follows that

$$\begin{aligned} \varphi(T\beta_{i-1}) - \varphi(\beta_{i-1}) &= \sum_{\pi k < \pi i} w(\nu^\varphi)_k - \sum_{j<i} w(\nu^\varphi)_j \\ &= \left(\sum_{\substack{k<i \\ \pi k < \pi i}} w(\nu^\varphi)_k + \sum_{\substack{k>i \\ \pi k < \pi i}} w(\nu^\varphi)_k \right) - \sum_{j<i} w(\nu^\varphi)_j \\ &= \sum_{\substack{i<k \\ \pi k < \pi i}} w(\nu^\varphi)_k - \sum_{\substack{j<i \\ \pi j > \pi i}} w(\nu^\varphi)_j \\ &= \left(L^\pi w(\nu^\varphi)\right)_i = z_i^\varphi. \end{aligned} \tag{2.36}$$

Now (2.32) follows from (2.33). □

The proof of Lemma 2.6 implies that if $[0, x) \in \mathcal{F}_{n+1}(T) \backslash \mathcal{F}_n(T)$, then x divides an atom of $\mathcal{F}_n(T)$, and

$$|\varphi_n(x) - \varphi_{n+1}(x)| \le 2 \left\| w^{(n)} \right\|_\infty.$$

In particular,

$$\|\varphi_n - \varphi_{n+1}\|_\infty \le 2 \left\| w^{(n)} \right\|_\infty. \tag{2.37}$$

PROPOSITION 2.11 ([**V84-I**], Theorem 8.6). *With notations as above suppose it is true that*

$$\sum_n \left\| w^{(n)} \right\| < \infty. \tag{2.38}$$

There exists $\varphi \in \mathbf{C}([0,|\lambda|])$ *such that*

$$\lim_{n\to\infty} \|\varphi_n - \varphi\|_\infty = 0$$

$$\varphi^w = \varphi|_{\mathcal{D}(T)}.$$

PROOF. It follows from (2.37) and (2.38) that the sequence $\{\varphi_n\}$ converges uniformly to a function φ on $[0,|\lambda|]$. φ is continuous at any point $x \in [0,|\lambda|)$ that is not an endpoint of an atom of $\mathcal{F}_n(T)$ for some n. These are the points that do **not** have the form $T^{-k}\beta_i(\lambda)$ for any $k \geq 0$ and $1 \leq i < m$. Define a function $\delta(\cdot)$ by

$$\delta(x) = \limsup_{y',y''\to x} |\varphi(y') - \varphi(y'')|.$$

$\delta(\cdot)$ is of the first Baire class and vanishes at the dense set of points of continuity of φ. Following [**V84-I**], pp. 1351-52, we consider two cases:

Case 1.: $x = T^l\beta_j$, $x > 0, l > 0, 1 \leq j < m$. In this case, the function

$$\sum_{i=1}^{m} z_i^\varphi \chi_{I_i^\lambda}(\cdot) \tag{2.39}$$

is continuous at $x, Tx, T^2x, \cdots$, and (2.32) implies $\delta(x) = \delta\left(T^k x\right)$, $k \geq 0$. Since the forward orbit of x is dense, it must be that $\delta(x) = \delta\left(T^k x\right) = 0, k \geq 0$.

Case 2.: $x = T^l\beta_j, l \leq 0, 1 \leq j < m$. In this case (2.39) is continuous at the points $T^k x$, $k < 0$, and therefore $\delta(x) = \delta\left(T^k x\right), k \leq 0$. Again it follows that $\delta(x) = 0$.

Cases one and two combine to imply that φ is continous on $(0,|\lambda|)$. Continuity at 0 is evident from (2.32) and continuity at $T0$. Continuity at $|\lambda|$ follows from $\varphi(|\lambda|) = \varphi_n(|\lambda|)$, $n > 0$ and the uniformity of convergence. □

REMARK 2.12. If $\varphi \in \mathcal{L}_{\mathcal{D}(T)}(T)$ extends to a continuous function on $[0,|\lambda|]$, then ν^φ extends to a set function on $\mathcal{F}_{|\lambda|}$. This implies φ belongs to the space $\mathcal{L}(T)$ in (2.2).

PROPOSITION 2.13. *Let* $w \in \mathbb{R}^m \backslash \{0\}$ *be such that*

$$\lim_{n\to\infty} \left(A^{(n)}(\lambda,\pi)\right)^{-1} w = 0. \tag{2.40}$$

Then

$$w \notin \ker(L^\pi). \tag{2.41}$$

PROOF. One knows from [**V78**] and [**V84-I**] that

$$A^{(n)}(\lambda,\pi)^{-1}\ker(L^\pi) = \ker(L^{\pi_n}) \tag{2.42}$$

isometrically. In particular, if $w \in \ker(L^\pi)$, and if (2.40) is true, then $w = 0$. □

DEFINITION 2.14. $\mathcal{Z}_0(T)$ is the set of $w \in \mathbb{R}^m$ such that (2.40) is true. $\mathcal{Z}(T)$ is the linear (isomorphic) image in $\mathbb{R}^m$ of $\mathcal{L}(T)$, under the map $\varphi \to w^{\nu^\varphi}, \varphi \in \mathcal{L}(T)$.

PROPOSITION 2.15. $\mathcal{Z}_0(T)$ *is an isotropic subspace of* $\mathbb{R}^m$*, relative to* $(\cdot,\cdot)_\pi$*, where*

$$(\xi,\eta)_\pi = \xi^t L^\pi \eta, \xi, \eta \in \mathbb{R}^m. \tag{2.43}$$

Moreover,

$$\dim\left(\mathcal{Z}_0(T)\right) \leq \frac{1}{2} Rank\left(L^{\pi}\right). \tag{2.44}$$

PROOF. Proposition 2.13 implies $\mathcal{Z}_0(T) \cap \ker(L^{\pi}) = \{0\}$. Let $w_k \in \mathcal{Z}_0(T)$, $k = 1, 2$. By ([**V78**]) it is true for each n that

$$\left(\left(A^{(n)}(\lambda,\pi)\right)^{-1} w_1, \left(A^{(n)}(\lambda,\pi)\right)^{-1} w_2\right)_{\pi_n} = (w_1, w_2)_{\pi}. \tag{2.45}$$

Now (2.40) and (2.45) imply

$$\begin{aligned}(w_1, w_2)_{\pi} &= \lim_{n\to\infty} \left(\left(A^{(n)}(\lambda,\pi)\right)^{-1} w_1, \left(A^{(n)}(\lambda,\pi)\right)^{-1} w_2\right)_{\pi_n} \\ &= 0.\end{aligned} \tag{2.46}$$

That is, $\mathcal{Z}_0(T)$ is isotropic. Since it is also complementary to $\ker(L^{\pi})$, (2.44) follows. $\square$

LEMMA 2.16. *Let $w \in \mathbb{R}^m$ be such that φ^w is continous at* 0 *as a function on $D(T)$. Then* (2.40) *is true.*

PROOF. Clear from (2.31). $\square$

PROPOSITION 2.17. *$\mathcal{Z}(T)$ is an isotropic subspace of $\mathcal{Z}_0(T)$, relative to $(\cdot,\cdot)_{\pi}$. Moreover,*

$$\dim\left(\mathcal{Z}(T)\right) = \dim\left(\mathcal{L}(T)\right) \leq \frac{1}{2} Rank\left(L^{\pi}\right). \tag{2.47}$$

PROOF. Lemma 2.16 implies that the map $\varphi \to w^{\nu^{\varphi}}$, which by Proposition 2.9 is an injection, has image in the isotropic subspace $\mathcal{Z}_0(T)$. Therefore, (2.47) holds. $\square$

3. Finitely additive invariant measures and coboundaries

By definition $T = T_{(\lambda,\pi)}$ is continuous from the right on $I^{\lambda} = [0, |\lambda|)$. There is a corresponding $S = S_{(\lambda,\pi)}$ that is continous from the left on $I_{\lambda} = (0, |\lambda|]$. Of course, T and S disagree only at the points $\beta_k(\lambda)$, $0 \leq k \leq m$. It is useful to observe

LEMMA 3.1. *With notations and assumptions as in Proposition* 2.10*, it is true that*

$$\varphi\left(S\beta_k(\lambda)\right) - \varphi\left(\beta_k(\lambda)\right) = z_k^{\varphi},\ 1 \leq k \leq m \tag{3.1}$$

PROOF. For all $1 \leq k \leq m$ it is true that

$$TI_k^{\lambda} = [T\beta_{k-1}(\lambda), S\beta_k(\lambda))$$

and then

$$\begin{aligned}&\varphi\left(\beta_k(\lambda)\right) - \varphi\left(\beta_{k-1}(\lambda)\right) = \\ &= \nu^{\varphi}\left(I_k^{\lambda}\right) = \nu^{\varphi}\left(TI_k^{\lambda}\right) = \\ &= \varphi\left(S\beta_k(\lambda)\right) - \varphi\left(T\beta_{k-1}(\lambda)\right)\end{aligned} \tag{3.2}$$

Since $\varphi\left(T\beta_{k-1}(\lambda)\right) - \varphi\left(\beta_{k-1}(\lambda)\right) = z_k^{\varphi}$, $1 \leq k \leq m$, (3.1) is true. $\square$

DEFINITION 3.2. $\mathcal{J}_{\mathcal{D}(T)}(T)$ is the set of $\varphi \in \mathbb{R}^{\mathcal{D}(T)\cup\{|\lambda|\}}$ for which $\varphi(0) = 0$ and there exists $z^{\varphi} \in \mathbb{R}^m$ such that

$$\varphi(Tx) - \varphi(x) = z_k^{\varphi}, x \in \mathcal{D}(T) \cap I_k^{\lambda}$$

$$\text{and} \tag{3.3}$$

$$\varphi(S\beta_k(\lambda)) - \varphi(\beta_k(\lambda)) = z_k^{\varphi}, 1 \le k \le m\ .$$

PROPOSITION 3.3. *With notations as above,*

$$\mathcal{J}_{\mathcal{D}(T)}(T) = \mathcal{L}_{\mathcal{D}(T)}(T) \cong \mathbb{R}^m. \tag{3.4}$$

PROOF. Proposition 2.10 and Lemma 3.1 imply that

$$\mathcal{L}_{\mathcal{D}(T)}(T) \subseteq \mathcal{J}_{\mathcal{D}(T)}(T).$$

For the reverse inclusion, suppose given $\varphi \in \mathcal{J}_{\mathcal{D}(T)}(T)$. Define $w = w(\nu^{\varphi}) \in \mathbb{R}^m$, and with respect to Rauzy induction define

$$w_j^{(n)} = \nu^{\varphi}(I_j^{\lambda^{(n)}}), 1 \le j \le n. \tag{3.5}$$

As before, let $a_j^{(n)}$ be the return time on $I_j^{\lambda^{(n)}}$. Since $\varphi \in \mathcal{J}_{\mathcal{D}(T)}(T)$, each element in the stack above $I_j^{\lambda^{(n)}}$ has the same ν^{φ} measure, and this implies

$$\sum_{j=1}^{m} A_{ij}^{(n)} w_j^{(n)} = w_i(\nu^{\varphi}), 1 \le i \le m, n > 0. \tag{3.6}$$

The relations (3.6) imply $\nu^{\varphi} \in \mathcal{I}_{\mathcal{D}(T)}(T)$ so that $\varphi \in \mathcal{L}_{\mathcal{D}(T)}(T)$. The isomorphism with $\mathbb{R}^m$ is a consequence of Proposition 2.9. □

REMARK 3.4. If $\varphi \in \mathbf{C}_0([0, |\lambda|])$ satisfies the first line of (3.3), it automatically satisfies the second line of (3.3).

Given a column vector $z \in \mathbb{R}^m$, consider first the coboundary equation (which is (8.7) in [**V84-I**])

$$\varphi \circ T - \varphi = \sum_{i=1}^{m} z_i \chi_{I_i^{\lambda}}. \tag{3.7}$$

By our continuing assumption, $T = T_{(\lambda,\pi)}$ satisfies the i.d.o.c., and Keane's theorem ([**Ke75**]) implies T is minimal. Therefore for a given vector $z \in \mathbb{R}^m$, (3.7) admits at most one solution $\varphi = \varphi_z \in \mathbf{C}_0([0, |\lambda|])$. It also makes sense to set $z = z^{\varphi}$ when (3.7) holds.

DEFINITION 3.5. $\mathcal{J}(T)$ is the set

$$\mathcal{J}(T) = \{\varphi \in \mathbf{C}_0([0, |\lambda|]) \mid (3.7) \text{ holds for some } z = z^{\varphi} \in \mathbb{R}^m\}. \tag{3.8}$$

Proposition 3.3 and the definitions imply

PROPOSITION 3.6. *The sets $\mathcal{L}(T)$ and $\mathcal{J}(T)$ are equal,*

$$\mathcal{L}(T) = \mathcal{J}(T). \tag{3.9}$$

DEFINITION 3.7. Let $\mathcal{Z}_0(T)$ and $\mathcal{Z}(T)$ be as in Definition 2.14. Define $\mathcal{X}_0(T)$ and $\mathcal{X}(T)$ to be the linear images

$$\begin{aligned}\mathcal{X}_0(T) &= L^\pi \mathcal{Z}_0(T)\\ \mathcal{X}(T) &= L^\pi \mathcal{Z}(T).\end{aligned} \tag{3.10}$$

Then define spaces $\mathcal{W}_0(T)$ and $\mathcal{W}(T)$ by

$$\begin{aligned}\mathcal{W}_0(T) &= \mathcal{X}_0(T) + \mathcal{Z}_0(T)\\ \mathcal{W}(T) &= \mathcal{X}(T) + \mathcal{Z}(T).\end{aligned} \tag{3.11}$$

REMARK 3.8. Proposition 2.13 implies that L^π is injective on $\mathcal{Z}_0(T)$ and $\mathcal{Z}(T)$. Proposition 2.15 implies that $\mathcal{Z}_0(T)$ and $\mathcal{Z}(T)$ are Lagrangian subspaces of $\mathcal{W}_0(T)$ and $\mathcal{W}(T)$, respectively, relative to the restriction of $(\cdot,\cdot)_\pi$.

PROPOSITION 3.9.

$$\mathcal{W}_0(T) \cap \ker(L^\pi) = \{0\}\,. \tag{3.12}$$

PROOF. Suppose $\xi \in \mathcal{W}_0(T) \cap \ker(L^\pi)$, so that

$$\xi = z + w,\ \ z \in \mathcal{X}_0(T),\ w \in \mathcal{Z}_0(T),\ \ \xi \in \ker(L^\pi)\,. \tag{3.13}$$

It follows from (3.13), (2.40) and the discussion of (2.42) that

$$\sup_n \left\| \left(A^{(n)}(\lambda,\pi)\right)^{-1} z \right\| \ < \ \infty. \tag{3.14}$$

Since $z \in \mathcal{X}_0(T)$, we also have $z = L^\pi w_0, w_0 \in \mathcal{Z}_0(T)$, so that

$$\lim_n \left\| \left(A^{(n)}(\lambda,\pi)\right)^{*} z \right\| \ = \ \lim_n \left\| L^\pi \left(A^{(n)}(\lambda,\pi)\right)^{-1} w_0 \right\| = 0. \tag{3.15}$$

Since for all $n > 0$

$$\|z\|^2 \ = \left(\left(A^{(n)}(\lambda,\pi)\right)^{-1} z, \left(A^{(n)}(\lambda,\pi)\right)^{*} z \right),$$

it follows from (3.13) and (2.42) that $\xi = w$. Finally, Proposition 2.13 implies $\xi = 0$. □

Define $H(\pi) \subseteq \mathbb{R}^m$ by

$$H(\pi) = (\ker(L^\pi))^\perp\,. \tag{3.16}$$

Denote the orthogonal projection on $H(\pi)$ by

$$Q_\pi :\ \mathbb{R}^m \to H(\pi),\ \ \ker(Q_\pi) = H(\pi)^\perp = \ker(L^\pi)\,. \tag{3.17}$$

Proposition 3.9 implies

COROLLARY 3.10. *Q_π is injective on $\mathcal{W}_0(T)$, and, in particular,*

$$\mathcal{X}_0(T) \cap Q_\pi \mathcal{Z}_0(T) = \{0\}. \tag{3.18}$$

In fact,

$$\mathcal{X}_0(T) \perp Q_\pi \mathcal{Z}_0(T). \tag{3.19}$$

PROOF. Suppose $z = L^{\pi}w, z' = \mathcal{Q}_{\pi}w', w, w' \in \mathcal{Z}_0(T)$. As $(w' - \mathcal{Q}_{\pi}w') \in \ker(L^{\pi})$, it is true that

$$L^{\pi}\mathcal{Q}_{\pi}w' = L^{\pi}w'.$$

Since $\mathcal{Z}_0(T)$ is isotropic relative to $(\cdot,\cdot)_{\pi}$, it follows that

$$(z, z') = -w^* L^{\pi}\mathcal{Q}_{\pi}w' =$$

$$= -w^* L^{\pi} w' = 0.$$

Therefore $\mathcal{X}_0(T) \perp Q_{\pi}\mathcal{Z}_0(T)$. □

PROPOSITION 3.11. *With notations as above,*

$$\lim_{n\to\infty} \left\| A^{(n)}(\lambda,\pi)^*\eta \right\| = \infty,\ \ \eta \in Q_{\pi}\mathcal{Z}_0(T)\backslash\{0\}. \tag{3.20}$$

PROOF. Let $\eta = \xi + w$, $\eta \in H(\pi)$, $\xi \in H(\pi)^{\perp}$, $w \in \mathcal{Z}_0(T)$. Then

$$\|\eta\|^2 = \left(\left(A^{(n)}(\lambda,\pi)\right)^{-1}\eta, \left(A^{(n)}(\lambda,\pi)\right)^*\eta\right) =$$

$$\left(\left(A^{(n)}(\lambda,\pi)\right)^{-1}(\xi + w), \left(A^{(n)}(\lambda,\pi)\right)^*\eta\right) =$$

$$= 0 + \left(\left(A^{(n)}(\lambda,\pi)\right)^{-1}w, \left(A^{(n)}(\lambda,\pi)\right)^*\eta\right) \leq$$

$$\leq \left\|\left(A^{(n)}(\lambda,\pi)\right)^{-1}w\right\| \left\|\left(A^{(n)}(\lambda,\pi)\right)^*\eta\right\|.$$

By assumption $\eta \neq 0$, and therefore

$$\liminf_{n\to\infty} \left\| A^{(n)}(\lambda,\pi)^*\eta \right\| \ \geq\ \liminf_{n\to\infty} \frac{\|\eta\|^2}{\left\|\left(A^{(n)}(\lambda,\pi)\right)^{-1}w\right\|} = \infty. \tag{3.21}$$

□

4. $\mathcal{L}(T)$ and invariant distributions

Notations are as in Section 1. Let $\varphi \in \mathcal{L}(T)$, $F \in \mathcal{Y}_{|\lambda|}$. Given $n > 0$, denote by $\mathcal{E}_n(T)$ the subset of $\mathcal{D}(T)$ (see (2.9)) consisting of endpoints of atoms of the algebra $\mathcal{F}_n(T)$ (see the paragraph containing (2.17)). Order the points of $\mathcal{E}_n(T)$ as

$$\mathcal{E}_n(T) = \{0 = t_0 < t_1 < \cdots < t_N = |\lambda|\}. \tag{4.1}$$

(We suppress dependence upon n, which is fixed for the moment.) Define a functional $\Lambda_{\varphi}^{(n)}(F)$ by

$$\begin{aligned}\Lambda_{\varphi}^{(n)}(F) &= \sum_{i=1}^{N}\nu^{\varphi}\left([t_{i-1}, t_i)\right)F(t_{i-1}) = \\ &= \sum_{i=1}^{N}\left(\varphi(t_i) - \varphi(t_{i-1})\right)F(t_{i-1}).\end{aligned} \tag{4.2}$$

It follows from Abel's formula and the assumption $\varphi(0) = 0$ that

$$\Lambda_{\varphi}^{(n)}(F) = -\sum_{i=1}^{N-1}\varphi(t_i)\left(F(t_{i-1}) - F(t_{i-1})\right) \ +\ \varphi\left(|\lambda|\right)F(t_{N-1}). \tag{4.3}$$

Conclude from minimality of T, continuity of φ and the existence of $F(|\lambda|^-) = \lim_{t\nearrow|\lambda|} F(t)$ that

$$\lim_{n\to\infty} \Lambda_\varphi^{(n)}(F) \;=\; -\int_{[0,|\lambda|]} \varphi dF \;+\; \varphi\left(|\lambda|\right) F(|\lambda|^-) \overset{def}{=} \Lambda_\varphi(F), \tag{4.4}$$

the limit existing.

For the same data, consider $\Lambda_\varphi^{(n)}(\mathcal{T}F)$, where $\mathcal{T}F = F \circ T$. By definition,

$$\Lambda_\varphi^{(n)}(\mathcal{T}F) = \sum_{i=1}^{N} \left(\varphi(t_i) - \varphi(t_{i-1})\right) F(Tt_{i-1}). \tag{4.5}$$

The set function ν_φ is invariant under both $T^{\pm1}$ on $\mathcal{F}_{|\lambda|}$. Therefore, for all integers $i \in [1, N]$

$$\nu^\varphi\left(T[t_{i-1}, t_i)\right) = \nu^\varphi\left([t_{i-1}, t_i)\right). \tag{4.6}$$

In particular, it is true that

$$\begin{gathered}\left(\varphi(t_i) - \varphi(t_{i-1})\right) F(Tt_{i-1}) = \left(\varphi(Tt_i) - \varphi(Tt_{i-1})\right) F(Tt_{i-1}),\\ t_i \notin D_1(T) \cup \{|\lambda|\}.\end{gathered} \tag{4.7}$$

Define δ_n to be the mesh of the partition (4.1), and set

$$\varepsilon_n = \sup_{|x-y|\le\delta_n} |\varphi(x) - \varphi(y)|. \tag{4.8}$$

It follows from (4.5), (4.7), (4.8) and (1.4) that

$$\begin{aligned}&\left|\Lambda_\varphi^{(n)}(\mathcal{T}F) - \sum_{\substack{1\le i\le N\\ t_i\notin D_1(T)\cup\{|\lambda|\}}} \left(\varphi(Tt_i) - \varphi(Tt_{i-1})\right) F(Tt_{i-1})\right|\\ &\qquad = \left|\sum_{t_i\in D_1(T)\cup\{|\lambda|\}} \left(\varphi(t_i) - \varphi(t_{i-1})\right) F(Tt_{i-1})\right|\\ &\qquad \le m\varepsilon_n \left\|F\right\|_{\mathcal{Y}_{|\lambda|}}.\end{aligned} \tag{4.9}$$

By definition, the intervals of the partition of $[0, |\lambda|)$ determined by $\mathcal{F}_n(T)$ divide into m classes. The j^{th} class is the set of images

$$T^a I_j^{\lambda^{(n)}},\; 0 \le a < a_j^{(n)} = \sum_{i=1}^{m} A_{ij}^{(n)}.$$

In terms of the ordering (4.1) there exist integers $l\,(j,a)$ such that

$$T^a I_j^{\lambda^{(n)}} = [t_{l(j,a)-1}, t_{l(j,a)}),\; 1 \le l\,(j,a) \le N.$$

If $t_{l(j,a)} \notin D_1(T) \overset{def}{=} \{\beta_k(\lambda) |\; 1 \le k \le m\}$, and if $a < a_j^{(n)} - 1$, then

$$T^{a+1} I_j^{\lambda^{(n)}} = [Tt_{l(j,a)-1}, Tt_{l(j,a)}) = [t_{l(j,a+1)-1}, t_{l(j,a+1)}). \tag{4.10}$$

Define quantities P and Q by

$$P = \sum_{\substack{1\le i\le N \\ t_i\notin D_1(T)\cup\{|\lambda|\}}} (\varphi(Tt_i) - \varphi(Tt_{i-1}))\, F(Tt_{i-1})$$

$$Q = \sum_{j=1}^{m} \sum_{\substack{a< a_j^{(n)}-1 \\ t_{l(j,a)}\notin D_1(T)}} \left(\varphi(t_{l(j,a+1)}) - \varphi(t_{l(j,a+1)-1})\right) F(t_{l(j,a+1)-1})$$

By the definition of ε_n, it is true that

$$|P-Q| \le m\varepsilon_n \|F\|_{\mathcal{Y}_{|\lambda|}}. \tag{4.11}$$

It is also true that

$$\left|\Lambda_\varphi^{(n)}(F) - Q\right| \le 2m\varepsilon_n \|F\|_{\mathcal{Y}_{|\lambda|}}. \tag{4.12}$$

Finally, (4.9), (4.11) and (4.12) imply

$$\left|\Lambda_\varphi^{(n)}(\mathcal{T}F) - \Lambda_\varphi^{(n)}(F)\right| \le 4m\varepsilon_n \|F\|_{\mathcal{Y}_{|\lambda|}}.$$

Letting $n\to\infty$, we have

Proposition 4.1. *If $\varphi\in\mathcal{L}(T)$, then*

$$\Lambda_\varphi(\mathcal{T}F) = \Lambda_\varphi(F), F\in\mathcal{Y}_{|\lambda|}. \tag{4.13}$$

*In the notation of (*1.7*) and (*2.2*)*

$$\mathcal{L}(T)\subseteq\mathcal{K}(T). \tag{4.14}$$

We shall next establish the reverse inclusion to (4.14). Fix $\varphi\in\mathcal{K}(T)$, and let $[x,y), a>0$ be such that T^b is a translation on $[x,y)$ and $T^b[x,y)\subseteq I^\lambda_{i(b,x,y)}$for $0\le b\le a$. Let $z\in(x,y)$ and define $F_b=\chi_{T^b[x,z)}\in\mathcal{Y}_{|\lambda|}$. Then $F_b\circ T^b=\chi_{[x,z)}$. By definition

$$\Lambda_\varphi(F_b) = -\int_{[0,|\lambda|)} \varphi dF_b + \varphi(|\lambda|)\, F_b(|\lambda|^-)$$

$$= -\int_{[0,|\lambda|)} \varphi dF_b = \varphi\left(T^b z\right) - \varphi\left(T^b x\right).$$

and

$$\Lambda_\varphi\left(F_b\circ T^b\right) = \varphi(z) - \varphi(x).$$

Since φ is continuous, we can let $z\nearrow y$ to find that

$$\nu^\varphi\left(T^b[x,y)\right) = \nu^\varphi([x,y)). \tag{4.15}$$

Define $w=w(\nu^\varphi)\in\mathbb{R}^m$, and with respect to Rauzy induction define

$$w_j^{(n)} = \nu^\varphi(I_j^{\lambda^{(n)}}), 1\le j\le n. \tag{4.16}$$

As before, let $a_j^{(n)}$ be the return time on $I_j^{\lambda^{(n)}}$. By (4.15) each element in the stack above $I_j^{\lambda^{(n)}}$ has the same ν^φ measure, and this readily implies

$$\sum_{j=1}^{m} A_{ij}^{(n)} w_j^{(n)} = w_i(\nu^\varphi), 1\le i\le m. \tag{4.17}$$

In terms of Notation 2.8,

$$\nu^{\varphi} \in \mathcal{I}_{\mathcal{D}(T)}(T).$$

Since φ is continuous, ν^{φ} is invariant on the algebra $\mathcal{F}_{|\lambda|}$. We have proved

PROPOSITION 4.2. *If* $\varphi \in \mathcal{K}(T)$, *then*

$$T\nu^{\varphi} = \nu^{\varphi}. \tag{4.18}$$

Therefore,

$$\mathcal{K}(T) \subseteq \mathcal{L}(T). \tag{4.19}$$

Propositions 2.10, 3.6, 4.1, 2.17 and 4.2 combine to imply

THEOREM 4.3. *With notations as in (*1.7*),(*2.2*), Definition* 3.5 *and Definition* 3.7,

$$\begin{gathered} \mathcal{J}(T) = \mathcal{K}(T) = \mathcal{L}(T) \\ \mathcal{X}(T) = L^{\pi}\mathcal{Z}(T). \end{gathered} \tag{4.20}$$

*In particular, (*1.8*) is true, i.e.,*

$$Dim\,(\mathcal{K}(T)) \leq \frac{1}{2} Rank\,(L^{\pi})$$

5. Zippered rectangles flow

Let $\mathcal{R} = \mathcal{R}\,(\pi)$ denote a fixed Rauzy class, and let $\Omega\,(\mathcal{R})$ denote the corresponding space of unit norm zippered rectangles (see [**V82**] and [**V84-II**], Section 3). A point of $\Omega\,(\mathcal{R})$ has the form $x = (\lambda, h, a, \pi)\,, \pi \in \mathcal{R}$, $h \in H(\pi)$, $a \in Z(h,\pi)$ the rectangular parallelpiped of "zipper heights", $\lambda \in (\mathbb{R}^{+})^{m}$, $\lambda \cdot h = 1$. Define $Q^{s}, s \in \mathbb{R}$, on $\Omega\,(\mathcal{R})$ by

$$Q^{s}x = \left(e^{s}\lambda, e^{-s}h, e^{-s}a, \pi\right),\ x = (\lambda, h, a, \pi) \in \Omega\,(\mathcal{R})\,. \tag{5.1}$$

The *zippered rectangles moduli space* is the set

$$\begin{gathered} \Omega_0\,(\mathcal{R}) = \left\{ Q^{s}y |\ 0 \leq s < \log \tfrac{1}{|\lambda^{(1)}|},\ y \in Y\,(\mathcal{R}) \right\} \\ \text{where} \\ Y\,(\mathcal{R}) = \{ y = (\lambda, h, a, \pi) \in \Omega\,(\mathcal{R}) \mid\ |\lambda| = 1 \}\,. \end{gathered} \tag{5.2}$$

($\lambda^{(1)}$ corresponds to the first step of Rauzy induction.)

NOTATION 5.1. Notations such as $\mathcal{X}\,(y)\,, \mathcal{Z}\,(y)$,... shall be understood to pertain to the interval exchange $T_y = T_{(\lambda,\pi)}$, $y = (\lambda, h, a, \pi)$.

The flow Q^{s} determines a flow Q_0^{s} on the moduli space. For purposes of the present discussion it is necessary to recall only that there is a cocycle $A\,(\cdot,\cdot)$ such that in the first two coordinates

$$\begin{gathered} \lambda(Q_0^{t}x) = A^{-1}(t,x)e^{t}\lambda\,(x) \\ h(Q_0^{t}x) = A^{*}(t,x)e^{-t}h\,(x)\,,\ x = (\lambda, h, a, \pi) \in \Omega_0\,(\mathcal{R})\,. \end{gathered} \tag{5.3}$$

As noted in [**V84-II, V84-III**], the multiplicative ergodic theorem applies to the cocycles $A^{-1}(t,x)$ and $A^*(t,x)$. Applying a consequence ([**Kai87**]) of the multiplicative ergodic theorem, one finds that

$$\lim_{t\to+\infty}\left(\left(A^{-1}\right)^*(t,x)A^{-1}(t,x)\right)^{\frac{1}{2t}} = B(x)$$

$$\tag{5.4}\lim_{t\to+\infty}\left(A(t,x)A^*(t,x)\right)^{\frac{1}{2t}} = B^{-1}(x),$$

the limit existing a.e. $x \in \Omega_0(\mathcal{R})$. The important fact is that the moduli space, $\Omega_0(\mathcal{R})$, has finite volume for its natural smooth invariant measure. However, the crossection, $Y(\mathcal{R})$, has infinite measure with respect to contraction of this volume form. The Poincaré map on this crossection is the natural extension of Rauzy induction. For each Rauzy class the zippered rectangles flow on $\Omega_0(\mathcal{R})$ serves as a finite extension of the Teichmüller geodesic flow on a component of a stratum of abelian differentials ([**V86**]).

In a remarkable work ([**Z96**]), Zorich introduces an "acceleration" of Rauzy induction. The important feature of the Zorich induction is that it admits an absolutely continuous invariant probability measure on a corresponding crossection, to which the multiplicative ergodic theorem applies directly, yielding an asymptotic for the associated Zorich cocycle. We shall observe that, notwithstanding the fact the measure on $Y(\mathcal{R})$ (from [**V82**]) is infinite, the multiplicative ergodic theorem does yield an asymptotic for the Rauzy cocycle.

With notation as in Rauzy induction, the time

$$\tag{5.5} t(y) = \log\frac{1}{|\lambda^{(1)}|}$$

is the return time for the Poincaré map of the flow on the global crossection $Y(\mathcal{R})$. The n^{th} return time, $t_n(y), n > 0$, satisfies

$$t_n(y) = \log\tfrac{1}{|\lambda^{(n)}|}, n > 0$$

$$\tag{5.6}\lim_{n\to\infty}\frac{\log\frac{1}{|\lambda^{(n)}|}}{n} = 1,\ a.e.\ \lambda.$$

(See [**V84-II**], Section 5.)

Moreover, as noted in [**V84-II, V84-III**], the Rauzy cocycle may be represented, in terms of the moduli flow cocycle, as

$$\tag{5.7} A(t_n(y),y) = A^{(n)}(\lambda,\pi),\ y = (\lambda,h,a,\pi)\in Y(\mathcal{R}).$$

Now (5.6) implies it is also true that

$$\lim_{n\to+\infty}\left(\left(\left(A^{(n)}(\lambda,\pi)\right)^{-1}\right)^*\left(A^{(n)}(\lambda,\pi)\right)^{-1}\right)^{\frac{1}{2n}} =$$

$$\tag{5.8} = \lim_{t\to+\infty}\left(\left(A^{-1}\right)^*(t_n,y)A^{-1}(t_n,y)\right)^{\frac{1}{2t_n}} = B(y)$$

and

$$\lim_{t\to+\infty}\left(\left(A^{(n)}(\lambda,\pi)\right)\left(A^{(n)}(\lambda,\pi)\right)^*\right)^{\frac{1}{2n}} = B^{-1}(y),$$

the limit existing a.e. $y \in Y(\mathcal{R})$.

The matrix $B(y)$ admits a diagonalization

$$B(y) = O(y)\Delta O^*(y),\ O(y) \in SO(m) \tag{5.9}$$

in which Δ is an a.e. constant positive diagonal matrix with diagonal entries nondecreasing and inversion-symmetric about 1. The logarithms of the diagonal entries comprise the Lyapunov spectrum of the Rauzy cocycle. $O(y)$ is unique up to $O(y)O$, where O lies in the centralizer of Δ in $SO(m)$. Δ has at least $\dim(\ker(L^\pi))$ diagonal entries that are 1. Assume that there are γ entries that exceed 1. It is a consequence of Forni's Theorem ([**F02**]) about the Lyapunov spectrum of the Kontsevich-Zorich cocycle that, generically,

$$\gamma = \frac{1}{2} rank\,(L^\pi)\,. \tag{5.10}$$

(The connection between the Kontsevich-Zorich cocycle and the zippered rectangles/Rauzy cocycles is discussed in Sections 7-8.) It follows that, generically, Δ has exactly $\dim(\ker(L^\pi))$ diagonal entries that are 1. Avila-Viana ([**AV07**]) have proved the remaining eigenvalues are simple, as conjectured by Kontsevich-Zorich ([**Ko97**]). Let $V^-(y)$ (resp. $V^+(y)$) be the linear subspace of $\mathbb{R}^m$ corresponding to the eigenvalues that are less than one (resp. greater than one). Proposition 2.11, Corollary 2.17 and Forni's Theorem (5.10) imply that, generically and in the Notation 5.1,

$$O(y)V^-(y) = \mathcal{Z}(y) \tag{5.11}$$

Moreover, $\mathcal{Z}(y) + \ker(L^\pi)$ accounts for all of the eigenvalues that are at most 1, and therefore

$$O(y)V^+(y) = (\mathcal{Z}(y) + \ker(L^\pi))^\perp = L^\pi \mathcal{Z}(y) = \mathcal{X}(y) \tag{5.12}$$

6. Multiplicative coboundaries

We shall raise a question concerning the multiplicative version of (3.7), which is employed in [**V84-I, V84-II**] to establish generic weak mixing for an infinite set of permutations. A very deep analysis of this cocycle is carried out by of Avila and Forni, [**AF07**], and used to establish generic weak mixing for all non-rotation irreducible permutations. Let $\varphi \in \mathcal{J}(T) = \mathcal{L}(T)$ with $z = z^\varphi \in \mathcal{Z}(T)$. Define $e(s) = e^{2\pi i s}$ and set

$$\begin{aligned} Z_i^\varphi &= e\left(z_i^\varphi\right) \\ \Phi &= e\left(\varphi\right). \end{aligned} \tag{6.1}$$

We have $\Phi \in \mathbf{C}\left([0, |\lambda|]\right)$ and

$$\Phi \circ T = \left(\sum_{i=1}^{m} Z_i^\varphi \chi_{I_i^\lambda}\right) \Phi. \tag{6.2}$$

In what follows we assume that $h \in \mathcal{W}(T) = \mathcal{X}(T) + \mathcal{Z}(T)$ ((3.11)), and that Ψ is a measurable solution to (6.2), with z^φ replaced by h:

$$\Psi \circ T = \left(\sum_{i=1}^{m} e\left(h_i\right) \chi_{I_i^\lambda}\right) \Psi. \tag{6.3}$$

Since $h \in \mathcal{W}(T)$, we may express h as

$$h = z + w, w \in \mathcal{Z}(T), z \in \mathcal{X}(T) = L^{\pi}\mathcal{Z}(T). \tag{6.4}$$

By definition, there exists $\varphi \in \mathcal{J}(T)$ with $z = z^{\varphi}$. Form $\Phi = e(\varphi)$ in (6.1), and define Ξ by

$$\Xi = \frac{\Psi}{\Phi}. \tag{6.5}$$

By construction, Ξ is a measurable solution to

$$\Xi(Tx) = e(w_i)\Xi(x), a.e. x \in I_i^{\lambda}, 1 \leq i \leq m. \tag{6.6}$$

Use $T_n = T_{I^{\lambda^{(n)}}}$ to denote the induced map on $I^{\lambda^{(n)}}$. We have

$$\Xi(T_n y) = e\left(\left(\left(A^{(n)}\right)^* w\right)_j\right)\Xi(y), a.e.\ y \in I_j^{\lambda^{(n)}}. \tag{6.7}$$

Remark 6.1. Let E_ε be a set of integers such that if $n \in E_\varepsilon$, then

$$\varepsilon < \frac{\lambda_k^{(n)}}{\lambda_l^{(n)}} < \frac{1}{\varepsilon},\ 1 \leq k, l \leq m,$$

and if b is the largest integer such that T^a is a translation on $I^{\lambda^{(n)}}$ for all $a \in [0, b]$, then

$$b|\lambda^{(n)}| > \varepsilon|\lambda|.$$

If E_ε is an infinite set, a density point argument shows that

$$\lim_{\substack{n\to\infty \\ n \in E_\varepsilon}} e\left(\left(\left(A^{(n)}\right)^* w\right)_j\right) = 1, 1 \leq j \leq m. \tag{6.8}$$

Definition 6.2. Let $\widehat{\mathcal{Z}(T)}$ be the orthogonal complement of $\mathcal{X}(T)$ in $H(\pi)$,

$$\widehat{\mathcal{Z}(T)} = (\mathcal{X}(T))^{\perp} \cap H(\pi). \tag{6.9}$$

Denote the integer lattice in $H(\pi)$ by $H_{\mathbb{Z}}(\pi)$. If $\varkappa \in H_{\mathbb{Z}}(\pi)\backslash\{0\}$, $z_\varkappa \in \mathcal{X}(T)$ and $w_\varkappa \in \widehat{\mathcal{Z}(T)}$ are uniquely determined by

$$\varkappa = z_\varkappa + w_\varkappa, z_\varkappa \in \mathcal{X}(T), w_\varkappa \in \widehat{\mathcal{Z}(T)}. \tag{6.10}$$

Since $\varkappa$ is a nonzero integer vector, (2.40) and (3.15) implies $w_\varkappa \neq 0$. With notations as in (6.1)-(6.2) for $z_\varkappa$, observe that if

$$\begin{gathered} W_i = e(-z_\varkappa)_i = e(w_\varkappa)_i \\ \Psi = e(\varphi_{-z_\varkappa}), \end{gathered} \tag{6.11}$$

then Ψ is a nontrivial continuous solution to

$$\Psi \circ T = \left(\sum_{i=1}^{m} W_i \chi_{I_i^{\lambda}}\right)\Psi. \tag{6.12}$$

Since (6.4) is an orthogonal sum, the image, which we denote by $\widehat{\mathcal{Z}(T)}_{\mathbb{Z}}$, of $H_{\mathbb{Z}}(\pi)$ under the orthogonal projection of $H(\pi)$ onto $\widehat{\mathcal{Z}(T)}$ is a dense subgroup, consisting of elements for which the associated multiplicative coboundary equation admits a *continuous* solution.

REMARK 6.3. Forni's Theorem (5.10) implies that generically, in the notation (3.17),

$$
\begin{aligned}
\widehat{\mathcal{Z}(T)} &= Q_\pi \mathcal{Z}(T) \\
H(\pi) = \mathcal{X}(T) \perp Q_\pi \mathcal{Z}(T) &= L^\pi \mathcal{Z}(T) \perp Q_\pi \mathcal{Z}(T).
\end{aligned} \tag{6.13}
$$

Question.: Does there exist for a.e. λ an element $w = w(\lambda, \pi) \in \widehat{\mathcal{Z}(T_{(\lambda,\pi)})} \setminus \widehat{\mathcal{Z}(T_{(\lambda,\pi)})}_{\mathbb{Z}}$ such that the associated multiplicative coboundary equation (6.12) admits a nontrivial measurable solution?

7. The Kontsevich-Zorich cocycle

Notations are as in the first paragraph of Section 5. In what follows we shall sketch a connection between the zippered rectangles flow over a single Rauzy class and the Kontsevich-Zorich cocycle over a single component of a stratum of squares. Associate to $\mathcal{R} = \mathcal{R}(\pi)$ the numbers $g = g(\mathcal{R})$ and $n = n(\mathcal{R})$, which are the common values

$$
\begin{aligned}
g(\mathcal{R}) &= \tfrac{1}{2} rank\left(L^\pi\right), \pi \in \mathcal{R} \\
n(\mathcal{R}) &= m + 1 - 2g(\mathcal{R}).
\end{aligned} \tag{7.1}
$$

Let $M_{g,n}$ be a fixed closed, oriented surface of genus $g = g(\mathcal{R})$ with n punctures. If $x = (\lambda, h, a, \pi) \in \Omega(\mathcal{R})$, let $U(x)$ denote the union of the set of rectangles that realize x as a set of zippered rectangles. $U(x)$ has 0 as its lower lefthand vertex. Rectangles $j, j{+}1$, $j \neq \pi^{-1}m, m$, in $U(x)$ are glued along a common vertical segment from the base to the zipper height. When $j = \pi^{-1}m$, the zipper height may exceed the height of the rectangle $\pi^{-1}m$, in which case the rectangles are glued to the full height of rectangle $\pi^{-1}m$. We understand $U(x)^o$ to include the interiors of the glued segments. The interval exchange $T_{(\lambda,\pi)}$ creates a natural pairing between remaining free edges, and when these edges are identified by translation, the result is a surface, $M_n(x)$, that is homeomorphic to $M_{g,n}$. (See [**V82**].) We shall use $M(x)$ to denote the closed surface with the punctures included. If $x, x' \in \Omega(\mathcal{R})$ are such that $\pi(x) = \pi(x')$, there is a natural homotopy class (rel punctures) of homeomorphisms between $M_n(x)$ and $M_n(x')$. For example, use PL maps between rectangles, and arrange that these respect zippers, corners and therefore identifications. For each $\pi \in \mathcal{R}$ select an element $x_0 \in \Omega(\mathcal{R})$ with $\pi(x_0) = \pi$. Then select a relative homotopy class of homeomorphisms $M_{g,n} \to M_n(x_0)$. Now for any x with $\pi(x) = \pi$ the class of $M_n(x_0) \to M_n(x)$ determines a class $M_{g,n} \to M_n(x)$. In particular, if $\alpha \in H^1(M_g, \mathbb{R})$, M_g the unpunctured surface, there is a natural class $c(\alpha, x) \in H^1(M(x), \mathbb{R})$, $M(x)$ the unpunctured surface. This serves as a trivialization

$$
\bigcup_{x \in \Omega(\mathcal{R})} \{x\} \times H^1(M(x), \mathbb{R}) \cong \Omega(\mathcal{R}) \times H^1(M_g, \mathbb{R}). \tag{7.2}
$$

Recall from [**V82**] the partially defined map

$$
x' = \mathcal{U}x = \left(A(c,\pi)^{-1} \lambda, A(c,\pi)^* h, J(h,c)a, c\pi\right),\ x \in \Omega(\mathcal{R}). \tag{7.3}
$$

The only requirement on x is that $T_{(\lambda,\pi)}$ admit the first step of Rauzy induction, i.e. that $\lambda_m \neq \lambda_{\pi^{-1}m}$. If one removes from $\Omega(\mathcal{R})$ an appropriate countable union of hyperplanes (see [**V82**]), the set, $\Omega^*(\mathcal{R})$, which remains satisfies

$$\mathcal{U}\Omega^*(\mathcal{R}) = \Omega^*(\mathcal{R}) = \mathcal{U}^{-1}\Omega^*(\mathcal{R}). \tag{7.4}$$

If x is in the domain of $\mathcal{U}$, e.g., if $x \in \Omega^*(\mathcal{R})$, there is a tautological biholomorphism

$$F_{x,\mathcal{U}} : M_n(\mathcal{U}x) \to M_n(x), \tag{7.5}$$

which will be recalled in Section 8. Set up the map $\mathcal{V} : \Omega^*(\mathcal{R}) \times H^1(M_g, \mathbb{R}) \to \Omega^*(\mathcal{R}) \times H^1(M_g, \mathbb{R})$ as

$$\mathcal{V}(x,\alpha) = \left(\mathcal{U}x, F^*_{x,\mathcal{U}}\alpha\right), \ \ (x,\alpha) \in \Omega^*(\mathcal{R}) \times H^1(M_g, \mathbb{R}). \tag{7.6}$$

Notation 7.1. Let $\Lambda^*(\mathcal{R})$ denote the set

$$\Lambda^*(\mathcal{R}) = \Omega^*(\mathcal{R}) \times H^1(M_g, \mathbb{R}). \tag{7.7}$$

In the notation of (5.1), we have on the domain of $\mathcal{U}$

$$Q^t\mathcal{U} = \mathcal{U}Q^t, \ t \in \mathbb{R}. \tag{7.8}$$

Lemma 7.2. *For all $(x,\alpha) \in \Lambda^*(\mathcal{R})$ and $t \in \mathbb{R}$ it is true that*

$$F^*_{x,\mathcal{U}}\alpha = F^*_{Q^t x,\mathcal{U}}\alpha, \ t \in \mathbb{R}, \alpha \in H^1(M_g, \mathbb{R}). \tag{7.9}$$

Proof. The matrix $B_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$ satisfies

$$B_t U(w) = U(Q^t w), \ \ t \in \mathbb{R}, w = x, \mathcal{U}x.$$

It induces a puncture-preserving Teichmüller map

$$M(w) \xrightarrow{b_t^w} M(Q^t w), \ \ t \in \mathbb{R}, w = x, \mathcal{U}x.$$

Note also that b_t^w may serve as a PL map to identify cohomology classes as above. Since the diagram

$$\begin{array}{ccc} M_n(\mathcal{U}x) & \xrightarrow{F_{x,\mathcal{U}}} & M_n(x) \\ \downarrow b_t^{\mathcal{U}x} & & \downarrow b_t^x \\ M_n(Q^t\mathcal{U}x) = M_n(\mathcal{U}Q^t x) & \xrightarrow{F_{Q^t x,\mathcal{U}}} & M_n(Q^t x) \end{array}$$

is commutative, (7.9) follows. $\square$

Notation 7.3. Define a flow $(\Lambda^*(\mathcal{R}), \{q^t\}_{t\in\mathbb{R}})$ by

$$q^t(x,\alpha) = (Q^t x, \alpha), \ t \in \mathbb{R}. \tag{7.10}$$

Also, denote by $\Lambda_0^*(\mathcal{R})$ the orbit space

$$\Lambda_0^*(\mathcal{R}) = \Lambda^*(\mathcal{R}) / \left\{\mathcal{V}^l \mid l \in \mathbb{Z}\right\}.$$

Lemma 7.2 now implies

Lemma 7.4. *With notations as above, we have*

$$q^t\mathcal{V} = \mathcal{V}q^t, \ t \in \mathbb{R}. \tag{7.11}$$

Therefore, q^t descends to a flow, which we also denote by q^t, on $\Lambda_0^*(\mathcal{R})$:

$$\left(\Lambda_0^*(\mathcal{R}), \{q^t\}_{t\in\mathbb{R}}\right). \tag{7.12}$$

In terms of $\Omega_0(\mathcal{R})$ in (5.2) and $\Omega^*(\mathcal{R})$ in (7.4), define $\Omega_0^*(\mathcal{R})$ by

$$\Omega_0^*(\mathcal{R}) = \Omega_0(\mathcal{R}) \cap \Omega^*(\mathcal{R}).$$

From [**V82**] one has the relation

$$\Omega_0^*(\mathcal{R}) \cong \Omega^*(\mathcal{R}) / \left\{\mathcal{U}^l \mid l \in \mathbb{Z}\right\}, \tag{7.13}$$

which implies a further relation

$$\Lambda_0^*(\mathcal{R}) \cong \Omega_0^*(\mathcal{R}) \times H^1(M_g, \mathbb{R}). \tag{7.14}$$

The Rauzy class $\mathcal{R}$ determines a component of a stratum of squares for the Teichmüller flow. Up to scale this component supports a unique smooth invariant measure. The flow $(\Lambda_0^*(\mathcal{R}), \{q^t\}_{t\in\mathbb{R}})$ is a finite extension of the *Kontsevich-Zorich Cocycle restricted to this component* ([**Ko97**], [**F02**]).

8. The Rauzy cocycle

Let $x \in \Omega(\mathcal{R})$, and let $U(x)$, as in Section 7, be the union of the set of rectangles that realize x as a set of zippered rectangles. The corners of the rectangles together with the zippers comprise the vertex set of a cell decomposition, $\tau(x)$, of $M(x)$. Each element of the set, $\mathcal{E}(x)$, of edges of this decomposition is either horizontal or vertical. The horizontal edges lie on a single outgoing horizontal separatrix for the natural associated horizontal foliation of $M(x)$. Denote by $\mathcal{Z}_h^1(x)$ the space of horizontal cocycles for $\tau(x)$, i.e., cocycles that assign zero to each vertical element of $\mathcal{E}(x)$. Recall from [**V90**], Section 4, that there is a canonical isomorphism onto the cohomology of the punctured surface,

$$\mathcal{Z}_h^1(x) \overset{\vartheta_x}{\to} H^1(M_n(x), \mathbb{R}). \tag{8.1}$$

If one views an element $\varkappa \in \mathcal{Z}_h^1(x)$ as an assignment of real signed masses to the horizontal edges, the condition of being a cocycle is just the condition that the total signed mass of the base of each rectangle coincides with the total signed mass of its top.

LEMMA 8.1. *Assume* $x = (\lambda, h, a, \pi) \in \Omega^*(\mathcal{R})$. *Then* $\mathcal{I}_{\mathcal{D}(T)}(T)$ *and* $\mathcal{Z}_h^1(x)$ *are canonically isomorphic.*

PROOF. Let $\nu \in \mathcal{I}_{\mathcal{D}(T)}(T)$. For each i, $1 \le i \le m$, I_i^λ is the base of a rectangle constituent, $Z_i(x)$, of $U(x)$. $T_{(\lambda,\pi)}I_i^\lambda$ is (glued to) the top of $Z_i(x)$. Both I_i^λ and $T_{(\lambda,\pi)}I_i^\lambda$ lie in the algebra $\mathcal{F}(\mathcal{D}(T))$, and therefore

$$\nu\left(I_i^\lambda\right) = \nu\left(T_{(\lambda,\pi)}I_i^\lambda\right). \tag{8.2}$$

As all horizontal edges of $\tau(x)$ also lie in $\mathcal{F}(\mathcal{D}(T))$, it is possible to use ν to define a horizontal cochain $\varkappa_\nu$, and (8.2) implies that $\varkappa_\nu \in \mathcal{Z}_h^1(x)$. Proposition 2.9 implies that the numbers (8.2) uniquely determine ν. Therefore, the map $\nu \to \varkappa_\nu$ is one-to-one. Proposition 2.9 also implies that

$$\begin{aligned} \dim\left(\mathcal{I}_{\mathcal{D}(T)}(T)\right) = m = \\ = 2g(\mathcal{R}) + n(\mathcal{R}) - 1 = \\ = \dim H^1(M_n(x), \mathbb{R}) = \dim \mathcal{Z}_h^1(x). \end{aligned} \tag{8.3}$$

Therefore, the map $\nu \to \varkappa_\nu$ is also onto. ☐

NOTATION 8.2. Let $x = (\lambda, h, a, \pi) \in \Omega^* (\mathcal{R})$. If $1 \leq i \leq m$, define a cycle $\gamma_i (x)$ to be the sum of the vertical edges above β_{i-1} and the horizontal edges, appropriately oriented, from $T\beta_{i-1}$ to β_{i-1}. Represent this vertical-horizontal decomposition by

$$\gamma_i (x) = \gamma_i^v (x) + \gamma_i^h (x) . \tag{8.4}$$

It is evident from the definitions that the intersection numbers satisfy

$$[\gamma_i (x) , \gamma_j (x)] = L_{ij}^{\pi}, \ 1 \leq i, j \leq m. \tag{8.5}$$

It follows from (7.1) that the canonical image of the linear span

$$\Xi (x) = \left\{ \sum_{i=1}^{m} \xi_i \gamma_i (x) \, | \xi = (\xi_1, \xi_2, \cdot \cdot \cdot, \xi_m) \in \mathbb{R}^m \right\} \tag{8.6}$$

under the forgetful map $M_n(x) \to M(x)$ is all of $H^1(M(x), \mathbb{R})$.

LEMMA 8.3. *Let $x = (\lambda, h, a, \pi) \in \Omega^* (\mathcal{R})$, and set $T = T_{(\lambda,\pi)}$. Let $\nu \in \mathcal{I}_{\mathcal{D}(T)} (T)$, and define $\varkappa_\nu \in \mathcal{Z}_h^1(x)$ as in the proof of Lemma 8.1. With notations (*2.29*) and (*2.43*) for w^ν and $(\cdot, \cdot)_\pi$, one has*

$$\varkappa_\nu \left(\sum_{i=1}^{m} \xi_i \gamma_i (x) \right) = (w^\nu, \xi)_\pi . \tag{8.7}$$

PROOF. Let $\varphi \in \mathcal{L}_{D(T)}(T)$ be such that $\upsilon = \upsilon^\varphi$. By Section 4 of [**?**] there exists a smooth, real closed 1-form θ_ν on $M(x)$ such that θ_ν represents $\varkappa_\nu$, i.e.,

$$\widehat{\theta_\nu} \left(\sum_{i=1}^{m} \xi_i \gamma_i (x) \right) \overset{def}{=} \sum_{i=1}^{m} \xi_i \int\limits_{\gamma_i(x)} \theta_\nu = \varkappa_\nu \left(\sum_{i=1}^{m} \xi_i \gamma_i (x) \right). \tag{8.8}$$

Since $\varkappa_\nu \in \mathcal{Z}_h^1(x)$, it is true for each i that θ_ν has integral zero over $\gamma_i^v (x)$. That is,

$$\int\limits_{\gamma_i(x)} \theta_\nu = \int\limits_{\gamma_i^h(x)} \theta_\nu.$$

As $\gamma_i^h (x)$ is oriented from from $T\beta_{i-1}$ to β_{i-1}, Proposition 2.10 implies that

$$\int\limits_{\gamma_i^h(x)} \theta_\nu = \varphi (\beta_{i-1}) - \varphi (T\beta_{i-1}) = - (L^\pi w^\nu)_i . \tag{8.9}$$

Now (8.7) follows from (8.8)-(8.9). □

We shall now recall how the map (7.5) arises. Recall from [**V82**] the definition of the equivalence relation $x \approx x'$: If the identity map on the connected set $U(x)^o \cap U(x')^o$ extends to a biholomorphism between $M(x)$ and $M(x')$, then one declares $x \approx x'$. It is proved in [**V82**] that if x is in the domain of $\mathcal{U}$, then $x \approx \mathcal{U}x$. The biholomorphism is tautological and maps punctures to punctures. We may therefore represent the map $M(\mathcal{U}x) \to M(x)$ as a map of the punctured surfaces,

$$F_{x,\mathcal{U}} : M_n(\mathcal{U}x) \to M_n(x). \tag{8.10}$$

Because $F_{x,\mathcal{U}}$ *is* the identity map on $U(x)^o \cap U(\mathcal{U}x)^o$ it is true that

$$F_{x,\mathcal{U}}^* \theta_\nu = \theta_\nu \text{ on } U(x)^o \cap U(\mathcal{U}x)^o. \tag{8.11}$$

In particular, if $R(\lambda,\pi) = \left(\pi^{(1)},\lambda^{(1)}\right)$ as in (2.15), then $F^*_{x,\mathcal{U}}\theta_\nu$ and θ_ν are equal on the segment $I^{\lambda^{(1)}}$. However, since θ_ν is smooth and ν is only additive, it cannot be expected that θ_ν will represent ν on all of $\mathcal{F}(\mathcal{D}(T))$, $T = T_{(\lambda,\pi)}$. In particular, if $R(\lambda,\pi)$ is as in (2.14), θ_ν may not assign the value $\nu(J)$ to the one or two basic intervals $J \subset I^{\lambda^{(1)}}$, for $T_{R(\lambda,\pi)}$, that are not basic intervals for $T_{(\lambda,\pi)}$. At the same time, while θ_ν, as a function on 1-cochains of $\tau(x)$, determines a cocycle $\widehat{\theta_\nu}$ which vanishes on vertical edges, i.e., $\widehat{\theta_\nu} \in \mathcal{Z}^1_h(x)$, one cannot expect that $\widehat{F^*_{x,\mathcal{U}}\theta_\nu} \in \mathcal{Z}^1_h(\mathcal{U}x)$. The lemma to follow is intended to address these problems.

LEMMA 8.4. *Let* $x = (\lambda,h,a,\pi) \in \Omega^*(\mathcal{R})$, *and set* $T = T_{(\lambda,\pi)}$. *If* $\nu \in \mathcal{I}_{\mathcal{D}(T)}(T)$, *define* $\kappa(\nu) \in \mathcal{I}_{\mathcal{D}(T_{R(\lambda,\pi)})}\left(T_{R(\lambda,\pi)}\right)$ *by*

$$\kappa(\nu) = \nu|_{\mathcal{F}(\mathcal{D}(T_{R(\lambda,\pi)}))},\ \nu \in \mathcal{I}_{\mathcal{D}(T)}(T).$$

There exists a smooth closed 1-form, θ^1_ν, *on* $M(x)$ *such that*

$$\begin{gathered} \widehat{\theta^1_\nu} = \varkappa_\nu \\ \widehat{F^*_{x,\mathcal{U}}\theta^1_\nu} = \varkappa_{\kappa(\nu)}. \end{gathered} \tag{8.12}$$

PROOF. Begin with any θ_ν as in the proof of Lemma 8.3. The rectangles that comprise $U(x)$ may be denoted

$$Z_k(x) = I^\lambda_k \times [0,h_k),\ x = (\lambda,h,a,\pi),\ 1 \le k \le m. \tag{8.13}$$

Define rectangles Y and W in terms of the Rauzy dichotomy,

$$\begin{gathered} Y = \begin{cases} Z_m(x),\ \lambda_m < \lambda_{\pi^{-1}m} \\ [|\lambda| - \lambda_{\pi^{-1}m}, |\lambda|) \times [0,h_m),\ \lambda_{\pi^{-1}m} < \lambda_m \end{cases} \\ W = \begin{cases} [\beta_{\pi^{-1}m} - \lambda_m, \beta_{\pi^{-1}m}) \times [0,h_{\pi^{-1}m}),\ \lambda_m < \lambda_{\pi^{-1}m} \\ Z_{\pi^{-1}m},\ \lambda_{\pi^{-1}m} < \lambda_m \end{cases}. \end{gathered} \tag{8.14}$$

The operation from [**V82**] that replaces x by $\mathcal{U}x$ amounts to translating the rectangle Y so that its base coincides with the top of the rectangle W. Define a vertical segment $\gamma \subset U(x)^o$ by

$$\gamma = \begin{cases} \text{(a) left side of } Y,\ \lambda_{\pi^{-1}m} < \lambda_m \\ \text{(b) left side of } W,\ \lambda_m < \lambda_{\pi^{-1}m} \end{cases}. \tag{8.15}$$

γ divides the base and top of one of the rectangles, Z_m or $Z_{\pi^{-1}m}$, into two pieces. The closure of γ contains just one vertex of $\tau(x)$ (at the bottom in Case(8.15-(a)), at the top in Case (8.15-(b))). Since θ_ν is closed, there exists on $M(x)$ a smooth function f that vanishes on the vertices of $\tau(x)$ and that satisfies $df = \theta_\nu$ on a neighborhood of the closure of γ. Since f vanishes on the vertex set of $\tau(x)$,

$$\widehat{df} = 0. \tag{8.16}$$

Define θ^1_ν by

$$\theta^1_\nu = \theta_\nu - df. \tag{8.17}$$

Since θ^1_ν vanishes on γ, and since $\widehat{df}=0$, we have

$$\begin{aligned} &\int_\gamma \theta^1_\nu = 0. \\ (8.18)\qquad & \\ &\widehat{\theta^1_\nu} = \widehat{\theta_\nu} \text{ on } \mathcal{E}(x) \end{aligned}$$

Let δ' be a vertical edge in $\tau(\mathcal{U}x)$. If $F_{x,\mathcal{U}}(\delta') = \delta \in \mathcal{E}(x)$, then $\widehat{F^*_{x,\mathcal{U}}\theta^1_\nu}(\delta') = 0$. There is precisely one vertical edge in $\tau(\mathcal{U}x)$ that does not map to one in $\mathcal{E}(x)$. With notations as in (8.15) that edge is $\gamma' = (F_{x,\mathcal{U}})^{-1}(\gamma)$. By construction, $\widehat{F^*_{x,\mathcal{U}}\theta^1_\nu}(\gamma') = 0$. Therefore, $\widehat{F^*_{x,\mathcal{U}}\theta^1_\nu} \in \mathcal{Z}^1_h(\mathcal{U}x)$. In order to prove that (8.12) holds, it is sufficient to prove that

$$\int_J \theta^1_\nu = \nu(J) \tag{8.19}$$

holds for one of the new intervals in $I^{\lambda^{(1)}}$ created by $R(\lambda,\pi)$. (When there are two new intervals created, their union is the old interval $I^\lambda_{\pi^{-1}m}$ and (8.19) for both is a consequence of (8.19) for one.) Suppose that $\lambda_m < \lambda_{\pi^{-1}m}$. In this case the base, J, of W is a new interval while the top of W is identified to I^λ_m. Since the closed form θ^1_ν has integral zero over the vertical sides of W, its integrals over the base and top are equal. The top integral has the value $\nu\left(I^\lambda_m\right) = \nu\left(T^{-1}_{(\lambda,\pi)} I^\lambda_m\right) = \nu(J)$. Therefore, the bottom integral has the same value, and (8.19) is true in this case. An analogous argument proves (8.19) is true for the base of Y when $\lambda_{\pi^{-1}m} < \lambda_m$. ☐

Let $x = (\lambda,h,a,\pi) \in \Omega^*(\mathcal{R})$, and set $T = T_{(\lambda,\pi)}$. Recall the canonical isomorphism (8.1), $\mathcal{Z}^1_h(x) \overset{\vartheta_x}{\to} H^1(M_n(x),\mathbb{R})$, and use it to define an isomorphism $\Phi^*_{x,\mathcal{U}} : \mathcal{Z}^1_h(x) \to \mathcal{Z}^1_h(\mathcal{U}x)$ by

$$\begin{aligned} &\Phi^*_{x,\mathcal{U}} : \mathcal{Z}^1_h(x) \to \mathcal{Z}^1_h(\mathcal{U}x) \\ (8.20)\qquad & \\ &\Phi^*_{x,\mathcal{U}} = \vartheta^{-1}_{\mathcal{U}x} \circ F^*_{x,\mathcal{U}} \circ \vartheta_x. \end{aligned}$$

It follows from (2.31) that with the notation (7.3), one has

$$\kappa(\varkappa_{\nu^w}) = \varkappa_{\nu^{A^{-1}(c,\pi)w}}, w \in \mathbb{R}^m. \tag{8.21}$$

Notation 8.2 and (8.6) associate to $x \in \Omega(\mathcal{R})$ a linear space, $\Xi(x)$, of 1-cycles from $\tau(x)$ that projects onto $H^1(M(x),\mathbb{R})$ under the forgetful map, $M_n(x) \to M(x)$. It follows from (8.5) that a cycle $\gamma = \sum_{i=1}^m \xi_i\gamma_i(x)$ lies in the kernel of this map if, and only if, the vector ξ lies in the kernel of the matrix L^π,

$$L^\pi\xi = 0.$$

Therefore,

$$H_1(M(x),\mathbb{R}) \cong \mathbb{R}^m/\ker(L^\pi).$$

With respect to the Euclidean inner product, the dual of $\mathbb{R}^m/\ker(L^\pi)$ is identified with both $H^1(M(x),\mathbb{R})$ and $(\ker(L^\pi))^\perp$. The fact that L^π is alternating implies that $H(\pi)$ in (3.16) satisfies

$$H^1(M(x),\mathbb{R}) \cong (\ker(L^\pi))^\perp = H(\pi) = L^\pi\mathbb{R}^m. \tag{8.22}$$

PROPOSITION 8.5. *Suppose that*

$$\mathcal{U}x = \left(A\left(c,\pi\right)^{-1}\lambda, A^*\left(c,\pi\right)h, J(h,c)a, c\pi\right),\ x \in \Omega^*\left(\mathcal{R}\right).$$

The canonical map

$$F^*_{x,\mathcal{U}} : H^1\left(M(x),\mathbb{R}\right) \to H^1\left(M(\mathcal{U}x),\mathbb{R}\right) \tag{8.23}$$

may be realized as a map $H(\pi) \to H(c\pi)$*, given by*

$$h' \to A\left(c,\pi\right)^* h',\ h' \in H(\pi). \tag{8.24}$$

PROOF. Let $h' \in H(\pi)$, and select any $w \in \mathbb{R}^m$ such that $h' = L^\pi w$. Proposition 2.9 implies that w determines $\nu^w \in \mathcal{I}_{\mathcal{D}(T)}\left(T\right)$, which, in turn, determines $\varkappa_{\nu^w} \in \mathcal{Z}^1_h(x)$. Lemma 8.4 implies there exists a smooth 1-form, θ^1_ν, such that

$$\vartheta_x\left(\varkappa_{\nu^w}\right) = \widehat{\theta^1_\nu}$$

$$\vartheta_{\mathcal{U}x}\left(\kappa\left(\varkappa_{\nu^w}\right)\right) = \widehat{F^*_{x,\mathcal{U}}\theta^1_\nu}.$$

Next, (8.20) and (8.21) imply that

$$\Phi^*_{x,\mathcal{U}}\left(\varkappa_{\nu^w}\right) = \kappa\left(\varkappa_{\nu^w}\right) = \varkappa_{\nu^{A^{-1}(c,\pi)w}}, w \in \mathbb{R}^m.$$

Lemma 8.3 implies that if $\gamma\left(\xi\right) = \sum_{i=1}^{m} \xi_i \gamma_i\left(x\right), \xi \in \mathbb{R}^m$, then

$$\varkappa_{\nu^w}\left(\gamma\left(\xi\right)\right) = \left(L^\pi w\right)\cdot\xi, \xi \in \mathbb{R}^m.$$

Similarly,

$$\varkappa_{\nu^{A^{-1}(c,\pi)w}}\left(\gamma\left(\xi'\right)\right) = \left(L^{c\pi}A\left(c,\pi\right)^{-1}w\right)\cdot\xi', \xi' \in \mathbb{R}^m.$$

By [**R79**] or [**V78**]

$$L^{c\pi}A\left(c,\pi\right)^{-1} = A^*\left(c,\pi\right)L^\pi,$$

and (8.24) follows. □

The flow (7.12) may be represented as a measurable flow on

$$\Lambda^*_0\left(\mathcal{R}\right) \cong \bigcup_{\pi\in\mathcal{R}} \Omega^*_0\left(\pi\right) \times H(\pi).$$

Let $(x, h_0) = \left(\left(\lambda, h, a, \pi\right), h_0\right) \in \Omega^*_0\left(\pi\right) \times H(\pi)$ and let $t \geq 0$ (for convenience). While $Q^t x \in \Omega^*\left(\pi\right)$ (but not necessarily $Q^t x \in \Omega^*_0\left(\pi\right)$), there exists an integer $n\left(x,t\right) \geq 0$ such that

$$\mathcal{U}^{n(x,t)}Q^t x \in \Omega^*_0\left(\mathcal{R}\right)$$

Identifying $H\left(\pi\right)$, $\pi \in \mathcal{R}$, with $H^1\left(M(x),\mathbb{R}\right)$, we have from the definition and Proposition 8.5 that

$$\mathcal{V}^{n(x,t)}q^t\left(x,h_0\right) = \left(\mathcal{U}^{n(x,t)}Q^t x, F^*_{x,\mathcal{U}^{n(x,t)}}h_0\right)$$

$$= \left(\mathcal{U}^{n(x,t)}Q^t x, \left(A^{(n(x,t))}\right)^* h_0\right).$$

The first two coordinates of $\mathcal{U}^{n(x,t)}Q^t x$ are

$$\lambda^{(t)} = e^t\left(A^{(n(x,t))}\right)^{-1}\lambda = A\left(t,x\right)^{-1}\lambda$$

and

$$h^{(t)} = \left(A^{(n(x,t))}\right)^* h = A\left(t,x\right)^* h$$

where $A(t,x)$ is as in (5.3). Just as in (5.6), it is true that

$$\lim_{t\to\infty}\frac{n(x,t)}{t}=1.$$

Therefore, asymptotics for the zippered rectangles cocycle are the same as for the Rauzy cocycle on the spaces $H(\pi)$. The Lyapunov spectrum for the former cocycle is the same as for the Kontsevich-Zorich cocycle, as determined by Forni ([**F02**]) and Avila-Viana ([**AV07**]).

References

[AF07] A. Avila, G. Forni, *Weak mixing for interval exchange transformations and translation flows*, Ann. of Math. (2) **165** (2007), no. 2, 637–664. MR2299743 (2010c:37002)

[AV07] A. Avila, M. Viana, *Simplicity of Lyapunov spectra: proof of the Zorich-Kontsevich conjecture*, Acta Math. **198** (2007), no. 1, 1–56. MR2316268 (2008m:37010)

[Bu09] A. Bufetov, *Finitely additive measures on the asymptotic foliations of a Markov compactum*, arXiv.org, February 2009.

[Bu10-1] A. Bufetov, *Hölder cocycles and ergodic integrals for translation flows on flat surfaces*, Electron. Res. Announc. Math. Sci. **17** (2010), 34–42. MR2653543 (2011j:37008)

[Bu10-2] A. Bufetov, *Limit Theorems for Translation Flows*, arXiv.org, March 2010.

[Bu08] A. Bufetov, *Suspension flows over Vershik's automorphisms*, arXiv.org, April 2008.

[F02] G. Forni, *Deviation of ergodic averages for area-preserving flows on surfaces of higher genus*, Ann. of Math. (2) **155** (2002), no. 1, 1–103. MR1888794 (2003g:37009)

[FHZ04] S. Ferenczi, C. Holton, and L. Zamboni, *The structure of three-interval exchange transformations III: Ergodic and spectral properties*, J. Anal. Math. **93** (2004), 103–138. MR2110326 (2005h:37011)

[Kai87] V. A. Kaĭmanovich, *Lyapunov exponents, symmetric spaces and a multiplicative ergodic theorem for semisimple Lie groups* (Russian), Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklova (LOMI) 164 (1987), Differentsialnaya Geom. Gruppy Li i Mekh. IX, 29–46, 196–197; translation in J. Soviet Math. **47** (1989), no. 2, 2387–2398. MR947327 (89m:22006)

[Ka73] A. B. Katok, *Invariant measures of flows on orientable surfaces* (Russian), Dokl. Akad. Nauk SSSR **211** (1973), 775–778. MR0331438 (48:9771)

[KS67] A. B. Katok, A. M. Stepin, *Approximations in ergodic theory* (Russian), Uspehi Mat. Nauk **22** (1967), no. 5(137), 81–106. MR0219697 (36:2776)

[Ke75] M. Keane, *Interval exchange transformations*, Math. Z. **141** (1975), 25–31. MR0357739 (50:10207)

[Ko97] M. Kontsevich, *Lyapunov exponents and Hodge theory*, The mathematical beauty of physics (Saclay, 1996), 318–332, Adv. Ser. Math. Phys., 24, World Sci. Publ., River Edge, NJ, 1997. MR1490861 (99b:58147)

[M82] H. Masur, *Interval exchange transformations and measured foliations*, Ann. of Math. (2) **115** (1982), no. 1, 169–200. MR644018 (83e:28012)

[R79] G. Rauzy, *Échanges d'intervalles et transformations induites*, Acta Arith. **34** (1979), no. 4, 315–328. MR543205 (82m:10076)

[Sa75] E. A. Sataev, *The number of invariant measures for flows on orientable surfaces* (Russian), Izv. Akad. Nauk SSSR Ser. Mat. **39** (1975), no. 4, 860–878. MR0391184 (52:12005)

[S81] M. Stewart, *Irregularities of uniform distribution*, Acta Math. Acad. Sci. Hungar. **37** (1981), no. 1–3, 185–221. MR616890 (82k:10072)

[V08] W. A. Veech, *The Forni cocycle*, J. Mod. Dyn. **2** (2008), no. 3, 375–395. MR2417477 (2010a:37038)

[V82] W. A. Veech, *Gauss measures for transformations on the space of interval exchange maps*, Ann. of Math. (2) **115** (1982), no. 1, 201–242. MR644019 (83g:28036b)

[V78] W. A. Veech, *Interval exchange transformations*, J. Analyse Math. **33** (1978), 222–272. MR516048 (80e:28034)

[V81] W. A. Veech, *Projective Swiss cheeses and uniquely ergodic interval exchange transformations*, Ergodic theory and dynamical systems, I (College Park, Md., 1979–80), 113–193, Progr. Math. 10, Birkhäuser, Boston, Mass., 1981. MR633764 (83g:28036a)

[V69] W. A. Veech, *Strict ergodicity in zero dimensional dynamical systems and the Kronecker-Weyl theorem mod 2*, Trans. Amer. Math. Soc. **140** (1969), 1–33. MR0240056 (39:1410)

[V84-I] W. A. Veech, *The metric theory of interval exchange transformations. I. Generic spectral properties*, Amer. J. Math. **106** (1984), no. 6, 1331–1359. MR765582 (87j:28024a)

[V84-II] W. A. Veech, *The metric theory of interval exchange transformations. II. Approximation by primitive interval exchanges*, Amer. J. Math. **106** (1984), no. 6, 1361–1387. MR765583 (87j:28024b)

[V84-III] W. A. Veech, *The metric theory of interval exchange transformations. III. The Sah-Arnoux-Fathi invariant*, Amer. J. Math. **106** (1984), no. 6, 1389–1422. MR765584 (87j:28024c)

[V86] W. A. Veech, *The Teichmüller geodesic flow*, Ann. of Math. (2) **124** (1986), no. 3, 441–530. MR866707 (88g:58153)

[V90] W. A. Veech, *Moduli spaces of quadratic differentials*, J. Analyse Math. **55** (1990), 117–171. MR1094714 (92e:32014)

[Ver82] A. M. Vershik, *A theorem on Markov periodic approximation in ergodic theory* (Russian), Boundary value problems of mathematical physics and related questions in the theory of functions, 14. Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklova (LOMI) **115** (1982), 72–82. MR660072 (84m:28030)

[Z96] A. Zorich, *Finite Gauss measure on the space of interval exchange transformations. Lyapunov exponents*, Ann. Inst. Fourier (Grenoble) **46** (1996), no. 2, 325–370. MR1393518 (97f:58081)

Mathematics Department MS 136, Rice University, 6100 S. Main St., Houston, Texas 77005

E-mail address: veech@rice.edu

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11250

Notes on the Schreier graphs of the Grigorchuk group

Yaroslav Vorobets

Dedicated to Anatoli Stepin on the occasion of his 70th birthday

Abstract. The paper is concerned with the space of the marked Schreier graphs of the Grigorchuk group and the action of the group on this space. In particular, we describe the invariant set of the Schreier graphs corresponding to the action on the boundary of the binary rooted tree and dynamics of the group action restricted to this invariant set.

1. Introduction

This paper is devoted to the study of two equivalent dynamical systems of the Grigorchuk group $\mathcal{G}$, the action on the space of the marked Schreier graphs and the action on the space of subgroups. The main object of study is going to be the set of the marked Schreier graphs of the standard action of the group on the boundary of the binary rooted tree and their limit points in the space of all marked Schreier graphs of $\mathcal{G}$.

Given a finitely generated group G with a fixed generating set S, to each action of G we associate its Schreier graph, which is a combinatorial object that encodes some information about orbits of the action. The marked Schreier graphs of various actions form a topological space $\mathrm{Sch}(G,S)$ and there is a natural action of G on this space. Any action of G corresponds to an invariant set in $\mathrm{Sch}(G,S)$ and any action with an invariant measure gives rise to an invariant measure on $\mathrm{Sch}(G,S)$. The latter allows to define the notion of a random Schreier graph, which is closely related to the notion of a random subgroup of G.

A principal problem is to determine how much information about the original action can be learned from the Schreier graphs. The worst case here is a free action, for which nothing beyond its freeness can be recovered. Vershik [**5**] introduced the notion of a totally nonfree action. This is an action such that all points have distinct stabilizers. In this case the information about the original action can be recovered almost completely. Further extensive development of these ideas was done by Grigorchuk [**3**].

The Grigorchuk group was introduced in [**1**] as a simple example of a finitely generated infinite torsion group. Later it was revealed that this group has intermediate growth and a number of other remarkable properties (see the survey [**2**]). In this paper we are going to use the branching property of the Grigorchuk group,

2010 *Mathematics Subject Classification.* Primary 05C99, 22F05, 37A15, 37B05.

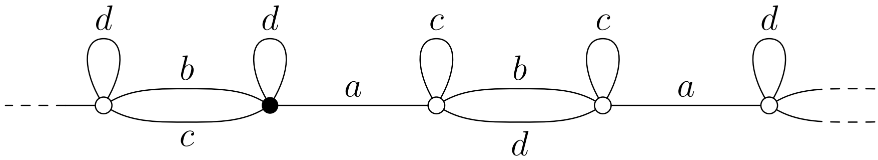

FIGURE 1. The marked Schreier graph of $0^\infty = 000\ldots$

which implies that its action on the boundary of the binary rooted tree is totally nonfree in a very strong sense.

The main results of the paper are summarized in the following two theorems. The first theorem contains a detailed description of the invariant set of the Schreier graphs. The second theorem is concerned with the dynamics of the group action restricted to that invariant set.

THEOREM 1.1. *Let $F : \partial\mathcal{T} \to \mathrm{Sch}(\mathcal{G}, \{a, b, c, d\})$ be the mapping that assigns to any point on the boundary of the binary rooted tree the marked Schreier graph of its orbit under the action of the Grigorchuk group. Then*

(i) *F is injective;*
(ii) *F is measurable; it is continuous everywhere except for a countable set, the orbit of the point $\xi_0 = 111\ldots$;*
(iii) *the Schreier graph $F(\xi_0)$ is an isolated point in the closure of $F(\partial\mathcal{T})$; the other isolated points are graphs obtained by varying the marked vertex of $F(\xi_0)$;*
(iv) *the closure of the set $F(\partial\mathcal{T})$ differs from $F(\partial\mathcal{T})$ in countably many points; these are obtained from three graphs $\Delta_0, \Delta_1, \Delta_2$ choosing the marked vertex arbitrarily;*
(v) *as an unmarked graph, $F(\xi_0)$ is a double quotient of each Δ_i $(i = 0, 1, 2)$; also, there exists a graph Δ such that each Δ_i is a double quotient of Δ.*

THEOREM 1.2. *Using notation of the previous theorem, let Ω be the set of non-isolated points of the closure of $F(\partial\mathcal{T})$. Then*

(i) *Ω is a minimal invariant set for the action of the Grigorchuk group $\mathcal{G}$ on $\mathrm{Sch}(\mathcal{G}, \{a, b, c, d\})$;*
(ii) *the action of $\mathcal{G}$ on Ω is a continuous extension of the action on the boundary of the binary rooted tree; the extension is one-to-one everywhere except for a countable set, where it is three-to-one;*
(iii) *there exists a unique Borel probability measure ν on $\mathrm{Sch}(\mathcal{G}, \{a, b, c, d\})$ invariant under the action of $\mathcal{G}$ and supported on the set Ω;*
(iv) *the action of $\mathcal{G}$ on Ω with the invariant measure ν is isomorphic to the action of $\mathcal{G}$ on $\partial\mathcal{T}$ with the uniform measure.*

The paper is organized as follows. Section 2 contains a detailed construction of the space of marked graphs. The construction is more general than that in [**3**]. Section 3 contains notation and definitions concerning group actions. In Section 4 we introduce the Schreier graphs of a finitely generated group, the space of marked Schreier graphs, and the action of the group on that space. In Section 5 we study the space of subgroups of a countable group and establish a relation of this space with the space of marked Schreier graphs. Section 6 is devoted to general considerations concerning groups of automorphisms of a regular rooted tree and their actions

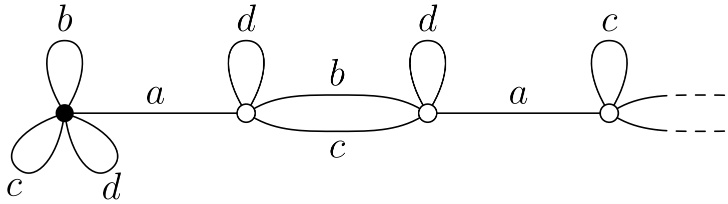

FIGURE 2. The marked Schreier graph of $1^\infty = 111\dots$

on the boundary of the tree. In Section 7 we apply the results of the previous sections to the study of the Grigorchuk group and prove Theorems 1.1 and 1.2. The exposition in Sections 2–6 is as general as possible, to make their results applicable to the actions of groups other than the Grigorchuk group.

The author is grateful to Rostislav Grigorchuk, Anatoly Vershik, and Oleg Ageev for useful discussions.

2. Space of marked graphs

A *graph* Γ is a combinatorial object that consists of *vertices* and *edges* related so that every edge joins two vertices or a vertex to itself (in the latter case the edge is called a *loop*). The vertices joined by an edge are its *endpoints*. Let V be the *vertex set* of the graph Γ and E be the set of its edges. Traditionally E is regarded as a subset of $V \times V$, i.e., any edge is identified with the pair of its endpoints. In this paper, however, we are going to consider graphs with multiple edges joining the same vertices. Also, our graphs will carry additional structure. To accomodate this, we regard E merely as a reference set whereas the actual information about the edges is contained in their *attributes*, which are functions on E. In a plain graph any edge has only one attribute: its endpoints, which are an unordered pair of vertices. Other types of graphs involve more attributes.

A *directed graph* has directed edges. The endpoints of a *directed edge* e are ordered, namely, there is a *beginning* $\alpha(e) \in V$ and an *end* $\omega(e) \in V$. Clearly, an undirected loop is no different from a directed one. An undirected edge joining two distinct vertices may be regarded as two directed edges e_1 and e_2 with the same endpoints and opposite directions, i.e., $\alpha(e_2) = \omega(e_1)$ and $\omega(e_2) = \alpha(e_1)$. This way we can represent any graph with undirected edges as a directed graph. Conversely, some directed graphs can be regarded as graphs with undirected edges (we shall use this in Section 7).

A *graph with labeled edges* is a graph in which each edge e is assigned a *label* $l(e)$. The labels are elements of a prescribed finite set. A *marked graph* is a graph with a distinguished vertex called the *marked vertex*.

The vertices of a graph are pictured as dots or small circles. An undirected edge is pictured as an arc joining its endpoints. A directed edge is pictured as an arrow going from its beginning to its end. The label of an edge is written next to the edge. Alternatively, one might think of labels as colors and picture a graph with labeled edges as a colored graph.

Let Γ be a graph and V be its vertex set. To any subset V' of V we associate a graph Γ' called a *subgraph* of Γ. By definition, the vertex set of the graph Γ' is V' and the edges are those edges of Γ that have both endpoints in V' (all attributes are retained). If Γ is a marked graph and the marked vertex is in V', it will also be the marked vertex of the subgraph Γ'. Otherwise the subgraph is not marked.

Suppose Γ_1 and Γ_2 are graphs of the same type. For any $i \in \{1,2\}$ let V_i be the vertex set of Γ_i and E_i be the set of its edges. The graph Γ_1 is said to be *isomorphic* to Γ_2 if there exist bijections $f : V_1 \to V_2$ and $\phi : E_1 \to E_2$ that respect the structure of the graphs. First of all, this means that f sends the endpoints of any edge $e \in E_1$ to the endpoints of $\phi(e)$. If Γ_1 and Γ_2 are directed graphs, we additionally require that $\alpha(\phi(e)) = f(\alpha(e))$ and $\omega(\phi(e)) = f(\omega(e))$ for all $e \in E_1$. If Γ_1 and Γ_2 have labeled edges, we also require that ϕ preserve labels. If Γ_1 and Γ_2 are marked graphs, we also require that f map the marked vertex of Γ_1 to the marked vertex of Γ_2. Assuming the above requirements are met, the mapping f of the vertex set is called an *isomorphism of the graphs* Γ_1 and Γ_2. If $\Gamma_1 = \Gamma_2$ then f is also called an *automorphism* of the graph Γ_1. We call the mapping ϕ a *companion mapping* of f. If neither of the graphs Γ_1 and Γ_2 admits multiple edges with identical attributes, the companion mapping is uniquely determined by the isomorphism f. Further, we say that the graph Γ_2 is a *quotient* of Γ_1 if all of the above requirements are met except the mappings f and ϕ need not be injective. Moreover, Γ_2 is a *k-fold quotient* of Γ_1 if f is k-to-1. Finally, we say that the graphs Γ_1 and Γ_2 coincide *up to renaming edges* if they have the same vertices and there is a one-to-one correspondence between their edges that preserves all attributes. An equivalent condition is that the identity map on the common vertex set is an isomorphism of these graphs.

A *path* in a graph Γ is a sequence of vertices $v_0, v_1, \dots, v_n$ together with a sequence of edges $e_1, \dots, e_n$ such that for any $1 \le i \le n$ the endpoints of the edge e_i are v_{i-1} and v_i. We say that the vertex v_0 is the beginning of the path and v_n is the end. The path is *closed* if $v_n = v_0$. The *length* of the path is the number of edges in the sequence (counted with repetitions), which is a nonnegative integer. The path is a *directed path* if the edges are directed and, moreover, $\alpha(e_i) = v_{i-1}$ and $\omega(e_i) = v_i$ for $1 \le i \le n$. If the graph Γ has labeled edges then the path is assigned a *code word* $l(e_1)l(e_2)\dots l(e_n)$, which is a string of labels read off the edges while traversing the path.

We say that a vertex v of a graph Γ is *connected to* a vertex v' if there is a path in Γ such that the beginning of the path is v and the end is v'. The length of the shortest path with this property is the *distance* from v to v'. The connectivity is an equivalence relation on the vertex set of Γ. The subgraphs of Γ corresponding to the equivalence classes are *connected components* of the graph Γ. A graph is *connected* if all vertices are connected to each other. Clearly, the connected components of any graph are its maximal connected subgraphs.

Let v be a vertex of a graph Γ. For any integer $n \ge 0$, the *closed ball* of radius n centered at v, denoted $\overline{B}_\Gamma(v,n)$, is the subgraph of Γ whose vertex set consists of all vertices in Γ at distance at most n from the vertex v. A graph is *locally finite* if every vertex is the endpoint for only finitely many edges. If the graph Γ is locally finite then any closed ball of Γ is a *finite graph*, i.e., it has a finite number of vertices and a finite number of edges.

Let $\mathcal{MG}$ denote the set of isomorphism classes of all marked directed graphs with labeled edges. For convenience, we regard elements of $\mathcal{MG}$ as graphs (i.e., we choose representatives of isomorphism classes). It is easy to observe that connectedness and local finiteness of graphs are preserved under isomorphisms. Let

$\mathcal{MG}_0$ denote the subset of $\mathcal{MG}$ consisting of connected, locally finite graphs. We endow the set $\mathcal{MG}_0$ with a topology as follows. The topology is generated by sets $\mathcal{U}(\Gamma_0, V_0) \subset \mathcal{MG}_0$, where Γ_0 runs over all finite graphs in $\mathcal{MG}_0$ and V_0 can be any subset of the vertex set of Γ_0. By definition, $\mathcal{U}(\Gamma_0, V_0)$ is the set of all isomorphism classes in $\mathcal{MG}_0$ containing any graph Γ such that Γ_0 is a subgraph of Γ and every edge of Γ with at least one endpoint in the set V_0 is actually an edge of Γ_0. In other words, there is no edge in Γ that joins a vertex from V_0 to a vertex outside the vertex set of Γ_0. For example, $\mathcal{U}(\Gamma_0, \emptyset)$ is the set of all graphs in $\mathcal{MG}_0$ that have a subgraph isomorphic to Γ_0. On the other hand, if V_0 is the entire vertex set of Γ_0 then $\mathcal{U}(\Gamma_0, V_0)$ contains only the graph Γ_0. As a consequence, every finite graph in $\mathcal{MG}_0$ is an isolated point. The following lemma implies that sets of the form $\mathcal{U}(\Gamma_0, V_0)$ constitute a base of the topology.

LEMMA 2.1. *Any nonempty intersection of two sets of the form $\mathcal{U}(\Gamma_0, V_0)$ can be represented as the union of some sets of the same form.*

PROOF. Let $\Gamma_1, \Gamma_2 \in \mathcal{MG}_0$ be finite graphs and V_1, V_2 be subsets of their vertex sets. Consider an arbitrary graph $\Gamma \in \mathcal{U}(\Gamma_1, V_1) \cap \mathcal{U}(\Gamma_2, V_2)$. For any $i \in \{1, 2\}$ let $f_i : W_i \to W_i'$ be an isomorphism of the graph Γ_i with a subgraph of Γ such that no edge of Γ joins a vertex from the set $f_i(V_i)$ to a vertex outside W_i'. Denote by Γ_0 the finite subgraph of Γ with the vertex set $W_0 = W_1' \cup W_2'$. Since the subgraphs of Γ with vertex sets W_1' and W_2' are both connected and both contain the marked vertex of Γ, the subgraph Γ_0 is also marked and connected. Besides, no edge of Γ joins a vertex from the set $V_0 = f_1(V_1) \cup f_2(V_2)$ to a vertex outside W_0. Hence $\Gamma \in \mathcal{U}(\Gamma_0, V_0)$. It is easy to observe that the entire set $\mathcal{U}(\Gamma_0, V_0)$ is contained in the intersection $\mathcal{U}(\Gamma_1, V_1) \cap \mathcal{U}(\Gamma_2, V_2)$. The lemma follows. $\square$

Next we introduce a distance function on $\mathcal{MG}_0$. Consider arbitrary graphs $\Gamma_1, \Gamma_2 \in \mathcal{MG}_0$. Let v_1 be the marked vertex of Γ_1 and v_2 be the marked vertex of Γ_2. We let $\delta(\Gamma_1, \Gamma_2) = 0$ if the graphs Γ_1 and Γ_2 are isomorphic (i.e., they represent the same element of $\mathcal{MG}_0$). Otherwise we let $\delta(\Gamma_1, \Gamma_2) = 2^{-n}$, where n is the smallest nonnegative integer such that the closed balls $\overline{B}_{\Gamma_1}(v_1, n)$ and $\overline{B}_{\Gamma_2}(v_2, n)$ are not isomorphic.

LEMMA 2.2. *The graphs Γ_1 and Γ_2 are isomorphic if and only if the closed balls $\overline{B}_{\Gamma_1}(v_1, n)$ and $\overline{B}_{\Gamma_2}(v_2, n)$ are isomorphic for any integer $n \geq 0$.*

PROOF. For any $i \in \{1, 2\}$ let V_i denote the vertex set of the graph Γ_i and E_i denote its set of edges. Further, for any integer $n \geq 0$ let $V_i(n)$ and $E_i(n)$ denote the vertex set and the set of edges of the closed ball $\overline{B}_{\Gamma_i}(v_i, n)$. First assume that the graph Γ_1 is isomorphic to Γ_2. Let $f : V_1 \to V_2$ be an isomorphism of these graphs and $\phi : E_1 \to E_2$ be its companion mapping. Clearly, $f(v_1) = v_2$. It is easy to see that any isomorphism of graphs preserves distances between vertices. It follows that f maps $V_1(n)$ onto $V_2(n)$ for any $n \geq 0$. Consequently, ϕ maps $E_1(n)$ onto $E_2(n)$. Hence the restriction of f to the set $V_1(n)$ is an isomorphism of the graphs $\overline{B}_{\Gamma_1}(v_1, n)$ and $\overline{B}_{\Gamma_2}(v_2, n)$.

Now assume that for every integer $n \geq 0$ the closed balls $\overline{B}_{\Gamma_1}(v_1, n)$ and $\overline{B}_{\Gamma_2}(v_2, n)$ are isomorphic. Let $f_n : V_1(n) \to V_2(n)$ be an isomorphism of these graphs and $\phi_n : E_1(n) \to E_2(n)$ be its companion mapping. Clearly, $f_n(v_1) = v_2$. Note that the closed ball $\overline{B}_{\Gamma_i}(v_i, n)$ is also the closed ball with the same center and radius in any of the graphs $\overline{B}_{\Gamma_i}(v_i, m)$, $m > n$. It follows that the restriction of the mapping f_m to the set $V_1(n)$ is an isomorphism of the graphs $\overline{B}_{\Gamma_1}(v_1, n)$ and $\overline{B}_{\Gamma_2}(v_2, n)$ while the restriction of ϕ_m to $E_1(n)$ is its companion mapping. Since the graphs Γ_1 and Γ_2 are locally finite, the sets $V_1(n), V_2(n), E_1(n), E_2(n)$ are finite. Hence there are only finitely many distinct restrictions $f_m|_{V_1(n)}$ or $\phi_m|_{E_1(n)}$ for any fixed n. Therefore one can find nested infinite sets of indices $I_0 \supset I_1 \supset I_2 \supset \dots$ such that the restriction $f_m|_{V_1(n)}$ is the same for all $m \in I_n$ and the restriction $\phi_m|_{E_1(n)}$ is the same for all $m \in I_n$. For any integer $n \geq 0$ let $f'_n = f_m|_{V_1(n)}$ and $\phi'_n = \phi_m|_{E_1(n)}$, where $m \in I_n$. By construction, f'_n is a restriction of f'_k and ϕ'_n is a restriction of ϕ'_k whenever $n < k$. Hence there exist maps $f : V_1 \to V_2$ and $\phi : E_1 \to E_2$ such that all f'_n are restrictions of f and all ϕ'_n are restrictions of ϕ. Since the graphs Γ_1 and Γ_2 are connected, any finite collection of vertices and edges in either graph is contained in a closed ball centered at the marked vertex. As for any $n \geq 0$ the mapping f'_n is an isomorphism of $\overline{B}_{\Gamma_1}(v_1, n)$ and $\overline{B}_{\Gamma_2}(v_2, n)$ and ϕ'_n is its companion mapping, it follows that f is an isomorphism of Γ_1 and Γ_2 and ϕ is its companion mapping. □

Lemma 2.2 implies that δ is a well-defined function on $\mathcal{MG}_0 \times \mathcal{MG}_0$. This is a distance function, which makes $\mathcal{MG}_0$ into an ultrametric space.

LEMMA 2.3. *The distance function δ is compatible with the topology on $\mathcal{MG}_0$.*

PROOF. The base of the topology on $\mathcal{MG}_0$ consists of the sets $\mathcal{U}(\Gamma_0, V_0)$. The base of the topology defined by the distance function δ is formed by open balls $\mathcal{B}(\Gamma_1, \epsilon) = \{\Gamma \in \mathcal{MG}_0 \mid \delta(\Gamma, \Gamma_1) < \epsilon\}$, where Γ_1 can be any graph in $\mathcal{MG}_0$ and $\epsilon > 0$. We have to show that any element of either base is the union of some elements of the other base.

First consider an open ball $\mathcal{B}(\Gamma_1, \epsilon)$. If $\epsilon > 1$ then $\mathcal{B}(\Gamma_1, \epsilon) = \mathcal{MG}_0$, which is the union of all sets $\mathcal{U}(\Gamma_0, V_0)$. Otherwise let n be the largest integer such that $\epsilon \leq 2^{-n}$. Clearly, $\mathcal{B}(\Gamma_1, \epsilon) = \mathcal{B}(\Gamma_1, 2^{-n})$. Let $\Gamma_0 = \overline{B}_{\Gamma_1}(v_1, n)$, where v_1 is the marked vertex of the graph Γ_1, and let V_0 be the set of all vertices of Γ_1 at distance at most $n-1$ from v_1. Consider an arbitrary graph $\Gamma \in \mathcal{MG}_0$ such that Γ_0 is a subgraph of Γ. Clearly, Γ_0 is also a subgraph of the closed ball $\overline{B}_{\Gamma}(v_1, n)$. If v is a vertex of Γ at distance k from the marked vertex v_1, then any vertex joined to v by an edge is at distance at most $k+1$ and at least $k-1$ from v_1. Moreover, if $k > 0$ then v is joined to a vertex at distance exactly $k-1$ from v_1. It follows that $\Gamma_0 = \overline{B}_{\Gamma}(v_1, n)$ if and only if no vertex from the set V_0 is joined in Γ to a vertex that is not a vertex of Γ_0. Thus $\mathcal{B}(\Gamma_1, 2^{-n}) = \mathcal{U}(\Gamma_0, V_0)$.

Now consider the set $\mathcal{U}(\Gamma_0, V_0)$, where Γ_0 is a finite graph in $\mathcal{MG}_0$ and V_0 is a subset of its vertex set. Denote by v_0 the marked vertex of Γ_0. Let n be the smallest integer such that every vertex of Γ_0 is at distance at most n from v_0 and every vertex from V_0 is at distance at most $n-1$ from v_0. Take any graph $\Gamma \in \mathcal{MG}_0$ such that Γ_0 is a subgraph of Γ and there is no edge in Γ joining a vertex from V_0 to a vertex outside the vertex set of Γ_0. Let $\Gamma_1 = \overline{B}_{\Gamma}(v_0, n)$ and V_1 be the set of all vertices of Γ at distance at most $n-1$ from v_0. By the above, $\mathcal{U}(\Gamma_1, V_1) = \mathcal{B}(\Gamma, 2^{-n})$. At the same time, $\mathcal{U}(\Gamma_1, V_1) \subset \mathcal{U}(\Gamma_0, V_0)$ since Γ_0 is a subgraph of Γ_1 and V_0 is a subset of V_1. Thus for any graph $\Gamma \in \mathcal{U}(\Gamma_0, V_0)$ the entire open ball $\mathcal{B}(\Gamma, 2^{-n})$ is contained in $\mathcal{U}(\Gamma_0, V_0)$. In particular, $\mathcal{U}(\Gamma_0, V_0)$ is the union of those open balls. □

Given a positive integer N and a finite set L, let $\mathcal{MG}(N, L)$ denote the subset of $\mathcal{MG}$ consisting of all graphs in which every vertex is the endpoint for at most N edges and every label belongs to L. Further, let $\mathcal{MG}_0(N, L) = \mathcal{MG}(N, L) \cap \mathcal{MG}_0$.

PROPOSITION 2.4. *$\mathcal{MG}_0(N,L)$ is a compact subset of the metric space $\mathcal{MG}_0$.*

PROOF. We have to show that any sequence of graphs $\Gamma_1, \Gamma_2, \dots$ in $\mathcal{MG}_0(N,L)$ has a subsequence converging to some graph in $\mathcal{MG}_0(N,L)$. For any positive integer n let V_n denote the vertex set of the graph Γ_n, E_n denote its sets of edges, and v_n denote the marked vertex of Γ_n. First consider the special case when each Γ_n is a subgraph of Γ_{n+1}. Let Γ be the graph with the vertex set $V = V_1 \cup V_2 \cup \dots$ and the set of edges $E = E_1 \cup E_2 \cup \dots$. We assume that any edge $e \in E_n$ retains its attributes (beginning, end, and label) in the graph Γ. The common marked vertex of the graphs Γ_n is set as the marked vertex of Γ. Note that any finite collection of vertices and edges of the graph Γ is already contained in some Γ_n. As the graphs $\Gamma_1, \Gamma_2, \dots$ belong to $\mathcal{MG}_0(N,L)$, it follows that $\Gamma \in \mathcal{MG}_0(N,L)$ as well. In particular, for any integer $k \geq 0$ the closed ball $\overline{B}_\Gamma(v_1,k)$ is a finite graph. Then it is a subgraph of some Γ_n. Clearly, $\overline{B}_\Gamma(v_1,k)$ is also a subgraph of the graphs $\Gamma_{n+1}, \Gamma_{n+2}, \dots$. Moreover, it remains the closed ball of radius k centered at the marked vertex in all these graphs. It follows that $\delta(\Gamma_m, \Gamma) < 2^{-k}$ for $m \geq n$. Since k can be arbitrarily large, the sequence $\Gamma_1, \Gamma_2, \dots$ converges to Γ in the metric space $\mathcal{MG}_0$.

Next consider a more general case when each Γ_n is isomorphic to a subgraph of Γ_{n+1}. This case is reduced to the previous one by repeatedly using the following observation: if a graph P_0 is isomorphic to a subgraph of a graph P then there exists a graph P' isomorphic to P such that P_0 is a subgraph of P'.

Finally consider the general case. For any graph in $\mathcal{MG}_0(N,L)$, the closed ball of radius k with any center contains at most $1+N+N^2+\dots+N^{k-1}$ vertices while the number of edges is at most N times the number of vertices. Hence for any fixed k the number of vertices and edges in the balls $\overline{B}_{\Gamma_n}(v_n,k)$ is uniformly bounded, which implies that there are only finitely many non-isomorphic graphs among them. Therefore one can find nested infinite sets of indices $I_0 \supset I_1 \supset I_2 \supset \dots$ such that the closed balls $\overline{B}_{\Gamma_n}(v_n,k)$ are isomorphic for all $n \in I_k$. Choose an increasing sequence of indices $n_0, n_1, n_2, \dots$ such that $n_k \in I_k$ for all k, and let Γ'_k be the closed ball of radius k in the graph Γ_{n_k} centered at the marked point v_{n_k}. Clearly, $\Gamma'_k \in \mathcal{MG}_0(N,L)$ and $\delta(\Gamma'_k, \Gamma_{n_k}) < 2^{-k}$. By construction, Γ'_k is isomorphic to a subgraph of Γ'_m whenever $k < m$. By the above the sequence $\Gamma'_0, \Gamma'_1, \Gamma'_2, \dots$ converges to a graph $\Gamma \in \mathcal{MG}_0(N,L)$. Since $\delta(\Gamma'_k, \Gamma_{n_k}) < 2^{-k}$ for all $k \geq 0$, the subsequence $\Gamma_{n_0}, \Gamma_{n_1}, \Gamma_{n_2}, \dots$ converges to the graph Γ as well. □

3. Group actions

Let M be an arbitrary nonempty set. Invertible transformations $\phi : M \to M$ form a transformation group. An *action* A of an abstract group G on the set M is a homomorphism of G into that transformation group. The action can be regarded as a collection of invertible transformations $A_g : M \to M$, $g \in G$, where A_g is the image of g under the homomorphism. The transformations are to satisfy $A_g A_h = A_{gh}$ for all $g, h \in G$. We say that A_g is the action of an element g within the action A. Alternatively, the action of the group G can be given as a mapping $A : G \times M \to M$ such that $A(g,x) = A_g(x)$ for all $g \in G$ and $x \in M$. Such a mapping defines an action of G if and only if the following two conditions hold:

- $A(gh,x) = A(g, A(h,x))$ for all $g, h \in G$ and $x \in M$;
- $A(1_G, x) = x$ for all $x \in M$, where 1_G is the unity of the group G.

A nonempty set $S \subset G$ is called a *generating set* for the group G if any element $g \in G$ can be represented as a product $g_1 g_2 \dots g_k$ where each factor g_i is an element of S or the inverse of an element of S. The elements of the generating set are called *generators* of the group G. The generating set S is *symmetric* if it is closed under taking inverses, i.e., $s^{-1} \in S$ whenever $s \in S$. If S is a generating set for G then any action A of the group G is uniquely determined by transformations A_s, $s \in S$.

Suppose G is a topological group. An action of G on a topological space M is a *continuous action* if it is continuous as a mapping of $G \times M$ to M. Similarly, an action of G on a measured space M is a *measurable action* if it is measurable as a mapping of $G \times M$ to M. A measurable action A of the group G on a measured space M with a measure μ is *measure-preserving* if the action of every element of G is measure-preserving, i.e., $\mu\bigl(A_g^{-1}(W)\bigr) = \mu(W)$ for all $g \in G$ and measurable sets $W \subset M$. In what follows, the group G will be a discrete countable group. In that case, an action A of G is continuous if and only if all transformations A_g, $g \in G$ are continuous. Likewise, the action A is measurable if and only if every A_g is measurable.

Given an action A of a group G on a set M, the *orbit* $O_A(x)$ of a point $x \in M$ under the action A is the set of all points $A_g(x)$, $g \in G$. A subset $M_0 \subset M$ is *invariant* under the action A if $A_g(M_0) \subset M_0$ for all $g \in G$. Clearly, the orbit $O_A(x)$ is invariant under the action. Moreover, this is the smallest invariant set containing x. The *restriction* of the action A to a nonempty invariant set M_0 is an action of G obtained by restricting every transformation A_g to M_0. Equivalently, one might restrict the mapping $A : G \times M \to M$ to the set $G \times M_0$. The action A is *transitive* if the only invariant subsets of M are the empty set and M itself. Equivalently, the orbit of any point is the entire set M. Assuming the action A is continuous, it is *topologically transitive* if there is an orbit dense in M, and *minimal* if every orbit of A is dense. The action is minimal if and only if the empty set and M are the only closed invariant subsets of M. Assuming the action A is measure-preserving, it is *ergodic* if any measurable invariant subset of M has zero or full measure. A continuous action on a compact space M is *uniquely ergodic* if there exists a unique Borel probability measure on M invariant under the action (the action is going to be ergodic with respect to that measure).

Given an action A of a group G on a set M, the *stabilizer* $\mathrm{St}_A(x)$ of a point $x \in M$ under the action A is the set of all elements $g \in G$ whose action fixes x, i.e., $A_g(x) = x$. The stabilizer $\mathrm{St}_A(x)$ is a subgroup of G. The action is *free* if all stabilizers are trivial. In the case when the action A is continuous, we define the *neighborhood stabilizer* $\mathrm{St}^o_A(x)$ of a point $x \in M$ to be the set of all $g \in G$ whose action fixes the point x along with its neighborhood (the neighborhood may depend on g). The neighborhood stabilizer $\mathrm{St}^o_A(x)$ is a normal subgroup of $\mathrm{St}_A(x)$.

Let $A : G \times M_1 \to M_1$ and $B : G \times M_2 \to M_2$ be actions of a group G on sets M_1 and M_2, respectively. The actions A and B are *conjugated* if there exists a bijection $f : M_1 \to M_2$ such that $B_g = f A_g f^{-1}$ for all $g \in G$. An equivalent condition is that $A(g, x) = B(g, f(x))$ for all $g \in G$ and $x \in M_1$. The bijection f is called a *conjugacy* of the action A with B. Two continuous actions of the same group are *continuously conjugated* if they are conjugated and, moreover, the conjugacy can be chosen to be a homeomorphism. Similarly, two measurable actions are *measurably conjugated* if they are conjugated and, moreover, the conjugacy f can be chosen so that both f and the inverse f^{-1} are measurable. Also, two measure-preserving actions are *isomorphic* if they are conjugated and, moreover, the conjugacy can be chosen to be an isomorphism of spaces with measure. The measure-preserving actions are *isomorphic modulo zero measure* if each action admits an invariant set of full measure such that the corresponding restrictions are isomorphic.

Given two actions $A : G \times M_1 \to M_1$ and $B : G \times M_2 \to M_2$ of a group G, the action A is an *extension* of B if there exists a mapping f of M_1 onto M_2 such that $B_g f = f A_g$ for all $g \in G$. The extension is k*-to-*1 if f is k-to-1. The extension is *continuous* if the actions A and B are continuous and f can be chosen continuous.

4. The Schreier graphs

Let G be a finitely generated group. Let us fix a finite symmetric generating set S for G. Given an action A of the group G on a set M, the *Schreier graph* $\Gamma_{\mathrm{Sch}}(G, S; A)$ of the action relative to the generating set S is a directed graph with labeled edges. The vertex set of the graph $\Gamma_{\mathrm{Sch}}(G, S; A)$ is M, the set of edges is $M \times S$, and the set of labels is S. For any $x \in M$ and $s \in S$ the edge (x, s) has beginning x, end $A_s(x)$, and carries label s. Clearly, the action A can be uniquely recovered from its Schreier graph. Given another action A' of G on some set M', the Schreier graph $\Gamma_{\mathrm{Sch}}(G, S; A')$ is isomorphic to $\Gamma_{\mathrm{Sch}}(G, S; A)$ if and only if the actions A and A' are conjugated. Indeed, a bijection $f : M \to M'$ is an isomorphism of the Schreier graphs if and only if $A'_s = f A_s f^{-1}$ for all $s \in S$, which is equivalent to f being a conjugacy of the action A with A'.

Any graph of the form $\Gamma_{\mathrm{Sch}}(G, S; A)$ is called a *Schreier graph* of the group G (relative to the generating set S). Notice that any graph isomorphic to a Schreier graph is also a Schreier graph up to renaming edges. This follows from the next proposition, which explains how to recognize a Schreier graph of G.

PROPOSITION 4.1. *A directed graph Γ with labeled edges is, up to renaming edges, a Schreier graph of the group G relative to the generating set S if and only if the following conditions are satisfied:*

(i) *all labels are in S;*
(ii) *for any vertex v and any generator $s \in S$ there exists a unique edge with beginning v and label s;*
(iii) *given a directed path with code word $s_1 s_2 \dots s_k$, the path is closed whenever the reversed code word $s_k \dots s_2 s_1$ equals 1_G when regarded as a product in G.*

PROOF. First suppose Γ is a Schreier graph $\Gamma_{\mathrm{Sch}}(G, S; A)$. Consider an arbitrary directed path in the graph Γ. Let v be the beginning of the path and $s_1 s_2 \dots s_k$ be its code word. Then the consecutive vertices of the path are $v_0 = v, v_1, \dots, v_k$, where $v_i = A_{s_i}(v_{i-1})$ for $1 \le i \le k$. Hence the end of the path is $A_{s_k} \dots A_{s_2} A_{s_1}(v) = A_g(v)$, where g denotes $s_k \dots s_2 s_1$ regarded as a product in G. Clearly, the path is closed whenever $g = 1_G$. Thus any Schreier graph of the group G satisfies the condition (iii). The conditions (i) and (ii) are trivially satisfied as well. It is easy to see that the conditions (i), (ii), and (iii) are preserved under isomorphisms of graphs. In particular, they hold for any graph that coincides with a Schreier graph up to renaming edges.

Now suppose Γ is a directed graph with labeled edges that satisfies the conditions (i), (ii), and (iii). Let M denote the vertex set of Γ. Given a word $w = s_1 s_2 \dots s_k$ over the alphabet S, we define a transformation $B_w : M \to M$

as follows. The condition (ii) implies that for any vertex $v \in M$ there is a unique directed path in Γ with beginning v and code word $s_k \dots s_2 s_1$ (the word w reversed). We set $B_w(v)$ to be the end of that path. For any words $w = s_1 s_2 \dots s_k$ and $w' = s'_1 s'_2 \dots s'_m$ over the alphabet S let ww' denote the concatenated word $s_1 s_2 \dots s_k s'_1 s'_2 \dots s'_m$. Then $B_{ww'}(v) = B_w(B_{w'}(v))$ for all $v \in M$. Any word over the alphabet S can be regarded as a product in the group G thus representing an element $g \in G$. Clearly, the concatenation of words corresponds to the multiplication in the group. The condition (iii) means that B_w is the identity transformation whenever the word w represents the unity 1_G. This implies that transformations B_w and $B_{w'}$ are the same if the words w and w' represent the same element $g \in G$. Indeed, let $w = s_1 s_2 \dots s_k$, $w' = s'_1 s'_2 \dots s'_m$ and consider the third word $z = s_k^{-1} \dots s_2^{-1} s_1^{-1}$. The word z represents the inverse g^{-1}. Therefore the words wz and zw' both represent the unity. Then $B_w = B_w B_{zw'} = B_{wzw'} = B_{wz} B_{w'} = B_{w'}$. Now for any $g \in G$ we let $A_g = B_w$, where w is an arbitrary word over the alphabet S representing g. By the above A_g is a well-defined transformation of M, A_{1_G} is the identity transformation, and $A_{gg'} = A_g A_{g'}$ for all $g, g' \in G$. Hence the transformations A_g, $g \in G$ constitute an action A of the group G on the vertex set M. By construction, for any $v \in M$ and $s \in S$ the vertex $A_s(v)$ is the end of the edge with beginning v and label s. In view of the conditions (i) and (ii), this means that the graph Γ coincides with the Schreier graph $\Gamma_{\mathrm{Sch}}(G, S; A)$ up to renaming edges. $\square$

For any $x \in M$ let $\Gamma_{\mathrm{Sch}}(G, S; A, x)$ denote the Schreier graph of the restriction of the action A to the orbit of x. We refer to $\Gamma_{\mathrm{Sch}}(G, S; A, x)$ as the Schreier graph of the orbit of x. It is easy to observe that $\Gamma_{\mathrm{Sch}}(G, S; A, x)$ is the connected component of the graph $\Gamma_{\mathrm{Sch}}(G, S; A)$ containing the vertex x. In particular, the Schreier graph of the action A is connected if and only if the action is transitive, in which case $\Gamma_{\mathrm{Sch}}(G, S; A, x) = \Gamma_{\mathrm{Sch}}(G, S; A)$ for all $x \in M$. Let $\Gamma^*_{\mathrm{Sch}}(G, S; A, x)$ denote a marked graph obtained from $\Gamma_{\mathrm{Sch}}(G, S; A, x)$ by marking the vertex x. We refer to it as the *marked Schreier graph* of the point x (under the action A). Notice that the point x and the restriction of the action A to its orbit are uniquely recovered from the graph $\Gamma^*_{\mathrm{Sch}}(G, S; A, x)$. Any graph of the form $\Gamma^*_{\mathrm{Sch}}(G, S; A, x)$ is called a marked Schreier graph of the group G (relative to the generating set S).

Let $\mathrm{Sch}(G, S)$ denote the set of isomorphism classes of all marked Schreier graphs of the group G relative to the generating set S. A graph $\Gamma \in \mathcal{MG}$ belongs to $\mathrm{Sch}(G, S)$ if it is a marked directed graph that is connected and satisfies conditions (i), (ii), (iii) of Proposition 4.1.

The group G acts naturally on the set of the marked Schreier graphs of G by changing the marked vertex. The action $\mathcal{A}$ is given by $\mathcal{A}_g\big(\Gamma^*_{\mathrm{Sch}}(G, S; A, x)\big) = \Gamma^*_{\mathrm{Sch}}(G, S; A, A_g(x))$, $g \in G$. It turns out that $\mathcal{A}$ is well defined as an action on $\mathrm{Sch}(G, S)$. Indeed, let $\Gamma^*_{\mathrm{Sch}}(G, S; B, y)$ be a marked Schreier graph isomorphic to $\Gamma^*_{\mathrm{Sch}}(G, S; A, x)$. Then any isomorphism f of the latter graph with the former one is simultaneously a conjugacy of the restriction of the action A to the orbit of x with the restriction of the action B to the orbit of y. Since $f(x) = y$, it follows that $f(A_g(x)) = B_g(y)$ for all $g \in G$. Hence for any $g \in G$ the map f is also an isomorphism of the graph $\Gamma^*_{\mathrm{Sch}}(G, S; A, A_g(x))$ with $\Gamma^*_{\mathrm{Sch}}(G, S; B, B_g(y))$.

PROPOSITION 4.2. *$\mathrm{Sch}(G, S)$ is a compact subset of the metric space $\mathcal{MG}_0$. The action of the group G (regarded as a discrete group) on $\mathrm{Sch}(G, S)$ is continuous.*

Proof. Let N be the number of elements in the generating set S. Then every vertex v of a graph $\Gamma \in \mathrm{Sch}(G,S)$ is the beginning of exactly N edges. Furthermore, v is the end of an edge with beginning v' and label s if and only if v' is the end of the edge with beginning v and label s^{-1}. It follows that v is also the end of exactly N edges. Hence any vertex of Γ is an endpoint for at most $2N$ edges. Therefore $\mathrm{Sch}(G,S) \subset \mathcal{MG}(2N,S)$. Since all marked Schreier graphs are connected, we have $\mathrm{Sch}(G,S) \subset \mathcal{MG}_0(2N,S) \subset \mathcal{MG}_0$.

Now let us show that the set $\mathrm{Sch}(G,S)$ is closed in the topological space $\mathcal{MG}_0$. Take any graph $\Gamma \in \mathcal{MG}_0$ not in that set. Then Γ does not satisfy at least one of the conditions (i), (ii), and (iii) in Proposition 4.1. First consider the case when the condition (i) or (iii) does not hold. Since the graph Γ is locally finite, it has a finite subgraph Γ_0 for which the same condition does not hold. Since Γ is connected, we can choose the subgraph Γ_0 to be marked and connected so that $\Gamma_0 \in \mathcal{MG}_0$. Clearly, the same condition does not hold for any graph Γ' such that Γ_0 is a subgraph of Γ'. It follows that the neighborhood $\mathcal{U}(\Gamma_0,\emptyset)$ of the graph Γ is disjoint from $\mathrm{Sch}(G,S)$. Next consider the case when Γ does not satisfy the condition (ii). Let v be the vertex of Γ such that for some generator $s \in S$ there are either several edges with beginning v and label s or no such edges at all. Since $\Gamma \in \mathcal{MG}_0$, there exists a finite connected subgraph Γ_0 of Γ that contains the marked vertex, the vertex v, and all edges for which v is an endpoint. Then $\Gamma_0 \in \mathcal{MG}_0$ and the open set $\mathcal{U}(\Gamma_0,\{v\})$ is a neighborhood of Γ. By construction, the condition (ii) fails in the entire neighborhood so that $\mathcal{U}(\Gamma_0,\{v\})$ is disjoint from $\mathrm{Sch}(G,S)$. Thus the set $\mathcal{MG}_0 \setminus \mathrm{Sch}(G,S)$ is open in $\mathcal{MG}_0$. Therefore the set $\mathrm{Sch}(G,S)$ is closed.

Since the closed set $\mathrm{Sch}(G,S)$ is contained in $\mathcal{MG}_0(2N,S)$, which is a compact set due to Proposition 2.4, the set $\mathrm{Sch}(G,S)$ is compact as well.

An action of the group G is continuous whenever the generators act continuously. To prove that the transformations $\mathcal{A}_s$, $s \in S$ are continuous, we are going to show that $\delta(\mathcal{A}_s(\Gamma),\mathcal{A}_s(\Gamma')) \leq 2\delta(\Gamma,\Gamma')$ for any graphs $\Gamma,\Gamma' \in \mathrm{Sch}(G,S)$ and any generator $s \in S$. If the graphs Γ and Γ' are isomorphic, then the graphs $\mathcal{A}_s(\Gamma)$ and $\mathcal{A}_s(\Gamma')$ are also isomorphic so that $\delta(\mathcal{A}_s(\Gamma),\mathcal{A}_s(\Gamma')) = \delta(\Gamma,\Gamma') = 0$. Otherwise $\delta(\Gamma,\Gamma') = 2^{-n}$ for some nonnegative integer n. Since the distance between any graphs in $\mathcal{MG}_0$ never exceeds 1, it is enough to consider the case $n \geq 2$. Let v denote the marked vertex of Γ and v' denote the marked vertex of Γ'. By definition of the distance function, the closed balls $\overline{B}_\Gamma(v,n-1)$ and $\overline{B}_{\Gamma'}(v',n-1)$ are isomorphic. Consider an isomorphism f of these graphs. Clearly, $f(v) = v'$. Let v_1 denote the marked vertex of the graph $\mathcal{A}_s(\Gamma)$ and v_1' denote the marked vertex of $\mathcal{A}_s(\Gamma')$. Then v_1 is the end of the edge with beginning v and label s in the graph Γ. Similarly, v_1' is the end of the edge with beginning v' and label s in Γ'. It follows that $f(v_1) = v_1'$. Since the vertex v_1 is joined to v by an edge, the closed ball $\overline{B}_\Gamma(v_1,n-2)$ is a subgraph of $\overline{B}_\Gamma(v,n-1)$. Note that $\overline{B}_\Gamma(v_1,n-2)$ remains the closed ball with the same center and radius in the graph $\overline{B}_\Gamma(v,n-1)$. Similarly, $\overline{B}_{\Gamma'}(v_1',n-2)$ is a subgraph of $\overline{B}_{\Gamma'}(v,n-1)$ and it is also the closed ball of radius $n-2$ centered at v_1' in the graph $\overline{B}_{\Gamma'}(v,n-1)$. Since $f(v_1) = v_1'$ and any isomorphism of graphs preserves distance between vertices, the restriction f_0 of f to the vertex set of $\overline{B}_\Gamma(v_1,n-2)$ is an isomorphisms of the graphs $\overline{B}_\Gamma(v_1,n-2)$ and $\overline{B}_{\Gamma'}(v_1',n-2)$. It remains to notice that the closed ball $\overline{B}_{\mathcal{A}_s(\Gamma)}(v_1,n-2)$ differs from $\overline{B}_\Gamma(v_1,n-2)$ in that the marked vertex is v_1 and, similarly, $\overline{B}_{\mathcal{A}_s(\Gamma')}(v_1',n-2)$ differs from $\overline{B}_{\Gamma'}(v_1',n-2)$ in that the marked vertex is v_1'. Therefore f_0 is also an isomorphism of $\overline{B}_{\mathcal{A}_s(\Gamma)}(v_1, n-2)$ and $\overline{B}_{\mathcal{A}_s(\Gamma')}(v'_1, n-2)$. By definition of the distance function, $\delta(\mathcal{A}_s(\Gamma), \mathcal{A}_s(\Gamma')) \le 2^{-(n-1)} = 2\delta(\Gamma, \Gamma')$. $\square$

Let A be an action of the group G on a set M. To any point $x \in M$ we associate three subgroups of G: the stabilizer $\mathrm{St}_A(x)$ of x, the stabilizer $\mathrm{St}_{\mathcal{A}}(\Gamma^*_x)$ of the marked Schreier graph $\Gamma^*_x = \Gamma^*_{\mathrm{Sch}}(G, S; A, x)$, and the neighborhood stabilizer $\mathrm{St}^o_{\mathcal{A}}(\Gamma^*_x)$ (if the action A is continuous then there is the fourth subgroup, the neighborhood stabilizer of x). Clearly, the graph $\Gamma^*_{A_g(x)}$ coincides with Γ^*_x if and only if $A_g(x) = x$. However this does not imply that the stabilizer of the graph is the same as the stabilizer of x. Since $\mathcal{A}$ is an action on isomorphism classes of graphs, we have $g \in \mathrm{St}_{\mathcal{A}}(\Gamma^*_x)$ if and only if the graph $\Gamma^*_{A_g(x)}$ is isomorphic to Γ^*_x.

Lemma 4.3. *(i)* $\mathrm{St}_A(x)$ *is a normal subgroup of* $\mathrm{St}_{\mathcal{A}}\big(\Gamma^*_{\mathrm{Sch}}(G, S; A, x)\big)$.

(ii) The quotient of $\mathrm{St}_{\mathcal{A}}\big(\Gamma^*_{\mathrm{Sch}}(G, S; A, x)\big)$ *by* $\mathrm{St}_A(x)$ *is isomorphic to the group of all automorphisms of the unmarked graph* $\Gamma_{\mathrm{Sch}}(G, S; A, x)$.

(iii) $\mathrm{St}_A(x)$ *is a subgroup of* $\mathrm{St}^o_{\mathcal{A}}\big(\Gamma^*_{\mathrm{Sch}}(G, S; A, x)\big)$.

Proof. Without loss of generality we can assume that the action A is transitive. For brevity, let Γ^* denote the marked graph $\Gamma^*_{\mathrm{Sch}}(G, S; A, x)$, Γ denote the unmarked graph $\Gamma_{\mathrm{Sch}}(G, S; A, x)$, and R denote the group of all automorphisms of Γ. Consider an arbitrary $f \in R$. For any vertex $y \in O_A(x)$ and any label $s \in S$ the unique edge of Γ with beginning y and label s has end $A_s(y)$. It follows that $f(A_s(y)) = A_s(f(y))$. Since the action A is transitive, the automorphism f commutes with transformations A_s, $s \in S$. Then f commutes with A_g for all $g \in G$. Notice that the automorphism f is uniquely determined by the vertex $f(x)$. Indeed, any vertex y of Γ is represented as $A_g(x)$ for some $g \in G$, then $f(y) = f(A_g(x)) = A_g(f(x))$. In particular, f is the identity if $f(x) = x$.

To prove the statements (i) and (ii), we are going to construct a homomorphism Ψ of the stabilizer $\mathrm{St}_{\mathcal{A}}(\Gamma^*)$ onto the group R with kernel $\mathrm{St}_A(x)$. An element $g \in G$ belongs to $\mathrm{St}_{\mathcal{A}}(\Gamma^*)$ if the graph Γ^* is isomorphic to $\Gamma^*_{\mathrm{Sch}}(G, S; A, A_g(x))$. An isomorphism of these marked graphs is an automorphism of the unmarked graph Γ that sends x to $A_g(x)$. Hence $g \in \mathrm{St}_{\mathcal{A}}(\Gamma^*)$ if and only if $A_g(x) = \psi_g(x)$ for some $\psi_g \in R$. By the above the automorphism ψ_g is uniquely determined by $A_g(x)$. Now we define a mapping $\Psi : \mathrm{St}_{\mathcal{A}}(\Gamma^*) \to R$ by $\Psi(g) = \psi_{g^{-1}}$. It is easy to observe that Ψ maps $\mathrm{St}_{\mathcal{A}}(\Gamma^*)$ onto R and the preimage of the identity under Ψ is $\mathrm{St}_A(x)$. Further, for any $g, h \in \mathrm{St}_{\mathcal{A}}(\Gamma^*)$ we have $\psi_{(gh)^{-1}}(x) = A^{-1}_{gh}(x) = A^{-1}_h(A^{-1}_g(x)) = A^{-1}_h(\psi_{g^{-1}}(x))$. Recall that the automorphism $\psi_{g^{-1}}$ commutes with the action A, in particular, $A^{-1}_h \psi_{g^{-1}} = \psi_{g^{-1}} A^{-1}_h$. Then $\psi_{(gh)^{-1}}(x) = \psi_{g^{-1}}(A^{-1}_h(x)) = \psi_{g^{-1}}(\psi_{h^{-1}}(x))$, which implies that $\Psi(gh) = \Psi(g)\Psi(h)$. Thus Ψ is a homomorphism.

We proceed to the statement (iii). Take any element $g \in \mathrm{St}_A(x)$. It can be represented as a product $s_1 s_2 \dots s_k$, where each s_i is in S. Let γ denote the unique directed path in Γ^* with beginning x and code word $s_k \dots s_2 s_1$. By construction, the end of the path γ is $A_g(x)$ so that the path is closed. Let Γ^*_0 denote the subgraph of Γ^* whose vertex set consists of all vertices of the path γ. Clearly, Γ^*_0 is a marked graph, finite and connected. Hence $\Gamma^*_0 \in \mathcal{MG}_0$. Any graph $\Gamma^*_1 \in \mathcal{U}(\Gamma^*_0, \emptyset)$ admits a closed directed path with beginning at the marked point and code word $s_k \dots s_2 s_1$. If $\Gamma^*_1 = \Gamma^*_{\mathrm{Sch}}(G, S; B, y)$, this implies that $B_g(y) = y$. Hence $g \in \mathrm{St}_B(y) \subset \mathrm{St}_{\mathcal{A}}\big(\Gamma^*_{\mathrm{Sch}}(G, S; B, y)\big)$. Thus the transformation $\mathcal{A}_g$ fixes the set $\mathcal{U}(\Gamma^*_0, \emptyset) \cap \mathrm{Sch}(G, S)$, which is an open neighborhood of the graph Γ^* in $\mathrm{Sch}(G, S)$. $\square$

Any group G acts naturally on itself by left multiplication. The action $\mathrm{adj}_G : G \times G \to G$, called *adjoint*, is given by $\mathrm{adj}_G(g_0, g) = g_0 g$. The Schreier graph of this action relative to any generating set S is the *Cayley graph* of the group G relative to S. Given a subgroup H of G, the adjoint action of the group G descends to an action on G/H. The action $\mathrm{adj}_{G,H} : G \times G/H \to G/H$ is given by $\mathrm{adj}_{G,H}(g_0, gH) = (g_0 g)H$. The Schreier graph of the latter action relative to a generating set S is denoted $\Gamma_{\mathrm{coset}}(G, S; H)$. It is called a *Schreier coset graph*. The *marked Schreier coset graph* $\Gamma^*_{\mathrm{coset}}(G, S; H)$ is the marked Schreier graph of the coset H under the action $\mathrm{adj}_{G,H}$. It is obtained from $\Gamma_{\mathrm{coset}}(G, S; H)$ by marking the vertex H.

Proposition 4.4. *A marked Schreier graph $\Gamma^*_{\mathrm{Sch}}(G, S; A, x)$ is isomorphic to a marked Schreier coset graph $\Gamma^*_{\mathrm{coset}}(G, S; H)$ if and only if $H = \mathrm{St}_A(x)$.*

Proof. Let H_0 denote the stabilizer $\mathrm{St}_A(x)$. Suppose $A_{g_1}(x) = A_{g_2}(x)$ for some $g_1, g_2 \in G$. Then $A_{g_2^{-1} g_1}(x) = A^{-1}_{g_2}(A_{g_1}(x)) = x$ so that $g_2^{-1} g_1 \in H_0$. Hence $g_2^{-1} g_1 H_0 = H_0$ and $g_1 H_0 = g_2 H_0$. Conversely, if $g_1 H_0 = g_2 H_0$ then $g_1 = g_2 h$ for some $h \in H_0$. It follows that $A_{g_1}(x) = A_{g_2}(A_h(x)) = A_{g_2}(x)$.

Let us define a mapping $f : G/H_0 \to O_A(x)$ by $f(gH_0) = A_g(x)$. By the above f is well defined and one-to-one. Clearly, it maps G/H_0 onto the entire orbit $O_A(x)$. For any $g_0, g \in G$ we have $f(g_0 g H_0) = A_{g_0 g}(x) = A_{g_0}(A_g(x)) = A_{g_0}(f(gH_0))$. Therefore f is a conjugacy of the action adj_{G,H_0} with the restriction of the action A to the orbit $O_A(x)$. It follows that f is also an isomorphism of the unmarked graphs $\Gamma_{\mathrm{coset}}(G, S; H_0)$ and $\Gamma_{\mathrm{Sch}}(G, S; A, x)$. As $f(H_0) = x$, the mapping f is an isomorphism of the marked graphs $\Gamma^*_{\mathrm{coset}}(G, S; H_0)$ and $\Gamma^*_{\mathrm{Sch}}(G, S; A, x)$ as well.

Since any isomorphism of Schreier graphs of the group G is also a conjugacy of the corresponding actions, it preserves stabilizers of vertices. In particular, marked Schreier graphs cannot be isomorphic if the stabilizers of their marked vertices do not coincide. For any subgroup H of G the stabilizer of the coset H under the action $\mathrm{adj}_{G,H}$ is H itself. Therefore the graph $\Gamma^*_{\mathrm{coset}}(G, S; H)$ is not isomorphic to $\Gamma^*_{\mathrm{Sch}}(G, S; A, x)$ if $H \neq \mathrm{St}_A(x)$. □

5. Space of subgroups

Let G be a discrete countable group. Denote by $\mathrm{Sub}(G)$ the set of all subgroups of G. We endow the set $\mathrm{Sub}(G)$ with a topology as follows. First we consider the product topology on $\{0,1\}^G$. The set $\{0,1\}^G$ is in a one-to-one correspondence with the set of all functions $f : G \to \{0,1\}$. Also, any subset $H \subset G$ (in particular, any subgroup) is assigned the indicator function $\chi_H : G \to \{0,1\}$ defined by

$$\chi_H(g) = \begin{cases} 1 \text{ if } g \in H, \\ 0 \text{ if } g \notin H. \end{cases}$$

This gives rise to a mapping $j : \mathrm{Sub}(G) \to \{0,1\}^G$, which is an embedding. Now the topology on $\mathrm{Sub}(G)$ is the smallest topology such that the embedding j is continuous. By definition, the base of this topology consists of sets of the form

$$U_G(S^+, S^-) = \{H \in \mathrm{Sub}(G) \mid S^+ \subset H \text{ and } S^- \cap H = \emptyset\},$$

where S^+ and S^- run independently over all finite subsets of G. Notice that $U_G(S_1^+, S_1^-) \cap U_G(S_2^+, S_2^-) = U_G(S_1^+ \cup S_2^+, S_1^- \cup S_2^-)$.

The topological space $\mathrm{Sub}(G)$ is ultrametric and compact (since $\{0,1\}^G$ is ultrametric and compact, and $j(\mathrm{Sub}(G))$ is closed in $\{0,1\}^G$). Suppose $g_1, g_2, g_3, \dots$ is a complete list of elements of the group G. For any subgroups $H_1, H_2 \subset G$ let $d(H_1, H_2) = 0$ if $H_1 = H_2$; otherwise let $d(H_1, H_2) = 2^{-n}$, where n is the smallest index such that g_n belongs to the symmetric difference of H_1 and H_2. Then d is a distance function on $\mathrm{Sub}(G)$ compatible with the topology.

Note that the above construction also applies to a finite group G, in which case $\mathrm{Sub}(G)$ is a finite set with the discrete topology.

The following three lemmas explore properties of the topological space $\mathrm{Sub}(G)$.

LEMMA 5.1. *The intersection of subgroups is a continuous operation on the space* $\mathrm{Sub}(G)$.

PROOF. We have to show that the mapping $I : \mathrm{Sub}(G) \times \mathrm{Sub}(G) \to \mathrm{Sub}(G)$ defined by $I(H_1, H_2) = H_1 \cap H_2$ is continuous. Take any finite sets $S^+, S^- \subset G$. Given subgroups $H_1, H_2 \subset G$, the intersection $H_1 \cap H_2$ is an element of the set $U_G(S^+, S^-)$ if and only if $H_1 \in U_G(S^+, S_1)$ and $H_2 \in U_G(S^+, S_2)$ for some sets S_1 and S_2 such that $S_1 \cup S_2 = S^-$. Clearly, the sets S_1 and S_2 are finite. It follows that

$$I^{-1}\big(U_G(S^+, S^-)\big) = \bigcup_{S_1, S_2 \,:\, S_1 \cup S_2 = S^-} U_G(S^+, S_1) \times U_G(S^+, S_2).$$

It remains to notice that any open subset of $\mathrm{Sub}(G)$ is a union of sets of the form $U_G(S^+, S^-)$ while any set of the form $U_G(S^+, S_1) \times U_G(S^+, S_2)$ is open in $\mathrm{Sub}(G) \times \mathrm{Sub}(G)$. $\square$

LEMMA 5.2. *For any subgroups H_1 and H_2 of the group G, let $H_1 \vee H_2$ denote the subgroup generated by all elements of H_1 and H_2. Then $\vee$ is a Borel measurable operation on* $\mathrm{Sub}(G)$.

PROOF. We have to show that the mapping $J : \mathrm{Sub}(G) \times \mathrm{Sub}(G) \to \mathrm{Sub}(G)$ defined by $J(H_1, H_2) = H_1 \vee H_2$ is Borel measurable. Take any $g \in G$ and consider arbitrary subgroups $H_1, H_2 \in \mathrm{Sub}(G)$ such that $J(H_1, H_2) \in U_G(\{g\}, \emptyset)$, i.e., $H_1 \vee H_2$ contains g. The element g can be represented as a product $g = h_1 h_2 \dots h_k$, where each h_i belongs to H_1 or H_2. Let S_1 denote the set of all elements of H_1 in the sequence $h_1, h_2, \dots, h_k$ and S_2 denote the set of all elements of H_2 in the same sequence. Then the element g belongs to $K_1 \vee K_2$ for any subgroups $K_1 \in U_G(S_1, \emptyset)$ and $K_2 \in U_G(S_2, \emptyset)$. Hence the pair (H_1, H_2) is contained in the preimage of $U_G(\{g\}, \emptyset)$ under the mapping J along with its open neighborhood $U_G(S_1, \emptyset) \times U_G(S_2, \emptyset)$. Thus the preimage $J^{-1}\big(U_G(\{g\}, \emptyset)\big)$ is an open set. Since the set $U_G(\emptyset, \{g\})$ is the complement of $U_G(\{g\}, \emptyset)$, its preimage under J is closed.

Given finite sets $S^+, S^- \subset G$, the set $U_G(S^+, S^-)$ is the intersection of sets $U_G(\{g\}, \emptyset)$, $g \in S^+$ and $U_G(\emptyset, \{h\})$, $h \in S^-$. By the above $J^{-1}\big(U_G(S^+, S^-)\big)$ is a Borel set, the intersection of an open set with a closed one. Finally, any open subset of $\mathrm{Sub}(G)$ is the union of some sets $U_G(S^+, S^-)$. Moreover, it is a finite or countable union since there are only countably many sets of the form $U_G(S^+, S^-)$. It follows that the preimage under J of any open set is a Borel set. $\square$

LEMMA 5.3. *Suppose H is a subgroup of G. Then $\mathrm{Sub}(H)$ is a closed subset of $\mathrm{Sub}(G)$. Moreover, the intrinsic topology on $\mathrm{Sub}(H)$ coincides with the topology induced by $\mathrm{Sub}(G)$.*

Proof. The intrinsic topology on $\mathrm{Sub}(H)$ is generated by all sets of the form $U_H(P^+,P^-)$, where P^+ and P^- are finite subsets of H. The topology induced by $\mathrm{Sub}(G)$ is generated by all sets of the form $U_G(S^+,S^-)\cap\mathrm{Sub}(H)$, where S^+ and S^- are finite subsets of G. Clearly, $U_G(S^+,S^-)\cap\mathrm{Sub}(H)=U_H(S^+,S^-\cap H)$ if $S^+\subset H$ and $U_G(S^+,S^-)\cap\mathrm{Sub}(H)=\emptyset$ otherwise. It follows that the two topologies coincide.

For any $g\in G$ the open set $U_G(\emptyset,\{g\})$ is also closed in $\mathrm{Sub}(G)$ as it is the complement of another open set $U_G(\{g\},\emptyset)$. Then the set $\mathrm{Sub}(H)$ is closed in $\mathrm{Sub}(G)$ since it is the intersection of closed sets $U_G(\emptyset,\{g\})$ over all $g\in G\setminus H$. □

Let A be an action of the group G on a set M. Let us consider the stabilizer $\mathrm{St}_A(x)$ of a point $x\in M$ under the action (see Section 3) as the value of a mapping $\mathrm{St}_A : M\to\mathrm{Sub}(G)$.

Lemma 5.4. *Suppose A is a continuous action of the group G on a Hausdorff topological space M. Then*

(i) *the mapping St_A is Borel measurable;*
(ii) *St_A is continuous at a point $x\in M$ if and only if the stabilizer of x under the action coincides with its neighborhood stabilizer: $\mathrm{St}^o_A(x)=\mathrm{St}_A(x)$;*
(iii) *if a sequence of points in M converges to the point x and the sequence of their stabilizers converges to a subgroup H, then $\mathrm{St}^o_A(x)\subset H\subset\mathrm{St}_A(x)$.*

Proof. For any $g\in G$ let $\mathrm{Fix}_A(g)$ denote the set of all points in M fixed by the transformation A_g. Let us show that $\mathrm{Fix}_A(g)$ is a closed set. Take any point $x\in M$ not in $\mathrm{Fix}_A(g)$. Since the points x and $A_g(x)$ are distinct, they have disjoint open neighborhoods X and Y, respectively. Since A_g is continuous, there exists an open neighborhood Z of x such that $A_g(Z)\subset Y$. Then $X\cap Z$ is an open neighborhood of x and $A_g(X\cap Z)$ is disjoint from $X\cap Z$. In particular, $X\cap Z$ is disjoint from $\mathrm{Fix}_A(g)$.

For any finite sets $S^+,S^-\subset G$ the preimage of the open set $U_G(S^+,S^-)$ under the mapping St_A is

$$\bigcap_{g\in S^+}\mathrm{Fix}_A(g)\setminus\bigcup_{h\in S^-}\mathrm{Fix}_A(h).$$

This is a Borel set as $\mathrm{Fix}_A(g)$ is closed for any $g\in G$. Since sets of the form $U_g(S^+,S^-)$ constitute a base of the topology on $\mathrm{Sub}(G)$, the mapping St_A is Borel measurable.

The mapping St_A is continuous at a point $x\in M$ if and only if x is an interior point in the preimage under St_A of any set $U_G(S^+,S^-)$ containing $\mathrm{St}_A(x)$. The latter holds true if and only if x is an interior point in any set $\mathrm{Fix}_A(g)$ containing this point. Clearly, x is an interior point of $\mathrm{Fix}_A(g)$ if and only if g belongs to the neighborhood stabilizer $\mathrm{St}^o_A(x)$. Thus St_A is continuous at x if and only if any element of $\mathrm{St}_A(x)$ belongs to $\mathrm{St}^o_A(x)$ as well.

Now suppose that a sequence $x_1,x_2,\dots$ of points in M converges to the point x and, moreover, the stabilizers $\mathrm{St}_A(x_1),\mathrm{St}_A(x_2),\dots$ converge to a subgroup H. Consider an arbitrary $g\in G$. In the case $g\in H$, the subgroup H belongs to the open set $U_G(\{g\},\emptyset)$. Since $\mathrm{St}_A(x_n)\to H$ as $n\to\infty$, we have $\mathrm{St}_A(x_n)\in U_G(\{g\},\emptyset)$ for large n. In other words, $x_n\in\mathrm{Fix}_A(g)$ for large n. Since the set $\mathrm{Fix}_A(g)$ is closed, it contains the limit point x as well. That is, $g\in\mathrm{St}_A(x)$. In the case $g\notin H$, the subgroup H belongs to the open set $U_G(\emptyset,\{g\})$. Then $\mathrm{St}_A(x_n)\in U_G(\emptyset,\{g\})$

for large n. In other words, $x_n \notin \mathrm{Fix}_A(g)$ for large n. Since $x_n \to x$ as $n \to \infty$, the action of g fixes no neighborhood of x. That is, $g \notin \mathrm{St}^o_A(x)$. □

The group G acts naturally on the set $\mathrm{Sub}(G)$ by conjugation. The action $\mathcal{C} : G \times \mathrm{Sub}(G) \to \mathrm{Sub}(G)$ is given by $\mathcal{C}(g, H) = gHg^{-1}$. This action is continuous. Indeed, one easily observes that $\mathcal{C}_g^{-1}\big(U_G(S_1, S_2)\big) = U_G(g^{-1}S_1g, g^{-1}S_2g)$ for all $g \in G$ and finite sets $S_1, S_2 \subset G$.

Proposition 5.5. *The action $\mathcal{C}$ of the group G on $\mathrm{Sub}(G)$ is continuously conjugated to the action $\mathcal{A}$ on the space $\mathrm{Sch}(G,S)$ of the marked Schreier graphs of G relative to a generating set S. Moreover, the mapping $f : \mathrm{Sub}(G) \to \mathrm{Sch}(G,S)$ given by $f(H) = \Gamma^*_{\mathrm{coset}}(G,S;H)$ is a continuous conjugacy.*

Proof. Proposition 4.4 implies that the mapping f is bijective.

Consider arbitrary element g and subgroup H of the group G. The stabilizer of the coset gH under the action $\mathrm{adj}_{G,H}$ consists of those $g_0 \in G$ for which $g_0gH = gH$. The latter condition is equivalent to $g^{-1}g_0g \in H$. Therefore the stabilizer is $gHg^{-1} = \mathcal{C}_g(H)$. As $\mathcal{A}_g\big(\Gamma^*_{\mathrm{coset}}(G,S;H)\big) = \Gamma^*_{\mathrm{Sch}}(G,S;\mathrm{adj}_{G,H}, gH)$, it follows from Proposition 4.4 that $\mathcal{A}_g(f(H)) = f(\mathcal{C}_g(H))$. Thus f conjugates the action $\mathcal{C}$ with $\mathcal{A}$.

Now we are going to show that for any finite sets $S^+, S^- \subset G$ the image of the open set $U_G(S^+, S^-)$ under the mapping f is open in $\mathrm{Sch}(G,S)$. Let $\Gamma^*_{\mathrm{Sch}}(G,S;A,x)$ be an arbitrary graph in that image. Any element $g \in G$ can be represented as a product $s_1s_2\dots s_k$, where $s_i \in S$. Let us fix such a representation and denote by γ_g the unique directed path in $\Gamma^*_{\mathrm{Sch}}(G,S;A,x)$ with beginning x and code word $s_k\dots s_2s_1$. Then the end of the path γ_g is $A_g(x)$. In particular, the path γ_g is closed if and only if $g \in \mathrm{St}_A(x)$. By Proposition 4.4, the preimage of the graph $\Gamma^*_{\mathrm{Sch}}(G,S;A,x)$ under f is $\mathrm{St}_A(x)$. Since $\mathrm{St}_A(x) \in U_G(S^+, S^-)$, the path γ_g is closed for $g \in S^+$ and not closed for $g \in S^-$. Let Γ_0 denote the smallest subgraph of $\Gamma^*_{\mathrm{Sch}}(G,S;A,x)$ containing all paths γ_g, $g \in S^+ \cup S^-$. Clearly, Γ_0 is a marked graph, finite and connected. Hence $\Gamma_0 \in \mathcal{MG}_0$. For any marked Schreier graph $\Gamma^*_{\mathrm{Sch}}(G,S;B,y)$ in $\mathcal{U}(\Gamma_0,\emptyset)$, the directed path with beginning y and the same code word as in γ_g is closed for all $g \in S^+$ and not closed for all $g \in S^-$. It follows that $\mathrm{St}_B(y) \in U_G(S^+, S^-)$. Therefore the graph $\Gamma^*_{\mathrm{Sch}}(G,S;A,x)$ is contained in $f(U_G(S^+, S^-))$ along with its neighborhood $\mathcal{U}(\Gamma_0,\emptyset) \cap \mathrm{Sch}(G,S)$.

Any open set in $\mathrm{Sub}(G)$ is the union of some sets $U_G(S^+, S^-)$. Hence it follows from the above that the mapping f maps open sets onto open sets. In other words, the inverse mapping f^{-1} is continuous. Since the topological spaces $\mathrm{Sub}(G)$ and $\mathrm{Sch}(G,S)$ are compact, f is continuous as well. □

Proposition 5.5 allows for a short (although not constructive) proof of the following statement.

Proposition 5.6. *Any subgroup of finite index of a finitely generated group is also finitely generated.*

Proof. Suppose G is a finitely generated group and H is a subgroup of G of finite index. Let S be a finite symmetric generating set for G. By Proposition 5.5, the space $\mathrm{Sub}(G)$ of subgroups of G is homeomorphic to the space $\mathrm{Sch}(G,S)$ of marked Schreier graphs of G relative to the generating set S. Moreover, there is a homeomorphism that maps the subgroup H to the marked Schreier coset graph $\Gamma^*_{\mathrm{coset}}(G,S;H)$. The vertices of the graph are cosets of H in G. Since H has finite index in G, the graph $\Gamma^*_{\mathrm{coset}}(G,S;H)$ is finite. Notice that any finite graph in the topological space $\mathcal{MG}_0$, which contains $\mathrm{Sch}(G,S)$, is an isolated point. It follows that H is an isolated point in $\mathrm{Sub}(G)$. Then there exist finite sets $S^+, S^- \subset G$ such that H is the only element of the open set $U_G(S^+,S^-)$. Let H_0 be the subgroup of G generated by the finite set S^+. Since $S^+ \subset H$ and $S^- \cap H = \emptyset$, the subgroup H_0 is disjoint from S^-. Thus $H_0 \in U_G(S^+,S^-)$ so that $H_0 = H$. $\square$

6. Automorphisms of regular rooted trees

Consider an arbitrary graph Γ. Let γ be a path in this graph, $v_0, v_1, \dots, v_m$ be consecutive vertices of γ, and $e_1, \dots, e_m$ be consecutive edges. A *backtracking* in the path γ occurs if $e_{i+1} = e_i$ for some i (then $v_{i+1} = v_{i-1}$). The graph Γ is called a *tree* if it is connected and admits no closed path of positive length without backtracking. In particular, this means no loops and no multiple edges. A *rooted tree* is a tree with a distinguished vertex called the *root*. Clearly, the root is a synonym for the marked vertex. For any integer $n \geq 0$ the *level* n (or the nth level) of the tree is defined as the set of vertices at distance n from the root. If $n \geq 1$ then any vertex on the nth level is joined to exactly one vertex on the level $n-1$ and, optionally, to some vertices on the level $n+1$. The rooted tree is called *k-regular* if every vertex on any level n is joined to exactly k vertices on level $n+1$. The 2-regular rooted tree is also called *binary*.

All k-regular rooted trees are isomorphic to each other. A standard model of such a tree is built as follows. Let X be a set of cardinality k referred to as the *alphabet* (usually $X = \{0, 1, \dots, k-1\}$). A *word* (or *finite word*) in the alphabet X is a finite string of elements from X (referred to as *letters*). The set of all words in the alphabet X is denoted X^*. X^* is a monoid with respect to the concatenation (the unit element is the empty word, denoted $\varnothing$). Moreover, it is the free monoid generated by elements of X. Now we define a plain graph $\mathcal{T}$ with the vertex set X^* in which two vertices w_1 and w_2 are joined by an edge if $w_1 = w_2 x$ or $w_2 = w_1 x$ for some $x \in X$. Then $\mathcal{T}$ is a k-regular rooted tree with the root $\varnothing$. The nth level of the tree $\mathcal{T}$ consists of all words of length n.

A bijection $f : X^* \to X^*$ is an automorphism of the rooted tree $\mathcal{T}$ if and only if it preserves the length of any word and the length of the common beginning of any two words. Given an automorphism f and a word $u \in X^*$, there exists a unique transformation $h : X^* \to X^*$ such that $f(uw) = f(u)h(w)$ for all $w \in X^*$. It is easy to see that h is also an automorphism of the tree $\mathcal{T}$. This automorphism is called the *section* of f at the word u and denoted $f|_u$. A set of automorphisms of the tree $\mathcal{T}$ is called *self-similar* if it is closed under taking sections. For any automorphisms f and h and any word $u \in X^*$ one has $(fh)|_u = f|_{h(u)} h|_u$ and $f^{-1}|_u = \big(f|_{f^{-1}(u)}\big)^{-1}$. It follows that any group of automorphisms generated by a self-similar set is itself self-similar.

Suppose G is a group of automorphisms of the tree $\mathcal{T}$. Let α denote the natural action of G on the vertex set X^*. Given a word $u \in X^*$, the section mapping $g \mapsto g|_u$ is a homomorphism when restricted to the stabilizer $\mathrm{St}_\alpha(u)$. If G is self-similar then this is a homomorphism to G. The self-similar group G is called *self-replicating* if for any $u \in X^*$ the mapping $g \mapsto g|_u$ maps the subgroup $\mathrm{St}_\alpha(u)$ onto the entire group G.

Suppose that letters of the alphabet X are canonically ordered: $x_1, x_2, \dots, x_k$. For any permutation π on X and automorphisms $h_1, h_2, \dots, h_k$ of the tree $\mathcal{T}$ we denote by $\pi(h_1, h_2, \ldots, h_k)$ a transformation $f : X^* \to X^*$ given by $f(x_i w) = \pi(x_i)h_i(w)$ for all $w \in X^*$ and $1 \le i \le k$. It is easy to observe that f is also an automorphism of $\mathcal{T}$ and $h_i = f|_{x_i}$ for $1 \le i \le k$. The expression $\pi(h_1, h_2, \ldots, h_k)$ is called the *wreath recursion* for f. Any self-similar set of automorphisms f_j, $j \in J$ satisfies a system of "self-similar" wreath recursions

$$f_j = \pi_j(f_{m(j,x_1)}, f_{m(j,x_2)}, \ldots, f_{m(j,x_k)}),\ j \in J,$$

where π_j, $j \in J$ are permutations on X and m maps $J \times X$ to J.

LEMMA 6.1. *Any system of self-similar wreath recursions over the alphabet X is satisfied by a unique self-similar set of automorphisms of the regular rooted tree $\mathcal{T}$.*

PROOF. Consider a system of wreath recursions $f_j = \pi_j(f_{m(j,x_1)}, \ldots, f_{m(j,x_k)})$, $j \in J$, where π_j, $j \in J$ are permutations on X and m is a mapping of $J \times X$ to J. We define transformations F_j, $j \in J$ of the set X^* inductively as follows. First $F_j(\varnothing) = \varnothing$ for all $j \in J$. Then, once the transformations are defined on words of a particular length $n \ge 0$, we let $F_j(x_i w) = \pi_j(x_i)F_{m(j,x_i)}(w)$ for all $j \in J$, $1 \le i \le k$, and words w of length n. By definition, each F_j preserves the length of words. Besides, it follows by induction on n that F_j is bijective when restricted to words of length n and that F_j preserves having a common beginning of length n for any two words. Therefore each F_j is an automorphism of the tree $\mathcal{T}$. By construction, the automorphisms F_j, $j \in J$ form a self-similar set satisfying the above system of wreath recursions. Moreover, they provide the only solution to that system. □

An *infinite path* in the tree $\mathcal{T}$ is an infinite sequence of vertices $v_0, v_1, v_2, \ldots$ together with a sequence of edges $e_1, e_2 \ldots$ such that the endpoints of any e_i are v_{i-1} and v_i. The vertex v_0 is the beginning of the path. Clearly, the path is uniquely determined by the sequence of vertices alone. The *boundary* of the rooted tree $\mathcal{T}$, denoted $\partial\mathcal{T}$, is the set of all infinite paths without backtracking that begin at the root. There is a natural one-to-one correspondence between $\partial\mathcal{T}$ and the set $X^{\mathbb{N}}$ of infinite words over the alphabet X. Namely, an infinite word $x_1x_2x_3\ldots$ corresponds to the path going through the vertices $\varnothing, x_1, x_1x_2, x_1x_2x_3, \ldots$. The set $X^{\mathbb{N}}$ is equipped with the product topology and the uniform Bernoulli measure. This allows us to regard the tree boundary $\partial\mathcal{T}$ as a compact topological space with a Borel probability measure (called *uniform*).

Suppose G is a group of automorphisms of the regular rooted tree $\mathcal{T}$. The natural action of G on the vertex set X^* gives rise to an action on the boundary $\partial\mathcal{T}$. The latter is continuous and preserves the uniform measure on $\partial\mathcal{T}$.

PROPOSITION 6.2 ([**3**]). *Let G be a countable group of automorphisms of a regular rooted tree $\mathcal{T}$. Then the following conditions are equivalent:*

(i) *the group G acts transitively on each level of the tree;*
(ii) *the action of G on the boundary $\partial\mathcal{T}$ of the tree is topologically transitive;*
(iii) *the action of G on $\partial\mathcal{T}$ is minimal;*
(iv) *the action of G on $\partial\mathcal{T}$ is ergodic with respect to the uniform measure;*
(v) *the action of G on $\partial\mathcal{T}$ is uniquely ergodic.*

Let G be a countable group of automorphisms of a regular rooted tree $\mathcal{T}$. Let α denote the natural action of G on the vertex set of the tree $\mathcal{T}$ and β denote the induced action of G on the boundary $\partial\mathcal{T}$ of the tree.

PROPOSITION 6.3. *The mapping* St_β *is continuous on a residual (dense* G_δ*) set.*

PROOF. For any $g \in G$ let $\mathrm{Fix}_\beta(g)$ denote the set of all points in $\partial\mathcal{T}$ fixed by the transformation β_g. If $g \in \mathrm{St}_\beta(\xi)$ but $g \notin \mathrm{St}^o_\beta(\xi)$, then ξ is a boundary point of the set $\mathrm{Fix}_\beta(g)$, and vice versa. Since $\mathrm{Fix}_\beta(g)$ is a closed set, its boundary is a closed, nowhere dense set. It follows that the set of points $\xi \in \partial\mathcal{T}$ such that $\mathrm{St}^o_\beta(\xi) = \mathrm{St}_\beta(\xi)$ is the intersection of countably many dense open sets (it is dense since $\partial\mathcal{T}$ is a complete metric space). By Lemma 5.4, the latter set consists of points at which the mapping St_β is continuous. □

The mapping St_β is Borel due to Lemma 5.4. If St_β is injective then, according to the descriptive set theory, it also maps Borel sets onto Borel sets (see, e.g., [**4**]). The following two lemmas show the same can hold under a little weaker condition.

LEMMA 6.4. *Assume that for any points* $\xi, \eta \in \partial\mathcal{T}$ *either* $\mathrm{St}_\beta(\xi) = \mathrm{St}_\beta(\eta)$ *or* $\mathrm{St}_\beta(\xi)$ *is not contained in* $\mathrm{St}_\beta(\eta)$*. Then the mapping* St_β *maps any open set, any closed set, and any intersection of an open set with a closed one onto Borel sets.*

PROOF. First let us show that St_β maps any closed subset C of the boundary $\partial\mathcal{T}$ onto a Borel subset of $\mathrm{Sub}(G)$. For any positive integer n let C_n denote the set of all words of length n in the alphabet X that are beginnings of infinite words in C. Further, let W_n be the union of sets $\mathrm{Sub}(\mathrm{St}_\alpha(w))$ over all words $w \in C_n$. Finally, let W be the intersection of the sets W_n over all $n \geq 1$. By Lemma 5.3, the set $\mathrm{Sub}(H)$ is closed in $\mathrm{Sub}(G)$ for any subgroup $H \in \mathrm{Sub}(G)$. Hence each W_n is closed as the union of finitely many closed sets. Then the intersection W is closed as well.

The stabilizer $\mathrm{St}_\beta(\xi)$ of an infinite word $\xi \in \partial\mathcal{T}$ is a subgroup of the stabilizer $\mathrm{St}_\alpha(w)$ of a finite word $w \in X^*$ whenever w is a beginning of ξ. It follows that $\mathrm{St}_\beta(\xi) \in W$ for all $\xi \in C$. By construction of the set W, any subgroup of an element of W is also an element of W. Hence W contains all subgroups of the groups $\mathrm{St}_\beta(\xi)$, $\xi \in C$.

Conversely, for any subgroup $H \in W$ there is a sequence of words $w_1, w_2, \dots$ such that $w_n \in C_n$ and $H \subset \mathrm{St}_\alpha(w_n)$ for $n = 1, 2, \dots$. Since the number of words of a fixed length is finite, one can find nested infinite sets of indices $I_1 \supset I_2 \supset \dots$ such that the beginning of length k of the word w_n is the same for all $n \in I_k$. Choose an increasing sequence of indices $n_1, n_2, \dots$ such that $n_k \in I_k$ for all k, and let w'_k be the beginning of length k of the word w_{n_k}. Then $w'_k \in C_k$ as $w_{n_k} \in C_{n_k}$. Besides, $\mathrm{St}_\alpha(w_{n_k}) \subset \mathrm{St}_\alpha(w'_k)$, in particular, the group H is a subgroup of $\mathrm{St}_\alpha(w'_k)$. By construction, the word w'_k is a beginning of w'_m whenever $k < m$. Therefore all w'_k are beginnings of the same infinite word $\xi' \in \partial\mathcal{T}$. Since every beginning of ξ' coincides with a beginning of some infinite word in C and the set C is closed, it follows that $\xi' \in C$. The stabilizer $\mathrm{St}_\beta(\xi')$ is the intersection of stabilizers $\mathrm{St}_\alpha(w'_k)$ over all $k \geq 1$. Hence H is a subgroup of $\mathrm{St}_\beta(\xi')$.

By the above a subgroup H of G belongs to the set W if and only if it is a subgroup of the stabilizer $\mathrm{St}_\beta(\xi)$ for some $\xi \in C$. The assumption of the lemma implies that stabilizers $\mathrm{St}_\beta(\xi)$, $\xi \in C$ can be distinguished as the maximal subgroups in the set W. That is, such a stabilizer is an element of W which is not a proper subgroup of another element of W. For any $g \in G$ we define a transformation ψ_g of $\mathrm{Sub}(G)$ by $\psi_g(H) = \langle g \rangle \vee H$, where $\langle g \rangle$ is a cyclic subgroup of G generated by g. The group $\psi_g(H)$ is generated by g and all elements of the group H. Clearly, a subgroup $H \in W$ is not maximal in W if and only if $\psi_g(H) \in W$ for some $g \notin H$. An equivalent condition is that H belongs to the set $W'_g = W \cap \psi_g^{-1}(W) \cap U_G(\emptyset, \{g\})$. It follows from Lemma 5.2 that the mapping ψ_g is Borel measurable. Therefore W'_g is a Borel set. Now the image of the set C under the mapping St_β is the difference of the closed set W and the union of Borel sets W'_g, $g \in G$. Hence this image is a Borel set.

Any open set $D \subset \partial\mathcal{T}$ is the union of a finite or countable collection of cylinders $Z_1, Z_2, \dots$, which are both open and closed sets. By the above each cylinder is mapped by St_β onto a Borel set in $\mathrm{Sub}(G)$. Then the union D is mapped onto the union of images of the cylinders, which is a Borel set as well. Further, for any closed set $C \subset \partial\mathcal{T}$ the intersection $C \cap D$ is the union of closed sets $C \cap Z_1, C \cap Z_2, \dots$. Hence it is also mapped by St_β onto a Borel set. □

LEMMA 6.5. *Under the assumption of Lemma* 6.4, *if the mapping* St_β *is finite-to-one, i.e., the preimage of any subgroup in* $\mathrm{Sub}(G)$ *is finite, then it maps Borel sets onto Borel sets.*

PROOF. Recall that the class $\mathfrak{B}$ of the Borel sets in $\partial\mathcal{T}$ is the smallest collection of subsets of $\partial\mathcal{T}$ that contains all closed sets and is closed under taking countable intersections, countable unions, and complements. Let $\mathfrak{U}$ denote the smallest collection of subsets of $\partial\mathcal{T}$ that contains all closed sets and is closed under taking countable intersections of nested sets and countable unions of any sets. Note that $\mathfrak{U}$ is well defined; it is the intersection of all collections satisfying these conditions. In particular, $\mathfrak{U} \subset \mathfrak{B}$. Further, let $\mathfrak{W}$ denote the collection of all Borel sets in $\partial\mathcal{T}$ mapped onto Borel sets in $\mathrm{Sub}(G)$ by the mapping St_β.

For any mapping $f : \partial\mathcal{T} \to \mathrm{Sub}(G)$ and any sequence $U_1, U_2, \dots$ of subsets of $\partial\mathcal{T}$ the image of the union $U_1 \cup U_2 \cup \dots$ under f is the union of images $f(U_1), f(U_2), \dots$. On the other hand, the image of the intersection $U_1 \cap U_2 \cap \dots$ under f is contained in $f(U_1) \cap f(U_2) \cap \dots$ but need not coincide with the latter when the mapping f is not one-to-one. The two sets do coincide if f is finite-to-one and $U_1 \supset U_2 \supset \dots$. Since the mapping St_β is assumed to be finite-to-one, it follows that the collection $\mathfrak{W}$ is closed under taking countable intersections of nested sets and countable unions of any sets. By Lemma 6.4, $\mathfrak{W}$ contains all closed sets. Therefore $\mathfrak{U} \subset \mathfrak{W}$.

To complete the proof, we are going to show that $\mathfrak{U} = \mathfrak{B}$, which will imply that $\mathfrak{W} = \mathfrak{B}$. Given a set $Y \in \mathfrak{U}$, let $\mathfrak{U}_Y$ denote the collection of all sets $U \in \mathfrak{U}$ such that the intersection $U \cap Y$ also belongs to $\mathfrak{U}$. For any sequence $U_1, U_2, \dots$ of elements of $\mathfrak{U}_Y$ we have

$$\Big(\bigcup\nolimits_{n\ge1} U_n\Big) \cap Y = \bigcup\nolimits_{n\ge1} (U_n \cap Y), \qquad \Big(\bigcap\nolimits_{n\ge1} U_n\Big) \cap Y = \bigcap\nolimits_{n\ge1} (U_n \cap Y).$$

Besides, the sets $U_1 \cap Y, U_2 \cap Y, \dots$ are nested whenever the sets $U_1, U_2, \dots$ are nested. It follows that the class $\mathfrak{U}_Y$ is closed under taking countable intersections of nested sets and countable unions of any sets. Consequently, $\mathfrak{U}_Y = \mathfrak{U}$ whenever $\mathfrak{U}_Y$ contains all closed sets. The latter condition obviously holds if the set Y is itself closed. Notice that for any sets $Y, Z \in \mathfrak{U}$ we have $Z \in \mathfrak{U}_Y$ if and only if $Y \in \mathfrak{U}_Z$. Since $\mathfrak{U}_Y = \mathfrak{U}$ for any closed set Y, it follows that $\mathfrak{U}_Z$ contains all closed sets for any $Z \in \mathfrak{U}$. Then $\mathfrak{U}_Z = \mathfrak{U}$ for any $Z \in \mathfrak{U}$. In other words, the class $\mathfrak{U}$ is closed under taking finite intersections. Combining finite intersections with countable intersections of nested sets, we can obtain any countable intersection of sets from $\mathfrak{U}$.

Namely, if $U_1, U_2, \dots$ are arbitrary elements of $\mathfrak{U}$, then their intersection coincides with the intersection of sets $Y_n = U_1 \cap U_2 \cap \dots \cap U_n$, $n = 1, 2, \dots$, which are nested: $Y_1 \supset Y_2 \supset \dots$. Therefore $\mathfrak{U}$ is closed under taking any countable intersections.

Let $\mathfrak{U}'$ be the collection of complements in $\partial\mathcal{T}$ of all sets from $\mathfrak{U}$. For any subsets $U_1, U_2, \dots$ of $\partial\mathcal{T}$ the complement of their union is the intersection of their complements $\partial\mathcal{T} \setminus U_1, \partial\mathcal{T} \setminus U_2, \dots$ while the complement of their intersection is the union of their complements. Since the class $\mathfrak{U}$ is closed under taking countable intersections and countable unions, so is $\mathfrak{U}'$. Further, any open subset of $\partial\mathcal{T}$ is the union of at most countably many cylinders, which are closed (as well as open) sets. Therefore $\mathfrak{U}$ contains all open sets. Then $\mathfrak{U}'$ contains all closed sets. Now it follows that $\mathfrak{U} \subset \mathfrak{U}'$. In other words, the class $\mathfrak{U}$ is closed under taking complements.

Thus the collection $\mathfrak{U}$ is closed under taking any countable intersections and complements. This implies that $\mathfrak{U} = \mathfrak{B}$. □

Let A be a continuous action of a countable group G on a compact metric space M. Let Ω denote the image of M under the mapping St_A.

LEMMA 6.6. *Assume that for any distinct points $x, y \in M$ the neighborhood stabilizer $\mathrm{St}^o_A(x)$ is not contained in the stabilizer $\mathrm{St}_A(y)$. Then the inverse of St_A, defined on the set Ω, can be extended to a continuous mapping of the closure of Ω onto M.*

PROOF. Since $\mathrm{St}^o_A(x)$ is a subgroup of $\mathrm{St}_A(x)$ for any $x \in M$, the assumption of the lemma implies that the mapping St_A is one-to-one so that the inverse is well defined on Ω. To prove that the inverse can be extended to a continuous mapping of the closure of Ω onto M, it is enough to show that any sequence $x_1, x_2, \dots$ of points in M is convergent whenever the sequence of stabilizers $\mathrm{St}_A(x_1), \mathrm{St}_A(x_2), \dots$ converges in $\mathrm{Sub}(G)$. Suppose that $\mathrm{St}_A(x_n) \to H$ as $n \to \infty$. Since M is a compact metric space, the sequence $x_1, x_2, \dots$ has at least one limit point. By Lemma 5.4(iii), any limit point x satisfies $\mathrm{St}^o_A(x) \subset H \subset \mathrm{St}_A(x)$. In particular, $\mathrm{St}^o_A(x) \subset \mathrm{St}_A(y)$ for any limit points x and y. Then $x = y$ due to the assumption of the lemma. It follows that the sequence $x_1, x_2, \dots$ is convergent. □

7. The Grigorchuk group

Let $X = \{0, 1\}$ be the binary alphabet, X^* be the set of finite words over X regarded as the vertex set of a binary rooted tree $\mathcal{T}$, and $X^{\mathbb{N}}$ be the set of infinite words over X regarded as the boundary $\partial\mathcal{T}$ of the tree $\mathcal{T}$.

We define the Grigorchuk group $\mathcal{G}$ as a self-similar group of automorphisms of the tree $\mathcal{T}$ (for alternative definitions, see [**2**]). The group is generated by four automorphisms a, b, c, d that, together with the trivial automorphism, form a self-similar set. Consider the following system of wreath recursions:

$$\begin{cases} a = (0\,1)(e, e), \\ b = (a, c), \\ c = (a, d), \\ d = (e, b), \\ e = (e, e). \end{cases}$$

By Lemma 6.1, this system uniquely defines a self-similar set of automorphisms of the tree $\mathcal{T}$. The automorphism e is clearly the identity (e.g., by Lemma 6.1). It is the unity of the group $\mathcal{G}$. We shall denote the unity by $1_{\mathcal{G}}$ to avoid confusion with a letter of the alphabet X. The set $S = \{a, b, c, d\}$ shall be considered the standard set of generators for the group $\mathcal{G}$.

All 4 generators of the Grigorchuk group are involutions. Indeed, the transformations $a^2, b^2, c^2, d^2, 1_{\mathcal{G}}$ form a self-similar set satisfying wreath recursions $a^2 = (1_{\mathcal{G}}, 1_{\mathcal{G}})$, $b^2 = (a^2, c^2)$, $c^2 = (a^2, d^2)$, $d^2 = (1_{\mathcal{G}}, b^2)$, and $1_{\mathcal{G}} = (1_{\mathcal{G}}, 1_{\mathcal{G}})$. Then Lemma 6.1 implies that $a^2 = b^2 = c^2 = d^2 = 1_{\mathcal{G}}$. This fact allows us to regard the Schreier graphs of the group $\mathcal{G}$ relative to the generating set S as graphs with undirected edges (as explained in Section 2).

Since $a^2 = 1_{\mathcal{G}}$, the automorphisms bcd, cdb, dbc, and $1_{\mathcal{G}}$ form a self-similar set satisfying wreath recursions $bcd = (1_{\mathcal{G}}, cdb)$, $cdb = (1_{\mathcal{G}}, dbc)$, $dbc = (1_{\mathcal{G}}, bcd)$, and $1_{\mathcal{G}} = (1_{\mathcal{G}}, 1_{\mathcal{G}})$. Lemma 6.1 implies that $bcd = cdb = dbc = 1_{\mathcal{G}}$. Then $bc = bcd^2 = d = d^2bc = bc$. It follows that $\{1_{\mathcal{G}}, b, c, d\}$ is a subgroup of $\mathcal{G}$ isomorphic to the Klein 4-group.

We denote by α the generic action of the group $\mathcal{G}$ on vertices of the binary rooted tree $\mathcal{T}$. The induced action on the boundary $\partial\mathcal{T}$ of the tree is denoted β. For brevity, we write $g(\xi)$ instead of $\beta_g(\xi)$. The action of the generator a is very simple: it changes the first letter in every finite or infinite word while keeping the other letters intact. In particular, the empty word is the only word fixed by a. To describe the action of the other generators, we need three observations. First of all, b, c, and d fix one-letter words. Secondly, any word beginning with 0 is fixed by d while b and c change only the second letter in such a word. Thirdly, the section mapping $g \mapsto g|_1$ induces a cyclic permutation on the set $\{b, c, d\}$. It follows that a finite or infinite word w is simultaneouly fixed by b, c, and d if it contains no zeros or the only zero is the last letter. Otherwise two of the three generators change the letter following the first zero in w (keeping the other letters intact) while the third generator fixes w. In the latter case, it is the position k of the first zero in w that determines the generator fixing w. Namely, $b(w) = w$ if $k \equiv 0 \bmod 3$, $c(w) = w$ if $k \equiv 2 \bmod 3$, and $d(w) = w$ if $k \equiv 1 \bmod 3$.

LEMMA 7.1. *The group $\mathcal{G}$ is self-replicating.*

PROOF. We have to show that for any word $w \in X^*$ the section mapping $g \mapsto g|_w$ maps the stabilizer $\mathrm{St}_\alpha(w)$ onto the entire group $\mathcal{G}$. Let W be the set of all words with this property. Clearly, $\varnothing \in W$ as $\mathrm{St}_\alpha(\varnothing) = \mathcal{G}$ and $g|_\varnothing = g$ for all $g \in \mathcal{G}$. Suppose $w_1, w_2 \in W$. Given an arbitrary $g \in \mathcal{G}$, there exists $g' \in \mathcal{G}$ such that $g'(w_2) = w_2$ and $g'|_{w_2} = g$. Further, there exists $g'' \in \mathcal{G}$ such that $g''(w_1) = w_1$ and $g''|_{w_1} = g'$. Then $g''(w_1w_2) = g''(w_1)g''|_{w_1}(w_2) = w_1w_2$ and $g''|_{w_1w_2} = (g''|_{w_1})|_{w_2} = g$. Since g is arbitrary, $w_1w_2 \in W$. That is, the set W is closed under concatenation.

Any automorphism of the tree $\mathcal{T}$ either interchanges the vertices 0 and 1 or fixes them both. Hence the stabilizer $\mathrm{St}_\alpha(0)$ coincides with $\mathrm{St}_\alpha(1)$. This stabilizer contains the elements b, c, d, aba, aca, ada. The wreath recursions for these elements are $b = (a, c)$, $c = (a, d)$, $d = (1_{\mathcal{G}}, b)$, $aba = (c, a)$, $aca = (d, a)$, $ada = (b, 1_{\mathcal{G}})$. It follows that the images of the group $\mathrm{St}_\alpha(0)$ under the section mappings $g \mapsto g|_0$ and $g \mapsto g|_1$ contain the generating set S. As the restrictions of these mappings to $\mathrm{St}_\alpha(0)$ are homomorphisms, both images coincide with $\mathcal{G}$. Therefore the words 0 and 1 are in the set W. By the above W is closed under concatenation and contains the empty word. This implies $W = X^*$. $\square$

The orbits of the actions α and β are very easy to describe.

LEMMA 7.2. *The group $\mathcal{G}$ acts transitively on each level of the binary rooted tree $\mathcal{T}$. Any two infinite words in $\partial\mathcal{T}$ are in the same orbit of the action β if and only if they differ in only finitely many letters.*

PROOF. For any infinite word $\xi \in \partial\mathcal{T}$ and any generator $h \in \{a, b, c, d\}$ the infinite word $h(\xi)$ differs from ξ in at most one letter. Any $g \in \mathcal{G}$ can be represented as a product $g = h_1 h_2 \dots h_k$, where each h_i is in $\{a, b, c, d\}$. It follows that for any $\xi \in \partial\mathcal{T}$ the infinite words $g(\xi)$ and ξ differ in at most k letters. Thus any two infinite words in the same orbit of the action β differ in only finitely many letters.

Now we are going to show that for any finite words $w_1, w_2 \in X^*$ of the same length there exists $g \in \mathcal{G}$ such that $g(w_1) = w_2$ and $g|_{w_1} = 1_{\mathcal{G}}$. Equivalently, $g(w_1\xi) = w_2\xi$ for all $\xi \in \partial\mathcal{T}$. This will complete the proof of the lemma. Indeed, the claim contains the statement that the group $\mathcal{G}$ acts transitively on each level of the tree $\mathcal{T}$. Moreover, it implies that two infinite words in $\partial\mathcal{T}$ are in the same orbit of the action β whenever they differ in a finite number of letters.

We prove the claim by induction on the length n of the words w_1 and w_2. The case $n = 0$ is trivial. Here w_1 and w_2 are the empty words so that we take $g = 1_{\mathcal{G}}$. Now assume that the claim is true for all pairs of words of specific length $n \geq 0$ and consider words w_1 and w_2 of length $n+1$. Let x_1 be the first letter of w_1 and x_2 be the first letter of w_2. Then $w_1 = x_1u_1$ and $w_2 = x_2u_2$, where u_1 and u_2 are words of length n. By the inductive assumption, there exists $h \in \mathcal{G}$ such that $h(u_1\xi) = u_2\xi$ for all $\xi \in \partial\mathcal{T}$. Since the group $\mathcal{G}$ is self-replicating, there exists $g_0 \in \mathcal{G}$ such that $g_0(x_1\eta) = x_1h(\eta)$ for all $\eta \in \partial\mathcal{T}$. In particular, $g_0(x_1u_1\xi) = x_1u_2\xi$ for all $\xi \in \partial\mathcal{T}$. It remains to take $g = g_0$ if $x_2 = x_1$ and $g = ag_0$ otherwise. Then $g(x_1u_1\xi) = x_2u_2\xi$ for all $\xi \in \partial\mathcal{T}$. □

LEMMA 7.3. *Suppose w_1 and w_2 are words in the alphabet $\{0, 1\}$ such that w_1 is not a beginning of w_2 while w_2, even with the last two letters deleted, is not a beginning of w_1. Then there exists $g \in \mathcal{G}$ that does not fix w_2 while fixing all words with beginning w_1.*

PROOF. First we consider a special case when $w_2 = 100$. To satisfy the assumption of the lemma, the word w_1 has to begin with 0. Then we can take $g = d$. Indeed, the transformation d fixes all words that begin with 0, which includes all words with beginning w_1. At the same time, $d(100) = 1b(00) = 10a(0) = 101 \neq 100$.

Next we consider a slightly more general case when w_2 is an arbitrary word of length 3. By Lemma 7.2, the group $\mathcal{G}$ acts transitively on the third level of the tree $\mathcal{T}$. Therefore $h(w_2) = 100$ for some $h \in \mathcal{G}$. The words $h(w_1)$ and $h(w_2)$ satisfy the assumption of the lemma since the words w_1 and w_2 do. By the above, $dh(w_2) \neq h(w_2)$ while $d(h(w_1)u) = h(w_1)u$ for all $u \in X^*$. Let $g = h^{-1}dh$. Then $g(w_2) \neq w_2$ while $g(w_1w) = w_1w$ for all $w \in X^*$.

Finally, consider the general case. Let w_0 be the longest common beginning of the words w_1 and w_2. Then $w_1 = w_0u_1$ and $w_2 = w_0u_2$, where the words u_1 and u_2 also satisfy the assumption of the lemma. In particular, u_1 is nonempty and the length of u_2 is at least 3. We have $u_2 = u_2'u_2''$, where $u_2', u_2'' \in X^*$ and the length of u_2' is 3. Since the first letters of the words u_1 and u_2' are distinct, these words satisfy the assumption of the lemma. By the above there exists $g_0 \in \mathcal{G}$ such that $g_0(u_2') \neq u_2'$ and $g_0(u_1u) = u_1u$ for all $u \in X^*$. Since the group $\mathcal{G}$ is self-replicating, there exists $g \in \mathcal{G}$ such that $g(w_0w) = w_0g_0(w)$ for all $w \in X^*$. Then g does not fix the word w_0u_2' while fixing all words with beginning w_1. Since w_0u_2' is a beginning of w_2, the transformation g does not fix w_2 as well. $\square$

LEMMA 7.4. *For any distinct points $\xi, \eta \in \partial\mathcal{T}$ the neighborhood stabilizer $\mathrm{St}^o_\beta(\xi)$ is not contained in $\mathrm{St}_\beta(\eta)$.*

PROOF. Let n denote the length of the longest common beginning of the distinct infinite words ξ and η. Let w_1 be the beginning of ξ of length $n+1$ and w_2 be the beginning of η of length $n+3$. It is easy to see that the words w_1 and w_2 satisfy the assumption of Lemma 7.3. Therefore there exists a transformation $g \in \mathcal{G}$ that does not fix w_2 while fixing all finite words with beginning w_1. Clearly, the action of g on $\partial\mathcal{T}$ fixes all infinite words with beginning w_1. As such infinite words form an open neighborhood of the point ξ, we have $g \in \mathrm{St}^o_\beta(\xi)$. At the same time, g does not fix the infinite word η since it does not fix its beginning w_2. Hence $g \notin \mathrm{St}_\beta(\eta)$ so that $\mathrm{St}^o_\beta(\xi) \not\subset \mathrm{St}_\beta(\eta)$. $\square$

LEMMA 7.5. $\mathrm{St}^o_\beta(\xi) = \mathrm{St}_\beta(\xi)$ *for any infinite word $\xi \in \partial\mathcal{T}$ containing infinitely many zeros.*

PROOF. We are going to show that, given an automorphism $g \in \mathcal{G}$ and an infinite word $\xi \in \partial\mathcal{T}$ with infinitely many zeros, one has $g|_w = 1_\mathcal{G}$ for a sufficiently long beginning w of ξ. This claim implies the lemma. Indeed, in the case $g(\xi) = \xi$ the action of g fixes all infinite words with beginning w, which form an open neighborhood of ξ.

Let R be the set of all $g \in \mathcal{G}$ such that the claim holds true for g and any $\xi \in \partial\mathcal{T}$ with infinitely many zeros. The set R contains the generating set S. Indeed, $a|_w = 1_\mathcal{G}$ for any nonempty word $w \in X^*$ and $b|_w = c|_w = d|_w = 1_\mathcal{G}$ for any word w that contains a zero which is not the last letter of w. Now suppose $g, h \in R$ and consider an arbitrary $\xi \in \partial\mathcal{T}$ with infinitely many zeros. Then $h|_w = 1_\mathcal{G}$ for a sufficiently long beginning w of ξ. Lemma 7.2 implies that the infinite word $h(\xi)$ also has infinitely many zeros. Since $h(w)$ is a beginning of $h(\xi)$ and $g \in R$, we have $g|_{h(w)} = 1_\mathcal{G}$ provided w is long enough. Since $(gh)|_w = g|_{h(w)}h|_w$, we have $(gh)|_w = 1_\mathcal{G}$ provided w is long enough. Thus $gh \in R$. That is, the set R is closed under multiplication. Since $S \subset R$ and all generators are involutions, it follows that $R = \mathcal{G}$. $\square$

The infinite word $\xi_0 = 111\dots$ (also denoted 1^∞) is an exceptional point for the action β.

LEMMA 7.6. *The quotient of $\mathrm{St}_\beta(\xi_0)$ by $\mathrm{St}^o_\beta(\xi_0)$ is the Klein 4-group. The coset representatives are $1_\mathcal{G}, b, c, d$.*

PROOF. Recall that $H = \{1_\mathcal{G}, b, c, d\}$ is a subgroup of $\mathcal{G}$ isomorphic to the Klein 4-group. Clearly, $H \subset \mathrm{St}_\beta(\xi_0)$. We are going to show that $H \cap \mathrm{St}^o_\beta(\xi_0) = \{1_\mathcal{G}\}$ and $\mathrm{St}_\beta(\xi_0) = \mathrm{St}^o_\beta(\xi_0)H$, which implies the lemma.

For any positive integer n let η_n denote the infinite word over the alphabet X that has a single zero in the position n. The sequence $\eta_1, \eta_2, \dots$ converges to ξ_0. One observes that any of the generators b, c, and d fixes η_n only if n leaves a specific remainder under division by 3 (0 for b, 2 for c, and 1 for d). It follows that $H \cap \mathrm{St}^o_\beta(\xi_0) = \{1_\mathcal{G}\}$.

Now let us show that any $g \in \mathrm{St}_\beta(\xi_0)$ is contained in the set $\mathrm{St}^o_\beta(\xi_0)H$. The proof is by strong induction on the length n of g, which is the smallest possible

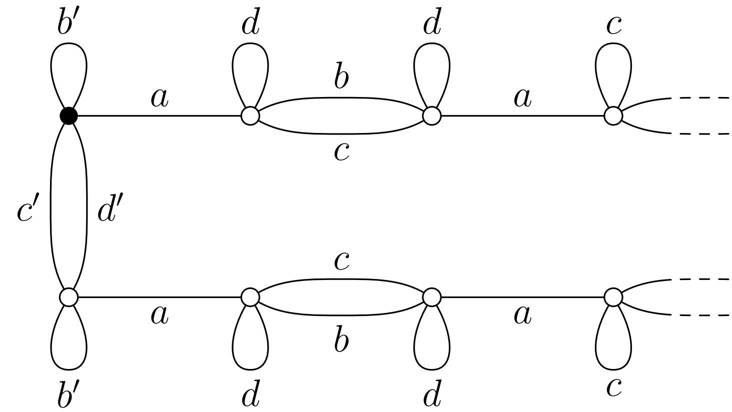

FIGURE 3. Limit graphs $\Delta_0^*, \Delta_1^*, \Delta_2^*$.

number of factors in an expansion $g = s_m \dots s_2 s_1$ such that each $s_i \in S$. The case $n = 0$ is trivial as $1_{\mathcal{G}}$ is the only element of length 0. Assume that the claim is true for all elements of length less than some $n > 0$ and consider an arbitrary element $g \in \mathrm{St}_\beta(\xi_0)$ of length n. We have $g = s_n \dots s_2 s_1$, where each s_i is a generator from S. Let $\xi_k = (s_k \dots s_2 s_1)(\xi_0)$, $k = 1, 2, \dots, n$. If $\xi_k = \xi_0$ for some $0 < k < n$, then $g_1 = s_n \dots s_{k+1}$ and $g_2 = s_k \dots s_2 s_1$ both fix ξ_0. Since the length of g_1 and g_2 is less than n, they belong to $\mathrm{St}^o_\beta(\xi_0)H$ by the inductive assumption. As $\mathrm{St}^o_\beta(\xi_0)H$ is a group, so does $g = g_1 g_2$. If $\xi_k \neq \xi_0$ for all $0 < k < n$, then $s_{i+1}|_{w_i} = 1_{\mathcal{G}}$ for any $0 \leq i < n$ and sufficiently long beginning w_i of the infinite word ξ_i. It follows that $g|_w = 1_{\mathcal{G}}$ for a sufficiently long beginning w of ξ_0. Thus $g \in \mathrm{St}^o_\beta(\xi_0)$. □

Recall that we consider the Schreier graphs of the group $\mathcal{G}$ relative to the generating set $S = \{a, b, c, d\}$ as graphs with undirected edges. The Schreier graphs of all orbits of the action β except $O_\beta(\xi_0)$ are similar. Any vertex is joined to two other vertices. Moreover, it is joined to one of the neighbors by a single edge labeled a and to the other neighbor by two edges. Also, there is one loop at each vertex. Hence the Schreier graph has a linear structure (see Figure 1) and all such graphs are isomorphic as graphs with unlabeled edges. The Schreier graph of the orbit of $\xi_0 = 1^\infty$ is different in that there are three loops labeled b, c, and d at the vertex ξ_0 (see Figure 2).

Let $F : \partial\mathcal{T} \to \mathrm{Sch}(\mathcal{G}, S)$ be the mapping that assigns to any point on the boundary of the binary rooted tree $\mathcal{T}$ its marked Schreier graph under the action β. Using notation of Section 4, $F(\xi) = \Gamma^*_{\mathrm{Sch}}(\mathcal{G}, S; \beta, \xi)$ for all $\xi \in \partial\mathcal{T}$.

LEMMA 7.7. *The graph $F(\xi_0)$ is an isolated point in the image $F(\partial\mathcal{T})$.*

PROOF. Let Γ_0 denote the marked graph with a single vertex and three loops labeled b, c, and d. Recall that $\mathcal{U}(\Gamma_0, \emptyset)$ is an open subset of $\mathcal{MG}_0$ consisting of all graphs in $\mathcal{MG}_0$ that have a subgraph isomorphic to Γ_0. Hence $\mathcal{U}(\Gamma_0, \emptyset) \cap \mathrm{Sch}(\mathcal{G}, S)$ is an open subset of $\mathrm{Sch}(\mathcal{G}, S)$. Given $\xi \in \partial\mathcal{T}$, the graph $F(\xi)$ belongs to that open subset if and only if $a(\xi) \neq \xi$ and $b(\xi) = c(\xi) = d(\xi) = \xi$. The latter conditions are satisfied only for $\xi = \xi_0$. The lemma follows. □

It turns out that the image $F(\partial\mathcal{T})$ is not closed in $\mathrm{Sch}(\mathcal{G},S)$. The following construction will help to describe the closure of $F(\partial\mathcal{T})$. Let us take two copies of the Schreier graph $\Gamma_{\mathrm{Sch}}(\mathcal{G},S;\beta,\xi_0)$. We remove two out of three loops at the vertex ξ_0 (loops with the same labels in both copies) and replace them with two edges joining the two copies. Let c' and d' denote labels of the removed loops and b' denote the label of the retained loop. Then b',c',d' is a permutation of b,c,d. To be rigorous, the new graph has the vertex set $O_\beta(\xi_0)\times\{0,1\}$, the set of edges $O_\beta(\xi_0)\times\{0,1\}\times S$, and the set of labels S. An arbitrary edge (ξ,i,s) has beginning (ξ,i) and label s. The end of this edge is $(s(\xi),i)$ unless $\xi=\xi_0$ and $s=c'$ or $s=d'$, in which case the end is $(s(\xi),1-i)=(\xi_0,1-i)$. There are three ways to perform the above construction depending on the choice of b'. We denote by Δ_0, Δ_1, and Δ_2 the graphs obtained when $b'=b$, $b'=d$, and $b'=c$, respectively. Further, for any $i\in\{0,1,2\}$ we denote by Δ_i^* a marked graph obtained from Δ_i by marking the vertex $(\xi_0,0)$ (see Figure 3).

Consider an arbitrary sequence of points $\eta_1,\eta_2,\dots$ in $\partial\mathcal{T}$ such that $\eta_n\to\xi_0$ as $n\to\infty$, but $\eta_n\neq\xi_0$. Let z_n denote the position of the first zero in the infinite word η_n.

LEMMA 7.8. *The marked Schreier graphs $F(\eta_n)$ converge to Δ_i^*, $0\le i\le 2$, as $n\to\infty$ if $z_n\equiv i \bmod 3$ for large n.*

PROOF. For any $n\ge 1$ we define a map $f_n: O_\beta(\xi_0)\times\{0,1\}\to O_\beta(\eta_n)$ as follows. Given $\xi\in O_\beta(\xi_0)$ and $x\in\{0,1\}$, let $f_n(\xi,x)$ be an infinite word obtained from η_n after replacing the first z_n-1 letters by the first z_n-1 letters of ξ and adding $x \bmod 2$ to the (z_n+1)-th letter. Clearly, $f_n(\xi_0,0)=\eta_n$. Let i_n be the remainder of z_n under division by 3. One can check that the restriction of f_n to the vertex set of the closed ball $\overline{B}_{\Delta_{i_n}^*}((\xi_0,0),N)$ is an isomorphism of this ball with the closed ball $\overline{B}_{F(\eta_n)}(\eta_n,N)$ whenever $N\le 2^{z_n-2}$. Therefore $\delta(F(\eta_n),\Delta_{i_n}^*)\to 0$ as $n\to\infty$. □

One consequence of Lemma 7.8 is that the graphs Δ_0, Δ_1, and Δ_2 are Schreier graphs of the group $\mathcal{G}$. By construction, each of these graphs admits a nontrivial automorphism, which interchanges vertices corresponding to the same vertex of $\Gamma_{\mathrm{Sch}}(\mathcal{G},S;\beta,\xi_0)$. This property distinguishes Δ_0, Δ_1, and Δ_2 from the Schreier graphs of orbits of the action β.

LEMMA 7.9. *The Schreier graphs $\Gamma_{\mathrm{Sch}}(\mathcal{G},S;\beta,\xi)$, $\xi\in\partial\mathcal{T}$ do not admit nontrivial automorphisms. The graphs Δ_0, Δ_1, and Δ_2 admit only one nontrivial automorphism.*

PROOF. It follows from Proposition 4.4 and Lemma 7.4 that marked Schreier graphs $F(\xi)$ and $F(\eta)$ are isomorphic only if $\xi=\eta$. Therefore the Schreier graphs $\Gamma_{\mathrm{Sch}}(\mathcal{G},S;\beta,\xi)$, $\xi\in\partial\mathcal{T}$ admit no nontrivial automorphisms.

The graphs Δ_0, Δ_1, and Δ_2 have linear structure. Namely, one can label their vertices by v_j, $j\in\mathbb{Z}$ so that each v_j is adjacent only to v_{j-1} and v_{j+1}. If f is an automorphism of such a graph, then either $f(v_j)=v_{n-j}$ for some $n\in\mathbb{Z}$ and all $j\in\mathbb{Z}$ or $f(v_j)=v_{n+j}$ for some $n\in\mathbb{Z}$ and all $j\in\mathbb{Z}$. Assume that some Δ_i has more than one nontrivial automorphism. Then we can choose f above so that the latter option holds with $n\neq 0$. Take any path in Δ_i that begins at v_0 and ends at v_n and let w be the code word of that path. Since $f^m(v_0)=v_{mn}$ and $f^m(v_n)=v_{(m+1)n}$ for any integer m, the path in Δ_i with beginning v_{mn} and code word w ends at

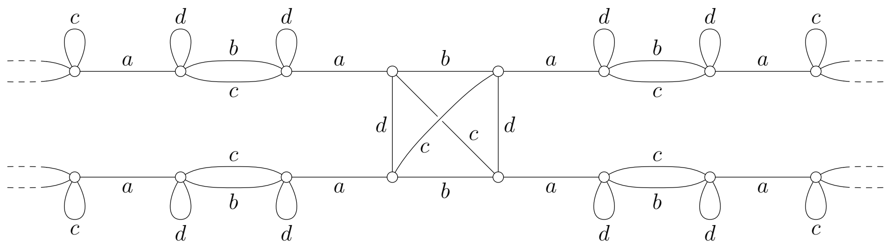

FIGURE 4. The Schreier coset graph of $\mathrm{St}^o_\beta(\xi_0)$.

$v_{(m+1)n}$. It follows that for any integer $m > 0$ the path with beginning v_0 and code word w^m ends at v_{mn}. In particular, this path is not closed. However every element of the Grigorchuk group $\mathcal{G}$ is of finite order (see [**2**]) so that for some $m > 0$ the reversed word w^m equals $1_\mathcal{G}$ when regarded as a product in $\mathcal{G}$. This conradicts with Proposition 4.1. Thus the graph Δ_i admits only one nontrivial automorphism. $\square$

LEMMA 7.10. *The Schreier graph $\Gamma_{\mathrm{Sch}}(\mathcal{G}, S; \beta, \xi_0)$ is a double quotient of each of the graphs Δ_0, Δ_1, and Δ_2. On the other hand, each of the graphs Δ_0, Δ_1, and Δ_2 is a double quotient of the Schreier coset graph $\Gamma_{\mathrm{coset}}(\mathcal{G}, S; \mathrm{St}^o_\beta(\xi_0))$.*

PROOF. The Schreier coset graph of the subgroup $\mathrm{St}^o_\beta(\xi_0)$ is shown in Figure 4. In view of Lemmas 7.6 and 7.9, the automorphism group of this graph is the Klein 4-group. The quotient of the graph by the entire automorphism group is the Schreier graph of the orbit of ξ_0. The quotients by subgroups of order 2 are the graphs Δ_0, Δ_1, and Δ_2. $\square$

Now it remains to collect all parts in Theorems 1.1 and 1.2.

PROOF OF THEOREM 1.1. We are concerned with the mapping $F : \partial\mathcal{T} \to \mathrm{Sch}(\mathcal{G}, S)$ given by $F(\xi) = \Gamma^*_{\mathrm{Sch}}(\mathcal{G}, S; \beta, \xi)$. Let us also consider a mapping $\psi : \partial\mathcal{T} \to \mathrm{Sub}(\mathcal{G})$ given by $\psi(\xi) = \mathrm{St}_\beta(\xi)$ and a mapping $f : \mathrm{Sub}(\mathcal{G}) \to \mathrm{Sch}(\mathcal{G}, S)$ given by $f(H) = \Gamma^*_{\mathrm{coset}}(\mathcal{G}, S; H)$. By Proposition 4.4, $F(\xi) = f(\psi(\xi))$ for all $\xi \in \partial\mathcal{T}$. By Proposition 5.5, f is a homeomorphism. Lemma 7.4 implies that the mapping ψ is injective. It is Borel measurable due to Lemma 5.4. Also, ψ is continuous at a point $\xi \in \partial\mathcal{T}$ if and only if $\mathrm{St}^o_\beta(\xi) = \mathrm{St}_\beta(\xi)$. Lemmas 7.5 and 7.6 imply that the latter condition fails only if the infinite word ξ contains only finitely many zeros. According to Lemma 7.2, an equivalent condition is that ξ is in the orbit of $\xi_0 = 1^\infty$ under the action β. Since the mapping F is f postcomposed with a homeomorphism, it is also injective, Borel measurable, and continuous everywhere except the orbit of ξ_0.

By Lemma 7.7, the graph $F(\xi_0)$ is an isolated point of the image $F(\partial\mathcal{T})$. Since $F(g(\xi)) = \mathcal{A}_g(F(\xi))$ for any $\xi \in \partial\mathcal{T}$ and $g \in \mathcal{G}$ and since the action $\mathcal{A}$ is continuous (see Proposition 4.2), the graph $F(g(\xi_0))$ is an isolated point of $F(\partial\mathcal{T})$ for all $g \in \mathcal{G}$. On the other hand, if $\xi \in \partial\mathcal{T}$ is not in the orbit of ξ_0, then the graph $F(\xi)$ is not an isolated point of $F(\partial\mathcal{T})$ as the mapping F is injective and continuous at ξ.

It follows from Lemma 7.9 that the image $F(\partial\mathcal{T})$ and the orbits $O_\mathcal{A}(\Delta^*_i)$, $i \in \{0, 1, 2\}$ are disjoint sets. Note that the orbit $O_\mathcal{A}(\Delta^*_i)$ consists of marked graphs obtained from the graph Δ_i by marking an arbitrary vertex. Lemma 7.8 implies the union of those 4 sets is the closure of $F(\partial\mathcal{T})$.

Finally, the statement (v) of Theorem 1.1 follows from Lemma 7.10. □

Proof of Theorem 1.2. Lemma 6.6 combined with Lemma 7.4 implies that the action of $\mathcal{G}$ on the closure of $F(\partial\mathcal{T})$ is a continuous extension of the action β. The extension is one-to-one everywhere except for the orbit $O_\beta(\xi_0)$ where it is four-to-one. Namely, for any $g \in \mathcal{G}$ the point $g(\xi_0)$ is covered by 4 graphs $F(g(\xi_0))$, $\mathcal{A}_g(\Delta_0^*)$, $\mathcal{A}_g(\Delta_1^*)$, and $\mathcal{A}_g(\Delta_0^*)$. According to Theorem 1.1, the graph $F(g(\xi_0))$ is an isolated point of the closure of $F(\partial\mathcal{T})$. When we restrict our attention to the set Ω of non-isolated points of the closure, we still have a continuous extension of the action β, but it is three-to-one on the orbit $O_\beta(\xi_0)$.

By Lemma 7.2, the group $\mathcal{G}$ acts transitively on each level of the binary rooted tree $\mathcal{T}$. Then Proposition 6.2 implies that the action β is minimal and uniquely ergodic, the only invariant Borel probability measure being the uniform measure on $\partial\mathcal{T}$. Since the action of $\mathcal{G}$ on the set Ω is a continuous extension of the action β that is one-to-one except for a countable set and since this action has no finite orbits, it follows that the action is minimal, uniquely ergodic, and isomorphic to β as the action with an invariant measure. □

References

1. R. Grigorchuk, *On Burnside's problem on periodic groups.* Funct. Anal. Appl. **14** (1980), 41–43. MR565099 (81m:20045)
2. R. Grigorchuk, *Solved and unsolved problems around one group.* L. Bartholdi (ed.) et al., Infinite groups: geometric, combinatorial and dynamical aspects. Basel, Birkhäuser. *Progress in Mathematics* **248**, 117–218 (2005). MR2195454 (2007d:20001)
3. R. Grigorchuk, *Some topics in the dynamics of group actions on rooted trees.* Proc. Steklov Inst. Math. **273** (2011), no. 1, 64–175.
4. A. S. Kechris, *Classical descriptive set theory.* Berlin, Springer. *Graduate Texts in Mathematics* 156 (1995). MR1321597 (96e:03057)
5. A. Vershik, *Nonfree actions of countable groups and their characters.* J. Math. Sci., NY **174** (2011), no. 1, 1–6. MR2749291 (2012b:37015)

Department of Mathematics, Texas A&M University, College Station, Texas 77843-3368

E-mail address: `yvorobet@math.tamu.edu`

Contemporary Mathematics
Volume **567**, 2012
http://dx.doi.org/10.1090/conm/567/11253

Minimal models for free actions

Benjamin Weiss

Abstract. Any infinite countable group G has a universal minimal action. The universality is in the sense that the invariant probability measures represent, up to isomorphism, all free actions of G on a Lebesgue space.

1. Introduction

In general by a **model** for a measure preserving action of a group G, say $(X, \mathcal{B}, \mu, T_g)$, we mean an isomorphic copy of this action which has some additional structure. The simplest example of this is the familiar representation of this measurable action by a continuous action on a compact metric space. Recall that this can be achieved by first constructing a separable subalgebra $\mathcal{A}$ of $L^\infty(X, \mathcal{B}, \mu)$ that is invariant under the group action and is dense in $L^2(X, \mathcal{B}, \mu)$, and then letting $\hat{X}$ be the maximal ideal space of $\mathcal{A}$. Another kind of model derives from the theorem of V. Rokhlin to the effect that when $G = \mathbb{Z}$ and the action is ergodic there is a countable partition $\mathcal{Q}$ of the space which is a generator. This is used to represent any such measure preserving transformation as a countably valued stochastic process. This result of Rokhlin can easily be extended to any countable group.

For $G = \mathbb{Z}$ there are much stronger theorems of this general nature. For example, the classic Jewett-Krieger theorem gives a uniquely ergodic model for any ergodic action. While many of these stronger results (cf [**GW**]) have been extended to amenable groups, they have not been extended to general groups. It is the purpose of this note to show that any free action of an infinite countable group has a minimal model - that is to say it is isomorphic to a minimal action of G on a compact metric space preserving an invariant measure. Even the question of the existence of a minimal action of G preserving a measure such that almost every orbit is free appears to have been open up to now.

Indeed I shall construct a **universal minimal action** of G. This is a minimal action of G whose invariant measures represent, up to isomorphism, all free probability preserving actions of G. For the case of $\mathbb{Z}$ this was carried out in [**W**]. In order to do this I first construct a minimal action of G whose invariant measures include a free factor of any free action of G. Then I apply a general result that shows how almost one to one minimal extensions can have a wide variety of invariant measures, which appears in [**FW**], to get the universal minimal action.

2010 *Mathematics Subject Classification.* Primary 37B05.

This work was supported by ISF grant #1157/08 and by BSF grant # 2006312.

Our main interest at that time was in $\mathbb{Z}$, and the proof for the general group was only sketched briefly. I give a more detailed proof of that result in §5. The next section is devoted to the construction of a a sequence of finite sets in G which will play a key role in the first construction. In §3 I give an easier example of the main construction which will be carried out in §4. The last section contains the main theorem.

I would like to thank Alekos Kechris for raising the question of whether an arbitrary group G has a minimal action preserving a free ergodic invariant measure. It was his question which gave the original impetus to this work. I also wish to thank Hanfeng Li for all of his valuable comments on earlier versions of this paper. Finally I am indebted to the referee for a very careful reading which resulted in the correction of several inaccuracies.

2. The construction of the K_n

Throughout the paper G will denote a fixed countably infinite group. The integers $\{0, 1, 2, ...n\}$ will be denoted by I_n and for $1 \leq m < n$ the mapping $\pi_{m,n}$ is the map that takes the integers greater than or equal to m to m and is the identity on the rest. If F is any subset of G then we will use the same notation for the mapping from I_n^F to I_m^F that is induced by $\pi_{m,n}$. If $z \in I_n^F$ then we will say that an integer i **occurs** at the subset $z[i] \subset F$ if $z(f) = i$ just when $f \in z[i]$. If K is a finite subset of the group we will say that the **occurrences**, $z[i]$, of an integer i in z are K**-disjoint** if the sets Kf are pairwise disjoint for $f \in z[i]$. They will be said to be K**-syndetic** if the union of the Kf as f ranges over $z[i]$ includes F.

Remark 2.1. The basic fact that we will use repeatedly is that if K is a symmetric finite subset of G and E is a maximal subset of F that is K-disjoint then it is K^2-syndetic, and conversely, if a K-disjoint subset of F is K^2-syndetic then it is maximally K-disjoint.

In the course of the construction of the minimal model we will construct an increasing sequence of finite symmetric subsets of G, K_n, and we shall need to consider special elements of I_n^G whose definition involves this sequence. The actual construction of the sets we defer to the next section.

For any finite set $F \subset G$, an element $z \in I_1^F$ will be called **1-acceptable** if there is a subset $C \subset K_1^2 F$ that satisfies:

(1) The set C is K_1-disjoint.
(2) $K_1^2 C \supset F$.
(3) $z[1] = C \cap F$.

These are the kind of sets that arise from the restriction to F of maximal K_1-disjoint subsets of supersets of F.

A $z \in I_2^F$ will be called **2-acceptable** if there are subsets $C_2 \subset C_1 \subset K_2^2 F$ that satisfy:

(1) The set C_1 is K_1-disjoint and the set C_2 is K_2-disjoint.
(2) For any $g \in K_2^2 F$ such that $K_1^2 g \subset K_2^2 F$ there is a $c_1 \in C_1$ such that $g \in K_1^2 c_1$, and for any $g \in C_1$ such that $K_2^2 g \subset K_2^2 F$ there is a $c_2 \in C_2$ such that $g \in K_2^2 c_2$.
(3) For $i = 1, 2$ we have $\pi_{i,2}(z)[i] = C_i \cap F$.

After these two cases the general pattern remains the same and we say that a $z \in I_n^F$ will be called **n-acceptable** if there are subsets

$$C_n \subset C_{n-1} \subset \dots C_2 \subset C_1 \subset C_0 = K_n^2 F$$

that satisfy:

(1) For $1 \le i \le n$ set C_i is K_i-disjoint.
(2) For $1 \le i \le n$ and any $g \in C_{i-1}$ such that $K_i^2 g \subset K_n^2 F$ there is a $c_i \in C_i$ such that $g \in K_i^2 c_i$.
(3) $\pi_{i,n}(z)[i] = C_i \cap F$.

These sets C_i are not necessarily unique - however we will usually consider one such choice to be associated with an n-acceptable word. Acceptability is robust in the sense that if $E \subset F$ and $z \in I_n^F$ is n-acceptable then the restriction of z to E is also n-acceptable. The other direction is also valid, namely any n-acceptable element of I_n^E is the restriction of an n-acceptable element of I_n^F. In fact something much more general is true as the next lemma shows. Before formulating the lemma it will be convenient to introduce one more notion. Finite sets in the group F_j will be said to be K**-separated** if the sets KF_j are disjoint.

LEMMA 2.2. *If the E_j are a collection of finite sets that are K_n^3-separated, and $z_j \in I_n^{E_j}$ are n-acceptable, then for any finite set F that contains the union of all the E_j's there is an n-acceptable $z \in I_n^F$ that for all j agrees with z_j on all $f \in E_j$.*

.

PROOF. We shall begin by proving this in the case $n = 1$. From the definition of 1-acceptable we are given sets $C_j \subset K_1^2 E_j$ that are K_1-disjoint such that $z_j[1] = E_j \cap C_j$. Since the sets E_j are K_1^3-separated the union $\bigcup_j C_j$ continues to be K_1-disjoint. It follows that we can extend this to a maximal K_1-disjoint subset $C \subset K_1^2 F$. Define $z \in I_1^F$ by setting $z[1] = C \cap F$, this yields an element z as required. On the one hand the fact that C is maximally K_1-disjoint in $K_1^2 F$ implies the second condition in the definition of 1-acceptable whereas the other two are trivially satisfied. The agreement with the initial z_j's on the E_j's follows from the fact that they satisfied the key second condition for each E_j.

For the general case we will denote the sets C_i that are involved in the n-acceptability of the z_j by $C_{i,j}$. We proceed in a sequence of steps beginning with the first level and repeat what we just did for $n = 1$. That means that we extend the union of the $C_{1,j}$ to a maximal K_1-disjoint subset of $K_n^2 F$ and call the resulting set C_1. Next we enlarge the union of the $C_{2,j}$'s to a maximal K_2-disjoint subset of C_1 and denote this by C_2. It is clear how to continue this all the way up to level n. This produces a sequence of sets $C_n \subset C_{n-1} \subset \dots C_2 \subset C_1 \subset K_n^2 F$, where C_{i+1} is a maximal K_{i+1}-disjoint subset of C_i. The required z is defined by setting $z[i] = (C_i \backslash C_{i+1}) \cap F$ for $i < n$ and $z[n] = C_n \cap F$. The fact that this z is n-acceptable follows from the construction. $\square$

The symmetric sets K_n that we will soon construct will have the property that $K_i^3 \subset K_{i+1}$ for all i and this will imply (we will assume that K_1 contains the identity) that $K_1^2 K_2^2 K_3^2 \cdots K_n^2 F_j \subset K_n^3 F_j$ for all n. This in turn will easily show that for an n-acceptable word $z \in I_n^F$ the sets $\pi_{i,n}(z)[i]$ are essentially K_i^3-syndetic, where essentially means that any $g \in F$ such that $K_i^3 g \subset F$ is contained in $K_i^3 z[i]$.

To define the K_n we will use an auxiliary sequence L_n and some enumeration of the elements of G. Set $L_1 = \{e\}$, the identity element of the group and let K_1 be any symmetric set that strictly contains L_1. Denote by W_1 the collection of all 1-acceptable words contained in $I_1^{K_1^5}$ and choose for $L_2 \supset K_1^5$ a finite symmetric subset of the group that is large enough to contain $|W_1|$ right translates of K_1^5 in the complement of K_1^7 that are K_1^3-separated. For K_2 take a symmetric subset of the group that contains $K_1^{14}L_2$ and the first two elements of G and is large enough so that

$$\frac{|K_1^{14}L_2|}{|K_2|} < \frac{1}{10}.$$

For the general step we suppose that for $1 \le i < n$ we have already defined symmetric sets $L_i \subset K_i$, and have defined W_i as the collection of all i-acceptable words in $I_i^{K_i^5}$ and that for $1 < i < n$ they satisfy:

- The set L_i contains K_{i-1}^5 and $|W_{i-1}|$ translates of K_{i-1}^5 in the complement of K_{i-1}^7, that are K_{i-1}^3-separated.
- The set K_i contains the first i elements of the group and $K_{i-1}^{14}L_i$.
- The set K_i is large enough so that

$$\frac{|K_{i-1}^{14}L_i|}{|K_i|} < \frac{1}{10^{i-1}}.$$

Let L_n be any symmetric set containing K_{n-1}^5 and large enough to accommodate $|W_{n-1}|$ translates of K_{n-1}^5 in the complement of K_{n-1}^6 that are K_{n-1}^3-separated and then choose K_n to be any symmetric set that contains the first n elements of the group together with $K_{n-1}^{14}L_n$ and is large enough so that

$$\frac{|K_{n-1}^{14}L_n|}{|K_n|} < \frac{1}{10^{n-1}}.$$

It is clear that this procedure constructs an infinite sequence of symmetric sets $L_n \subset K_n$ that satisfy the above conditions and that their union is all of G. For a finitely generated group we could have taken K_1 to be a symmetric set of generators together with the identity and then all of the subsequent L_n and K_n could be taken to be appropriate powers of K_1.

3. Invariant measures on the 1-acceptable sequences

For Z a finite set or any compact space, Z^G with the product topology is also compact and there is a natural action of G on the space - obtained from the action of G on itself by multiplication on the right. This is what we will call the **shift** action σ_g. The definition is for a $u \in Z^G$ set $\sigma_g(u)(h) = u(hg)$. The minimal system that we will construct will be a closed shift invariant subset of such a shift action. As a first step we consider the space $Y \subset I_1^G$ that consists of all 1-acceptable elements y, i.e. those y such that the occurrences of 1 in y are a maximal K_1-disjoint set. Under the shift action of G this is a closed and invariant subspace and one can ask does it support an invariant measure. For an amenable group the answer is trivially yes, but it is not so clear how to construct such an invariant measure for a general group. In fact we won't do this directly but rather begin with a free measure preserving action of G on some probability space $(X, \mathcal{B}, \mu, T_g)$. For a specific example one can take the full shift I^G where I is any finite set and product measure of any probability distribution on I. For a finite subset F in the group a measurable subset B in X is said to be an F**-set** if the sets $T_f(B)$ for $f \in F$ are pairwise disjoint. The following lemma is a well-known basic consequence of the freeness of the action :

LEMMA 3.1. *If $(X, \mathcal{B}, \mu, T_g)$ is a free action and F is any nonempty finite subset of G then any positive measure subset $B \subset X$ contains an F-set with positive measure.*

Fix a nonempty symmetric set F and apply the lemma to X to get an F-set with positive measure A_1. Furthermore, one can clearly arrange that $\mu(A_1) > \frac{c_0}{2}$ where c_0 is the supremum of the μ-measures of F-sets in X. We now examine the $B_1 = X \setminus \bigcup_{f \in F^2} T_f(A_1)$. If $\mu(B_1) = 0$ we stop and set $A = A_1$. If it is positive we apply the lemma again to B_1 to obtain an F-set $C_1 \subset B_1$ of positive measure such that $\mu(C_1) > \frac{c_1}{2}$ where c_1 is the supremum of the μ-measures of the F-sets in B_1. We set $A_2 = A_1 \bigcup C_1$ and observe that A_2 is again an F-set. Next we set $B_2 = X \setminus \bigcup_{f \in F^2} T_f(A_2)$ and if $\mu(B_2) = 0$ we stop and set $A = A_2$. Otherwise we continue as before applying the lemma to B_2. This procedure needn't terminate at a finite step, and if it doesn't then we set A to be the union of the A_i. It is then easy to see that A is still an F-set and that now we are certain that $\mu(\bigcup_{f \in F^2} T_f(A)) = 1$. This argument establishes the following lemma:

LEMMA 3.2. *If $(X, \mathcal{B}, \mu, T_g)$ is a free action and F is any finite symmetric subset of G then there is an F-set A in X such that*

$$\mu(\bigcup_{f \in F^2} T_f(A)) = 1.$$

Applying this lemma to the set K_1 gives a measurable subset $A \subset X$ which is a K_1-set and has the property that $\mu(\bigcup_{f \in K_1^2} T_f(A)) = 1$. If ϕ denotes the indicator function of A then we define a map from X to I_1^G by setting

$$\Phi(x)(h) = \phi(T_h(x)).$$

This mapping is equivariant for the shift action on I_1^G, since

$$\Phi(T_g(x))(h) = \phi(T_h(T_g(x))) = \phi(T_{hg}(x)) = \Phi(x)(hg) = \sigma_g(\Phi(x))(h)$$

and it follows that $\nu = \Phi \circ \mu$ defines an invariant measure for the shift action. The properties of A ensure that for μ-almost every $x \in X$ the set of $g \in G$ such that $T_g(x) \in A$ is both K_1-disjoint and K_1^2-syndetic and hence is a maximal K_1-disjoint set. This implies that the closed support of ν is contained in Y and gives an example of the kind of measure that we were looking for. This simple construction will be amplified in the next section in order to get the desired minimal model for the free action.

4. Constructing the partition

Throughout this section we fix a free action of G, $(X, \mathcal{B}, \mu, T_g)$, and proceed to construct a sequence of partitions $\mathcal{Q}_n$ of X, where the n-th partition is indexed by I_n, and we view the partition as measurable function from X to I_n. The sequence will converge to a limiting countable partition $\mathcal{Q}$ indexed by $\mathbb{N}$ (which we take to include zero). The one point compactification of $\mathbb{N}$ will be denoted by I_∞, and we shall use the partition $\mathcal{Q}$ to define a mapping Φ from X to ${I_\infty}^G$ as we did before, and matters will be arranged so that the closed support of the measure $\Phi \circ \mu$ will be a minimal set in which all orbits are free. In the next section we shall show how to use this minimal factor to get a proper minimal model.

Step 1. The first step in the construction is exactly what we did in the previous section, with the function $\mathcal{Q}_1$ being the indicator function of the set A constructed there. Since this set will be modified in the course of the construction we denote it by A_1^1 and its complement by A_0^1, where the subscript indicates that the value of $\mathcal{Q}_1$ for points in A_i^1 is i, and the superscript indicates the stage of the construction. We will be using the sets $L_n \subset K_n$ that were constructed in §2.

Step 2. Applying Lemma 3.2 we find a set $A_2^2 \subset X$ that is a K_2-set and satisfies $\mu(\bigcup_{f\in K_2^2} T_f(A_2^2)) = 1$. The set A_2^2 is partitioned into a finite number of pieces E_l, according to the common refinement of $T_g^{-1}(\mathcal{Q}_1 \cap T_g(A_2^2))$ as g ranges over all the members of K_2. Thus on each piece, E_l , for each $g \in K_2$ either $T_g(E_l)$ is contained in A_1^1 or in its complement. In the subsequent discussion we will use the following abbreviation. For a finite subset F in G and a measurable subset C in X we will denote by $T_F(C)$ the union $\bigcup_{f\in F} T_f(C)$.

The modifications to the sets A_i^1 will be carried out separately on the sets $T_{K_2}(E_l)$. Fix one of these sets and denote it simply by E. To begin with remove from A_1^1 all of the sets $T_{K_1^{10}L_2}(E)$. For each word $w \in W_1 \subset I_1^{K_1^5}$ there is a distinct translate of K_1^5 in L_2 and we consider that word to be written there. Put the symbol 1 at the identity of the group and applying Lemma 2.2 we get a 1-acceptable word in $I_1^{L_2}$. For future reference we denote the word obtained from it by changing the 1 which was written at e to a 2 by s_2. On $K_2 \backslash K_1^{10}L_2$ the partition $\mathcal{Q}_1$ defines a word $\hat{z}$, namely $\hat{z}(g) = i$ just in case $T_g(E) \subset A_i^1$. This word is 1-acceptable and therefore we can apply Lemma 2.2 again to patch this together with what we have already defined on L_2 to get a 1-acceptable word w_E defined now on all of K_2. This word w_E agrees with what is defined by $\mathcal{Q}_1$ on most of the space and we use it to define our new partition $\mathcal{Q}_2$. This is done by putting the sets $T_g(E)$ in A_i^2 just when $w_E(g) = i$ for all $e \neq g \in K_2$. After this is done for all of the sets E_l we have completed the definition of $\mathcal{Q}_2$. Matters have been arranged so that for μ-almost every x the words w defined by $w(g) = \mathcal{Q}_2(T_g(x))$ are 2-acceptable over all of G, and furthermore all possible 1-acceptable $I_1^{K_1^3}$ words appear K_2^2-syndetically in these w's. This is accomplished by having the fixed pattern that was constructed above, s_2, appearing in the L_2 neighborhood of every occurrence of 2. This s_2 was constructed so that all possible 1-acceptable words in $I_1^{K_1^3}$ appear in it. Note too that the distance between $\mathcal{Q}_1$ and $\mathcal{Q}_2$ in the L^1-metric is at most 0.1. In the next stage the change will be even smaller and this will ensure the eventual convergence of the partitions.

Remark 4.1. Before turning to the general case we will describe the next step since a new feature appears here which will require the introduction of yet another important notion. A 2-acceptable element w of I_2^F, where $F \subset G$ is a finite subset, will be said to be **2-welcome** if whenever $f \in F$ and $w(f) = 2$ we have $w(gf) = s_2(g)$ for all $g \in L_2$ with $gf \in F$. The partition $\mathcal{Q}_2$ was constructed in such a way that for μ-almost every $x \in X$ the words defined by $\mathcal{Q}_2$ on the orbit of x are 2-welcome. These are the words w given by $w(g) = \mathcal{Q}_2(T_g(x))$.

Step 3. Applying Lemma 3.2 we find a set $A_3^3 \subset X$ that is a K_3-set and satisfies $\mu(\bigcup_{f\in K_3^2} T_f(A_3^3)) = 1$. The set A_3^3 is partitioned into a finite number of pieces E_l (since this is a temporary notation we do not indicate the stage of the construction - these are not the same as the E_l in Step 2) according to the common refinement of $T_g^{-1}(\mathcal{Q}_2 \cap T_g(A_3^3))$ as g ranges over all the members of K_3. Thus on each piece, E_l , for each $g \in K_3$ the set $T_g(E_l)$ is contained in a unique set of $\mathcal{Q}_2$.

The modifications to the sets A_i^2 will be carried out separately on the sets $T_{K_3}(E_l)$. Fix one of these sets and denote it simply by E. We will proceed to modify the word z that $\mathcal{Q}_2$ defines on K_3 by $z(g) = i$ just when $T_g(E) \subset A_i^2$. The changes will be made in four stages.

I. In the first we consider the set $K_2^2 L_3$. Denote by $\hat{W}_2$ the 2-welcome words in W_2 and write each one of them in a distinct translate of K_2^5 that is contained in L_3 and then write s_2 on the set L_2 itself. These words are all in particular 2-acceptable and the sets where we have written them are sufficiently separated from each other and from L_2 so that we can apply Lemma 2.2 to obtain a 2-acceptable word $y \in I_1^{K_2^2 L_3}$ that agrees with s_2 on L_2, with the 2-welcome words on the translates of K_2^5. This word is the same for all of the sets E.

II. In the second stage we will modify this word so that it is also 2-welcome. This is done by applying to the new occurrences of 2, where we haven't guaranteed that the L_2 neighborhood coincides with s_2, the procedure that was described in Step 2. More explicitly, we erase the current symbols in the $K_1^{10} L_2$ neighborhoods of the new occurrences of 2 and write the word s_2 there. Then we fill in so as to get a 2-acceptable word using Lemma 2.2 which will now be 2-welcome. While this procedure might make some changes in the words we wrote on the translates of K_2^5 these cannot take place in the central K_2^3 part so that we still will see there all possible 2-welcome words in K_2^3 that are restrictions of of 2-welcome words on K_2^5. For the later applications it suffices to consider only such words. Furthermore, since at the identity we already had s_2 there is no need to make any change there. We obtain in this way a 2-welcome word w in $I_2^{K_2^2 L_3}$, and this word, restricted to L_3 with the central 2 changed to 3 will be denoted by s_3. It will be the same for all of the E_l. So far what we have done is independent of E.

III. In the third stage we restrict the original z which does depend on E to $K_3 \backslash K_2^{10} L_3$ which is K_2^3 separated from $K_2^2 L_3$ and apply Lemma 2.2 to this restriction and to $w \in I_2^{K_2^2 L_3}$ to obtain a 2-acceptable word u_E on K_3 with the following properties:

- The word u_E is 2-acceptable.
- The word u_E agrees with the original word z on $K_3 \backslash K_2^{10} L_3$ and is 2-welcome there.
- On L_3 the word u_E coincides with $\pi_{2,3}(s_3)$ and on $K_2^2 L_3$ it is 2-welcome.

IV. In the last stage we will modify u_E so that it becomes 2-welcome. This is done by the same procedure that was used in the second stage. As before the modifications won't affect the values of u_E in the "interior" of the two sets where it is already 2-welcome. The 2-welcome word $z_E \in I_2^{K_3}$ that we obtain in this way is used to define $\mathcal{Q}_3$ as follows. The sets $T_g(E)$ are included in A_i^3 (for $i \in \{0, 1, 2\}$ just when $w_E(g) = i$ for all $e \neq g \in K_3$.

After this is done for all of the sets E_l we have completed the definition of $\mathcal{Q}_3$. Matters have been arranged so that for μ-almost every x the words w defined by $w(g) = \mathcal{Q}_3(T_g(x))$ are 3-acceptable over all of G, and furthermore all possible

2-welcome $I_2^{K_2^2}$ words appear K_3^3-syndetically in these w's. These words have two additional features that define what we will call **3-welcome**.

DEFINITION 4.2. A word w in I_3^F will be said to be **3-welcome** if it satisfies:

- It is 3-acceptable.
- For all $f \in F$ and $w(f) = 2$ we have $w(gf) = s_2(g)$ for all $g \in L_2$ with $gf \in F$.
- For all $f \in F$ and $w(f) = 3$ we have $w(gf) = s_3(g)$ for all $g \in L_3$ with $gf \in F$.

Note too that the distance between $\mathcal{Q}_2$ and $\mathcal{Q}_3$ in the L^1-metric is at most 0.01.

Step $n+1$. We will describe now how to pass from $\mathcal{Q}_n$ to $\mathcal{Q}_{n+1}$ but first we need to define what we mean by ***i*-welcome**. The definition is inductive and will use the special i-welcome words $s_i \in I_i^{L_i}$ with the property that they contain copies of all the $(i-1)$-welcome words in $I_{i-1}^{K_{i-1}^3}$. To unify the notation and start the induction s_1 is set to be the word with the value 1 at e, the identity of the group, and a word is **1-welcome** just if it is 1-acceptable. This is consistent with the general notation since $L_1 = \{e\}$.

DEFINITION 4.3. For $i > 1$ a word w in I_i^F will be said to be ***i*-welcome** if it satisfies:

- It is i-acceptable.
- For all $f \in F$ and all $1 \le j \le i$ if $w(f) = j$ we have $w(gf) = s_j(g)$ for all $g \in L_j$ with $gf \in F$.

Our inductive assumption is that the partition $\mathcal{Q}_n = \{A_0^n, A_1^n, \ldots A_n^n\}$ has been defined in such a way that it satisfies the following properties:

- For all $0 < i \le n$ the set A_i^n is a K_i-set.
- For almost every $x \in X$ and $F \subset G$ the word $w \in I_n^F$ defined by $w(f) = \mathcal{Q}_n(T_f(x))$ is n-welcome and for all $1 \le i < n$ the word $\pi_{i,n}(w)$ is i-welcome.

Applying Lemma 3.2 we find a set $A_{n+1}^{n+1} \subset X$ that is a K_{n+1}-set and satisfies $\mu(\bigcup_{f \in K_{n+1}^2} T_f(A_{n+1}^{n+1})) = 1$. The set A_{n+1}^{n+1} is partitioned into a finite number of pieces E_l according to the common refinement of $T_g^{-1}(\mathcal{Q}_n \cap T_g(A_{n+1}^{n+1}))$ as g ranges over all the members of K_{n+1}. Thus on each piece, E_l , for each $g \in K_{n+1}$ the set $T_g(E_l)$ is contained in a unique set of $\mathcal{Q}_n$.

The modifications to the sets A_i^n will be carried out separately on the sets $T_{K_{n+1}}(E_l)$. Fix one of these sets and denote it simply by E. Our procedure now mimics what we did in Step 3. We will proceed to modify the word z that $\mathcal{Q}_n$ defines on K_{n+1} by $z(g) = i$ just when $T_g(E) \subset A_i^n$. The changes will be made in four stages which parallel what we did in Step 3.

Stage I. In the first we consider the set $K_n^3 L_{n+1}$ and construct an entirely new word there. Denote by $\hat{W}_{n+1}$ the n-welcome words in W_{n+1} and write each one of them in a distinct translate of K_n^5 that is contained in $L_{n+1} \backslash K_n^{10}$. By the inductive assumption on the partition $\mathcal{Q}_n$ there clearly are words $t_n \in I_n^{K_n^4}$ that are n-welcome and satisfy $t_n(e) = n$ and we copy one of them on K_n^2. Note that this t_n extends s_n. Now to the translates of K_n^5 and the central K_n^4 with the n-welcome words just defined we apply Lemma 2.2 to obtain an n-acceptable word

$y \in K_n^3 L_{n+1}$ which agrees with these n-welcome words on the translates of K_n^5 and on K_n^2.

Stage II. We now need to modify this word y so that it becomes n-welcome over the entire set $K_n^3 L_{n+1}$. This argument is a bit more complicated and we separate it out as a lemma.

LEMMA 4.4. *If the $K_n^3 F_j$ for $j \in J$ are a collection of finite sets in G that are K_n^2-separated with their union contained in F and u is an n-acceptable word in I_n^F such that its restrictions to the $K_n^3 F_j$, $u_j \in I_n^{K_n^3 F_j}$ are all n-welcome, then there is a an n-welcome word $w \in I_n^F$ which agrees with the u_j's on the F_j's and furthermore $u[n] = w[n]$.*

PROOF. The proof is carried out by induction on n. For the purposes of the induction it is convenient to have a stronger lemma in which the hypothesis is strengthened by changing $K_n^3 F_j$ to $K_1^2 K_2^2 K_3^2 \cdots K_n^2 F_j$. That this is indeed a strengthening follows from the fact that our sequence of K_i's satisfies $K_i^3 \subset K_{i+1}$ for all i and this easily implies that $K_1^2 K_2^2 K_3^2 \cdots K_n^2 F_j \subset K_n^3 F_j$. For $n = 1$ there is nothing to prove since 1-acceptable and 1-welcome coincide. Suppose that we have already established the lemma for all $i < n$. Consider $\pi_{n-1,n}(u)$ and apply the inductive hypothesis to obtain a $\hat{v} \in I_{n-1}^F$ which agrees with the $\pi_{n-1,n}(u_j)$ on the $K_n^2 F_j$'s, which is $n-1$-welcome and whose sets of occurrences of $n-1$ coincides with that of $\pi_{n-1,n}(u)$. We define v by changing its value to n exactly on the set $u[n]$. This v agrees with the u_j's on the $K_n^2 F_j$'s and is therefore already n-welcome there.

To continue we will need an extension of s_n to $K_{n-1}^3 L_n$ which continues to be n-welcome. The existence of such an extension follows from the fact that the words defined by $\mathcal{Q}_n$ are n-welcome. We denote one such extension by $\hat{s}_n \in I_n^{K_{n-1}^3 L_n}$.

Returning to v we proceed to modify it in two steps. In the first, if $v(f) = n$ and $f \in \bigcup_{j \in J} K_n^2 F_j$ we know that $v(gf) = s_n(g)$ whenever g L_n and gf is also in the union. If $K_{n-1}^6 L_n f \backslash K_{n-1}^3 L_n f$ has elements not in the union we change the value of v there to be zero and change it so that it agrees with $\hat{s}_n$ on all of $K_{n-1}^3 L_n f$. For the occurrences of n at an f that is not in the foregoing union we change the value of v to zero on $K_{n-1}^6 L_n f \backslash K_{n-1}^3 L_n f$ while setting v to be $\hat{s}_n$ on $K_{n-1}^3 L_n f$. It might be that we make changes at elements of some $K_n^2 F_j$, however clearly no changes are being made on $\bigcup_{j \in J} K_n F_j$. Denote the modified v by $\hat{v}$. If we now denote $\pi_{n-1,n}(\hat{v})$ by v_{n-1} then we can summarize its properties as follows:

- For all $f \in F$ where $\hat{v}(f) = n$ and all $e \neq g \in L_n$ with $gf \in F$ we have $v_{n-1}(gf) = \hat{v}(gf) = s_n(g)$.
- For all $f \in K_n F_j$, and a fortiori all $f \in K_{n-1}^3 F_j$ we have $v_{n-1}(f) = u_j(f)$.

Note that due to our setting the value of v to be zero on the "annuli" $K_{n-1}^6 L_n f \backslash K_{n-1}^3 L_n f$ we no longer have a word that is n-acceptable on all of F. However, applying Lemma 2.2 to v_{n-1} we can modify it maintaining the above two properties and have it also be $n-1$-acceptable. It follows that we can apply the inductive case of the current lemma, adding to the collection of F_j's the sets $L_n f$ for all $f \in \hat{v}[n] \backslash \bigcup_{j \in J} K_{n-1}^3 F_j$ in order to get a word $\hat{w} \in I_{n-1}^F$ which is $n-1$-welcome. Redefining $\hat{w}$ to be equal to n at all occurrences of n in $\hat{v}$ will give us the w that is n-welcome and agrees with the u_j's on the F_j's as required. $\square$

Applying the lemma to the word y of the first stage we get an n-welcome word w on $K_n^3 L_{n+1}$ and we denote by s_{n+1} its restriction to L_{n+1} with the value at e changed to $n+1$. This modified restriction contains copies of all n-welcome words in $I_n^{K_n^2}$ that are restrictions of n-welcome words on K_n^5, since we apply the lemma to translates of $K_n^5 = K_n^3 \cdot K_n^2$ and the n-welcome words on the K_n^2 are not changed. This word w is independent of the set E.

Stage III. In the third stage we come back to the set E which was used to define z and restrict z to $K_n^3(K_{n+1} \backslash K_n^{10} L_{n+1})$ which is K_n^3 separated from $K_n^3 L_{n+1}$ and apply Lemma 2.2 to this restriction and to $w \in I_n^{K_n^3 L_{n+1}}$ to obtain an n-acceptable word u_E on K_{n+1} with the following properties:

- The word u_E is n-acceptable.
- The word u_E agrees with the original word z on $K_n^3(K_{n+1} \backslash K_n^{10} L_{n+1})$ and is n-welcome there.
- On L_{n+1} the word u_E coincides with $\pi_{n,n+1}(s_{n+1})$ and on $K_n^3 L_{n+1}$ it is n-welcome.

Stage IV. In the last and final stage we will modify u_E so that it becomes n-welcome. This is done by an application of the lemma from Stage II and gives us an n-welcome word w_E which still coincides with s_{n+1} on $L_{n+1} \backslash \{e\}$ and with the original z on $K_{n+1} \backslash K_n^{10} L_{n+1}$. The n-welcome word $w_E \in I_n^{K_{n+1}}$ that we obtain in this way is used to define $\mathcal{Q}_{n+1}$ as follows. The sets $T_g(E)$ are included in A_i^{n+1} for $0 \le i \le n$ just when $w_E(g) = i$ for all $e \ne g \in K_{n+1}$. The set A_{n+1}^{n+1} was already defined at the outset. After this is done for all of the sets $E_l \subset A_{n+1}^{n+1}$ we have completed the definition of $\mathcal{Q}_{n+1}$.

The construction has guaranteed that for μ-almost every x the words w defined by $w(g) = \mathcal{Q}_{n+1}(T_g(x))$ are $(n+1)$-welcome over all of G, and furthermore all possible n-welcome words in $I_n^{K_n^2}$ that extend to K_n^5 appear K_{n+1}^3-syndetically in these w's. Furthermore our assumptions on the relative size of K_{n+1} and $K_n^{10} L_{n+1}$ ensure that the distance between $\mathcal{Q}_n$ and $\mathcal{Q}_{n+1}$ is at most $\frac{1}{10^n}$.

Passing to a limit we obtain a countable partition $\mathcal{Q}$ indexed by $\mathbb{N}$. The natural projections from $\mathbb{N}$ to I_n defined by $\pi_n(j) = j$ for all $0 \le j \le n$ and $\pi_n(j) = n$ for all $j > n$ map $\mathcal{Q}$ to partitions $\hat{\mathcal{Q}}_n$ that have all of the properties that the $\mathcal{Q}_n$ had, and this easily implies that if we define $\Phi : X \to I_\infty^G$ by $\Phi(x)(g) = \mathcal{Q}(T_g(x))$ then the closed support of $\Phi \circ \mu$ is a minimal set. Furthermore we can be certain of at least one free orbit in this minimal set namely the orbit of the point where some coordinate is equal to ∞ since at most one coordinate can equal ∞. We have established the following theorem:

THEOREM 4.5. *If G is a countably infinite group and $(X, \mathcal{B}, \mu, T_g)$ is a free action then there is a countable partition, $\mathcal{Q}$, of X with values in $\mathbb{N}$ and defining Φ by $\Phi(x)(g) = \mathcal{Q}(T_g(x))$ the closed support of $\Phi \circ \mu$ is a minimal set containing a free orbit.*

We conclude with a remark that will be used later on to get the universality of our model.

REMARK 4.6. Note that the definitions of n-acceptable and n-welcome were independent of the free action and were defined in terms of the group alone. This means that we have actually constructed a single minimal subshift $U \subset I_\infty^G$ and every free measure preserving action of G has a factor with a model on that space.

The space U is defined by the property that for all $u \in U$ and $n > 0$ all of the words that one sees in $\pi_n(u)$ are n-welcome. This means that in fact the minimal set of the theorem does not depend on the action or on the various choices that were made in the course of the construction.

5. Almost 1-1 extensions

The minimal system that we constructed in the preceding section with its invariant measure is in general not isomorphic to the system that we started with. In order to get a larger minimal system with an invariant measure which will be isomorphic to the initial free action we will use a general theorem that appears in [**FW**]. The proof given there of the theorem for a general group is somewhat sketchy in that it relies on the more detailed proof for the case $G = \mathbb{Z}$. For the reader's convenience I will give a more detailed account of the proof in this section. First I recall the notion of an almost $1-1$ extension. The context is that of actions of G as homeomorphisms of a compact metric space X. If (X, T_g) is one such action then another one (Y, S_g) is called an **extension** if there is a continuous mapping, $\pi : Y \to X$ that takes Y onto X and is equivariant, i.e. for all $g \in G$ we have $\pi S_g = T_g \pi$. In this case the system (X, T_g) is called a **factor** of (Y, S_g).

DEFINITION 5.1. An extension $\pi : Y \to X$ is said to be **almost** $1 - 1$ if π is one-to-one on a residual subset of X.

For minimal systems (X, T_g) the existence of a single point $x_0 \in X$ for which $\pi^{-1}(x_0)$ consists of a single point is sufficient to guarantee that the extension is almost $1 - 1$. It is well known that in a topological measure space one can have sets that are large topologically, but small in the sense of the measure. The classic example of this is the non-normal numbers in the unit interval which are residual but have Lebesgue measure zero. The theorem that we shall establish in this section is an extreme instance of this phenomenon. Throughout this section by "measure" we will always mean a Borel probability measure.

THEOREM 5.2. *Let G be a countably infinite group acting minimally on a totally disconnected space X with a free orbit x_0, and (Y, S_g) , an extension of (X, T_g) by π, which has some recurrent point $y_0 \in \pi^{-1}(x_0)$ whose orbit is dense in Y. Then there exists an almost $1 - 1$ minimal extension $(\hat{Y}, \hat{S}_g)$ of (X, T_g) given by $\hat{\pi} : \hat{Y} \to X$ and a Borel subset $Y_0 \subset Y$ with a map $\theta : Y_0 \to \hat{Y}$ satisfying:*

(1) *$\theta S_g = \hat{S}_g \theta$ for all $g \in G$, $\theta\hat{\pi} = \pi$.*
(2) *θ is one to one on Y_0.*
(3) *For any G invariant measure, μ on Y we have $\mu(Y_0) = 1$.*

In fact the set Y_0 will be of the form $\pi^{-1}(X_0)$ for some set $X_0 \subset X$ and 3. may be replaced by $\mu(X_0) = 1$ for any G invariant measure μ on X.

Here is an outline of the proof. Fix a point $y_0 \in Y$ with a dense orbit. Using it we shall construct a function f from Y to Y with some special properties, and then define a function F from Y to $X \times Y^G$ by the formula:

$$F(y) = (\pi(y), \omega(y))$$

where $\omega(y)$ is given by $\omega(y)(g) = f(S_g(y))$. This function F will not be continuous but it will be equivariant, where the action on $X \times Y^G$ is the product action of the given action on X with the shift on Y^G. The space $\hat{Y}$ will be the orbit closure of $F(y_0)$ in $X \times Y^G$, and the function θ will be the restriction of F to a set Y_0 which will be defined as $\pi^{-1}(X_0)$ where X_0 will also be constructed along with the function f.

In order to define the function f we will construct a subset $J \subset G$, and for $g \in J$ clopen neighborhoods U_g of $T_g(x_0)$ that are pairwise disjoint, and such that their union $U = \bigcup_{g \in J} U_g$ includes the entire orbit of x_0. In terms of these sets the function f is defined by the formula

$$f(y) = \begin{cases} y \text{ if } \pi(y) \notin U, & ; \\ S_g(y_0) \text{ if } \pi(y) \in U_g \text{ for some } g \in J, & . \end{cases}$$

The set J will be constructed as a disjoint union $J = \bigcup_{1 \le m} J_m$ together with an auxiliary sequence of sets I_m and will have a number of additional properties that will be formulated in terms of a sequence of symmetric sets $e \in B_n \subset G$ that increase to all of G and satisfy the following condition: B_{n+1} contains the disjoint union of B_n^2 and a translate of B_n say, $B_n a_n$.

P1.: For all m the set I_{m+1} is disjoint from $\bigcup_{k \le m} J_m$ and for each m the sets $B_{m+1}h$ for $h \in I_m$ are pairwise disjoint. The set J_m is defined to be all bh with $b \in B_m$ and $h \in I_m$ such that

$$T_{bh}(x_0) \notin \bigcup_{l<m} \bigcup_{g \in J_l} U_g.$$

P2.: For a fixed m and elements $g, h \in J_m$ the neighborhoods U_g, U_h satisfy

$$U_h = T_h T_{g^{-1}} U_g.$$

P3.: For any $g \in I_m$ and $b \in B_m$ either $bg \in J_m$ or $T_b T_g(x_0) \in U_h$ for some $h \in \bigcup_{i<m} J_i$.

P4.: For fixed m and $g = bh \in J_m$ with $b \in B_m$ and $h \in I_m$, the sets $\{T_{b'} T_{b^{-1}}(U_g) : b' \in B_{m+1}\}$ are disjoint.

P5.: For each fixed m the union

$$V_m = \bigcup_{b \in B_m, g \in J_m} T_b(U_g)$$

has enough disjoint translates in X so as to force its $\mu(V_m) < \frac{1}{10^m}$ for any G invariant measure μ on X.

P6.: If $y \notin \bigcup_{k \ge m} \pi^{-1}(V_k)$ then there is some $c \in I_m$ such that $S_c(y_0)$ is so close to y that for all $b \in B_m$ we will have that

$$d(f(S_b S_c(y_0)), f(S_b(y))) < \frac{1}{m}$$

where d is the metric on Y.

Here is brief explanation of how these properties enable one to prove the theorem. Following this explanation we will proceed to a detailed construction of the function f. The fact that there is only a single point in the orbit closure of $F(y_0)$ that projects onto x_0 follows easily form the definition of f and the fact that the open set U contains the entire G orbit of x_0. Properties **P5** and **P6** together show that for a set $X_0 \subset X$ that has full measure for any G invariant measure, all $x \in X_0$ and any $y \in \pi^{-1}(x)$ we can find a sequence of elements $c_n \in G$ such that $F(S_{c_n}(y_0))$ will converge to $F(y)$. In addition **P1** will be used to show that F is a one to one mapping on each such fiber $\pi^{-1}(x)$. A detailed argument will be given later on.

Putting this together we will get that the restriction of F to $\pi^{-1}(X_0)$ will be a one to one equivariant mapping which can serve as the θ in the theorem.

The construction of the sets I_m, J_m and the clopen neighborhoods U_g is carried out inductively, and the function f will be a pointwise limit of functions f_m which are constructed at the same time. We assume that we have already constructed these objects for all $k < m$ with the necessary properties and show how to construct the objects for m. The inductive properties that we assume have already been achieved are a more detailed version of the properties **P1 - P6** above:

P$_k$**1.:** The set I_k is disjoint from $\bigcup_{l<k} J_l$ and will contain the first element g of G in some fixed ordering of G such that $T_g(x_0) \notin \bigcup_{l<k} \bigcup_{g\in J_l} U_g$. The sets $B_{k+1}I_k$ are pairwise disjoint and the sets J_k are defined to be all elements of the form bh with $b \in B_k$ and $h \in I_k$ such that

$$T_{bh}(x_0) \notin \bigcup_{l<k} \bigcup_{g\in J_l} U_g.$$

P$_k$**2.:** For $g, h \in J_k$ the neighborhoods U_g, U_h satisfy

$$U_h = T_h T_{g^{-1}} U_g$$

and the $\{U_g : g \in \bigcup_{l\leq k} J_l\}$ are pairwise disjoint.

P$_k$**3.:** For any $g \in I_k$ and $b \in B_k$ either $bg \in J_k$ or $T_b T_g(x_0) \in U_h$ for some $h \in \bigcup_{l<k} J_l$.

P$_k$**4.:** The sets $\{U_g : g \in J_k\}$ are pairwise disjoint and disjoint from $\bigcup_{l<k} \bigcup_{h\in J_l} U_h$ and in addition the sets $\{T_{b'} T_{b^{-1}}(U_g) : b' \in B_{k+1}\}$ are pairwise disjoint for any fixed $g = bh \in J_k$, with $b \in B_k$ and $h \in I_k$.

P$_k$**5.:** The union

$$V_k = \bigcup_{b\in B_k, g\in J_k} T_b(U_g)$$

has enough disjoint translates in X so as to force $\mu(V_k) < \frac{1}{10^k}$ for any G invariant measure μ on X.

P$_k$**6.:** Let α_k be the minimum distance between the elements of the partition of X defined by the sets $\{U_h : h \in \bigcup_{l<k} J_l\}$ and the complement of their union. Let $\delta_k < \frac{1}{k}$ be small enough so that $d(y, y') < \delta_k$ implies $d(\pi(y), \pi(y')) < \alpha_k$. For any $y \in Y$ there is some some $c \in I_k$ such that for all $b \in B_k$ we will have that

$$d(S_b S_c(y_0), S_b(y)) < \delta_k.$$

The function f_k is defined as follows:

$$f_k(y) = \begin{cases} y \text{ if } \pi(y) \notin \bigcup_{l\leq k} \bigcup_{g\in J_l} U_g, & ; \\ S_g(y_0) \text{ if } \pi(y) \in U_g \text{ for some } g \in \bigcup_{l\leq k} J_l, & . \end{cases}$$

With this definition, if $\pi(y) \notin V_k$ then for the c that **P**$_k$**6** guarantees we can compare $f_k(S_b S_c(y_0))$ with $f_k(S_b(y))$ by analyzing where $T_b(x)$ is for $x = \pi(y)$. From the definition of α_k and δ_k it follows that either both $T_b(x)$ and $T_b T_c(x_0)$ belong to the same U_g with $g \in \bigcup_{l<k} J_l$ or both are in the complement of their union. In the former case by the definition of f_k it will follow that $f_k(S_b S_c(y_0)) = f_k(S_b(y)) = S_g(y_0)$. In the latter case, our assumption that $\pi(y) \notin V_k$ precludes the

possibility that $\pi(S_b(y))$ belongs to $\bigcup_{g\in J_k} U_g$ and so we would have that $f_k(S_b(y)) = S_b(y)$ while $f_k(S_bS_c(y_0)) = S_bS_c(y_0)$, and then $\mathbf{P}_k\mathbf{6}$ guarantees that

$$d(f_k(S_bS_c(y_0)), f_k(S_b(y))) < \frac{1}{k}.$$

Thus this equation holds for all $b \in B_k$ in any case.

Furthermore, the same analysis using the properties $P_l\mathbf{1}$-$P_l\mathbf{6}$ for $l < k$ will show that if $\pi(y) \notin \bigcup_{l\le r\le k} V_r$ then using the $c_l \in I_l$ that was guaranteed earlier we would still have

$$d(f_k(S_bS_{c_l}(y_0)), f_k(S_b(y))) < \frac{1}{l}$$

for all $b \in B_l$. In more detail, this last equation certainly holds for f_l, and by property $P_l\mathbf{1}$ we know that for all $h \in B_lI_l$ we have $T_h(x_0) \in \bigcup_{m\le l}\bigcup_{g\in J_m} U_g$. Thus $f_k(S_bS_{c_l}(y_0)) = f_l(S_bS_{c_l}(y_0))$ and since $\pi(y) \notin \bigcup_{l\le r\le k} V_r$ we also have that $f_k(S_b(y)) = f_l(S_b(y))$ from which the above follows.

If we then would define f to be the pointwise limit of the f_k's then this limiting function will have the properties that we described above in the outline of the proof. It remains to explain the inductive step - that is how to get the new I_{k+1}, J_{k+1} and U_g's. Begin by letting α_{k+1} be the minimum distance between the elements of the partition of X defined by the sets $\{U_h : h \in \bigcup_{l\le k} J_l\}$ and the complement of their union. Let $\delta_{k+1} < \frac{1}{k+1}$ be small enough so that $d(y,y') < \delta_{k+1}$ implies $d(\pi(y),\pi(y')) < \alpha_{k+1}$. Next use the fact that the orbit of y_0 is dense in Y to define I_{k+1} to be a finite subset of G that satisfies:

- **(i):** The set I_{k+1} is disjoint from $\bigcup_{l\le k} J_l$ and contains the first element g in G (according to a fixed enumeration of G) such that $T_g(x_0) \notin \bigcup_{l\le k}\bigcup_{g\in J_l} U_g$ is not in $\bigcup_{l\le k} J_l$.
- **(ii):** The balls with radius δ'_{k+1} centered at the points $\{S_c(y_0) : c \in I_{k+1}\}$ cover Y. Here δ'_{k+1} is chosen to be small enough so that $d(y,y') < \delta'_{k+1}$ implies $d(S_by, S_by') < \delta_{k+1}$ for all $b \in B_{k+1}$.
- **(iii):** The sets $B_{k+2}c$ for $c \in I_{k+1}$ are pairwise disjoint.

Define J_{k+1} to be all elements $bc \in G$ with $b \in B_{k+1}$ and $c \in I_{k+1}$ such that $T_{bc}(x_0) \notin \bigcup_{l\le k}\bigcup_{g\in J_l} U_g$ and set

$$\hat{J}_{k+1} = B_{k+2}I_{k+1}.$$

Let $H_{k+1} = \{h_n : 1 \le n \le 10^{k+1}\}$ be a finite subset of G such that the sets $h_n\hat{J}_{k+1}$ are pairwise disjoint.

To define the new clopen neighborhoods start with a fixed $g_0 \in J_{k+1}$ and choose a clopen neighborhood U_{g_0} so small that for all $g \in \hat{J}_{k+1}$ the sets $T_{gg_0^{-1}}U_{g_0}$ are pairwise disjoint, and are either disjoint from $\bigcup_{l\le k}\bigcup_{g\in J_l} U_g$ or are entirely contained in a single such U_g. In addition we want the neighborhoods to be small enough so that if we define

$$V_{k+1} = \bigcup_{b\in B_{k+1}, g\in J_{k+1}} T_b(U_g)$$

then the sets h_nV_{k+1} for $h_n \in H_{k+1}$ are pairwise disjoint. In this last formula for $g \in J_{k+1}$ we define $U_g = T_{gg_0^{-1}}U_{g_0}$. Note that by the definition of J_{k+1} these sets are disjoint from $\bigcup_{l\le k}\bigcup_{g\in J_l} U_g$ and satisfy of course that $T_g(x_0) \in U_g$. The size of

H_{k+1} guarantee that for any G invariant measure μ on X we have

$$\mu(V_{k+1}) < \frac{1}{10^{k+1}}.$$

Finally we can define f_{k+1} by the formula:

$$f_{k+1}(y) = \begin{cases} y \text{ if } \pi(y) \notin \bigcup_{l\leq k+1} \bigcup_{g\in J_l} U_g, & ; \\ S_g(y_0) \text{ if } \pi(y) \in U_g \text{ for some } g \in \bigcup_{l\leq k+1} J_l, & . \end{cases}$$

It is completely straightforward to check that with these definitions all of the properties $\mathbf{P}_{k+1}\mathbf{1}$ - $\mathbf{P}_{k+1}\mathbf{6}$ are satisfied. The fact that the space $\hat{Y}$ is minimal is a consequence of the fact that it is defined as the orbit closure of a point, $F(y_0)$, on whose orbit the projection is one to one and the factor (X, T_g) is minimal. Indeed any orbit closure in $\hat{Y}$ must project onto all of X and therefore must contain $F(y_0)$ and hence all of $\hat{Y}$. The set X_0 is defined by

$$X_0 = X \backslash \bigcap_k \bigcup_{m>k} V_m.$$

and $Y_0 = \pi^{-1}(X_0)$. These sets have full measure for any invariant measures on X and Y respectively. If $\pi(y) = x \in X_0$ then $\mathbf{P6}$ will show that $\hat{Y}$ contains $F(y)$. Furthermore, F will be one to one on this fiber. To see this last point we observe that we just have to check that for $x \in X_0$ there is some $g \in G$ with $T_g(x) \notin U$. Now if $x \in X_0$ then for some k we have $x \notin V_m$ for all $m > k$. This implies that $T_b(x) \notin \bigcup_{m>k} \bigcup_{h\in J_m} U_h$ for all $b \in B_{k+1}$. Now the orbit section $T_b(x)$ for $b \in B_{k+1}$ cannot be covered by the sets U_g for $g \in \bigcup_{l\leq k} J_l$. This is easily seen as follows. If some point in the orbit section above lies in a set U_g for some $g \in J_k$ then recalling that B_{k+1} contains a full translate of B_k, namely what we called $B_k a_k$, that is disjoint from B_k^2 and using $\mathbf{P}_k\mathbf{1}$ and $\mathbf{P}_k\mathbf{4}$ we would find in that orbit section of B_{k+1} a full orbit section of B_k that is in the complement of $\bigcup_{g\in J_k}$. We can now continue with a downward induction on k to find some $b \in B_{k+1}$ with $T_b(x) \notin U$ and this is enough to guarantee the required injectivity. Together with our earlier discussion this completes the proof of the theorem.

Remark 5.3. In case there is a factor intermediate between (Y, S_g) and (X, T_g), say (Z, R_g), with equivariant projections $\pi_2 : Y \to Z$ and $\pi_1 : Z \to X$ such that $\pi = \pi_1\pi_2$ then with this construction there is a naturally defined intermediate factor $\hat{Z}$ between $\hat{Y}$ and X that is also a minimal almost $1-1$ extension of X that captures the entire invariant measure structure of (Z, R_g).

6. Minimal models for free actions

Using what we have already done we can now easily find a minimal model for any measure preserving free action of a countable group G. Even for amenable groups this result is new. While for ergodic actions there is an analogue of the Jewett-Krieger theorem which provides uniquely ergodic models and they are of course minimal for non ergodic free actions this result simply doesn't apply. For the case of $\mathbb{Z}$ I constructed in [**W**] a minimal system which was a model for all ergodic non-periodic actions [1], however this kind of construction has not been carried out for a general amenable group. Once again we start with a fixed free action of

[1] The theorem is stated there for free actions but the proof is carried out in detail only for ergodic actions.

a countably infinite group $(X, \mathcal{B}, \mu, T_g)$. It is well known that we may assume that X is a compact metric space and that the action is by homeomorphisms and that furthermore there is a dense orbit. To see this let $\theta : X \to [0,1]$ be a Borel measurable mapping from X to the unit interval that is one to one and onto, and let $\Theta : X \to [0,1]^G$ be defined by $\Theta(x)(g) = \theta(T_g(x))$. This is an equivariant mapping from X to $[0,1]^G$ with the shift action , which is a Borel measurable bijection and $\nu = \Theta \circ \mu$ is a shift invariant measure such that $([0,1]^G, \mathcal{B}, \nu, \sigma_g)$ is isomorphic to $(X, \mathcal{B}, \mu, T_g)$. To avoid new notation we assume henceforth that $X = [0,1]^G, \nu = \mu$, and $T_g = \sigma_g$.

Letting U be the universal subshift in Remark 4.6, form the space $Y = U \times X$ with the product action which we will denote by S_g. Let π denote the projection onto the first coordinate. Since the second coordinate is the full shift space and the action on the first coordinate is minimal it is easy to see that there is a $y_0 \in Y$ with a dense orbit that sits above a point with a free orbit. We can apply theorem 5.2 to get a minimal extension $\hat{Y}$ of U with the properties stated in the theorem.

I claim that this space is a universal minimal space. To see this we will construct a measure λ on Y that is invariant under the action S_g and is a natural way isomorphic to the original action of G on $(X, \mathcal{B}, \mu)$ that we started with. As before we denote by Φ the mapping that we constructed from X to U, and we map X to Y by sending x to the pair $(\Phi(x), x)$. This map is one to one and equivariant and takes μ to the required measure λ. According to the theorem this invariant measure can be mapped to the space of the minimal extension $\hat{Y}$ and demonstrates the universality of $\hat{Y}$. We state this as:

THEOREM 6.1. *For any infinite countable group G there is a minimal action of G on a subshift $\hat{Y} \subset ([0,1] \times I_\infty)^G$, and for any free action $(X, \mathcal{B}, \mu, T_g)$ there is an invariant measure $\hat{\mu}$ on $\hat{Y}$ such that the action $(\hat{Y}, \mathcal{B}, \hat{\mu}, \sigma_g)$ is measure theoretically isomorphic to the action $(X, \mathcal{B}, \mu, T_g)$.*

References

[FW] H. Furstenberg and B. Weiss, *On almost 1-1 extensions*, Israel J. of Mathematics **65** (1989), 311-322. MR1005015 (90g:28020)

[GW] E. Glasner and B. Weiss, *On the interplay between measurable and topological dynamics*, Handbook of dynamical systems. **1**B, 597–648, Elsevier B. V., Amsterdam, 2006. MR2186250 (2006i:37005)

[W] B. Weiss, *Countable generators in Dynamics - Universal Minimal Models in Measure and Measurable Dynamics*, Measure and measurable dynamics (Rochester, NY, 1987), 321–326, Contemp. Math., **9**4, Amer. Math. Soc., 1989. MR1013000 (90g:28025)

HEBREW UNIVERSITY OF JERUSALEM, JERUSALEM, ISRAEL 91904
E-mail address: weiss@math.huji.ac.il

Selected Titles in This Series

567 **Lewis Bowen, Rostislav Grigorchuk, and Yaroslav Vorobets, Editors,** Dynamical Systems and Group Actions, 2012

566 **Antonio Campillo, Gabriel Cardona, Alejandro Melle-Hernández, Wim Veys, and Wilson A. Zúñiga-Galindo, Editors,** Zeta Functions in Algebra and Geometry, 2012

565 **Susumu Ariki, Hiraku Nakajima, Yoshihisa Saito, Ken-ichi Shinoda, Toshiaki Shoji, and Toshiyuki Tanisaki, Editors,** Algebraic Groups and Quantum Groups, 2012

564 **Valery Alexeev, Angela Gibney, Elham Izadi, János Kollár, and Eduard Looijenga, Editors,** Compact Moduli Spaces and Vector Bundles, 2012

563 **Primitivo B. Acosta-Humánez, Federico Finkel, Niky Kamran, and Peter J. Olver, Editors,** Algebraic Aspects of Darboux Transformations, Quantum Integrable Systems and Supersymmetric Quantum Mechanics, 2012

562 **P. Ara, K. A. Brown, T. H. Lenagan, E. S. Letzter, J. T. Stafford, and J. J. Zhang, Editors,** New Trends in Noncommutative Algebra, 2012

561 **Óscar Blasco, José A. Bonet, José M. Calabuig, and David Jornet, Editors,** Topics in Complex Analysis and Operator Theory, 2012

560 **Weiping Li, Loretta Bartolini, Jesse Johnson, Feng Luo, Robert Myers, and J. Hyam Rubinstein, Editors,** Topology and Geometry in Dimension Three, 2011

559 **Guillaume Bal, David Finch, Peter Kuchment, John Schotland, Plamen Stefanov, Gunther Uhlmann, and Gunther Uhlmann, Editors,** Tomography and Inverse Transport Theory, 2011

558 **Martin Grohe and Johann A. Makowsky, Editors,** Model Theoretic Methods in Finite Combinatorics, 2011

557 **Jeffrey Adams, Bong Lian, and Siddhartha Sahi, Editors,** Representation Theory and Mathematical Physics, 2011

556 **Leonid Gurvits, Philippe Pébay, J. Maurice Rojas, and David Thompson, Editors,** Randomization, Relaxation, and Complexity in Polynomial Equation Solving, 2011

555 **Alberto Corso and Claudia Polini, Editors,** Commutative Algebra and Its Connections to Geometry, 2011

554 **Mark Agranovsky, Matania Ben-Artzi, Greg Galloway, Lavi Karp, Simeon Reich, David Shoikhet, Gilbert Weinstein, and Lawrence Zalcman, Editors,** Complex Analysis and Dynamical Systems IV, 2011

553 **Mark Agranovsky, Matania Ben-Artzi, Greg Galloway, Lavi Karp, Simeon Reich, David Shoikhet, Gilbert Weinstein, and Lawrence Zalcman, Editors,** Complex Analysis and Dynamical Systems IV, 2011

552 **Robert Sims and Daniel Ueltschi, Editors,** Entropy and the Quantum II, 2011

551 **Jesus Araujo-Gomez, Bertin Diarra, and Alain Escassut, Editors,** Advances in Non-Archimedean Analysis, 2011

550 **Y. Barkatou, S. Berhanu, A. Meziani, R. Meziani, and N. Mir, Editors,** Geometric Analysis of Several Complex Variables and Related Topics, 2011

549 **David Blázquez-Sanz, Juan J. Morales-Ruiz, and Jesús Rodríguez Lombardero, Editors,** Symmetries and Related Topics in Differential and Difference Equations, 2011

548 **Habib Ammari, Josselin Garnier, Hyeonbae Kang, and Knut Sølna, Editors,** Mathematical and Statistical Methods for Imaging, 2011

547 **Krzysztof Jarosz, Editor,** Function Spaces in Modern Analysis, 2011

546 **Alain Connes, Alexander Gorokhovsky, Matthias Lesch, Markus Pflaum, and Bahram Rangipour, Editors,** Noncommutative Geometry and Global Analysis, 2011

545 **Christian Houdré, Michel Ledoux, Emanuel Milman, and Mario Milman, Editors,** Concentration, Functional Inequalities and Isoperimetry, 2011

544 **Carina Boyallian, Esther Galina, and Linda Saal, Editors,** New Developments in Lie Theory and Its Applications, 2011

CONM 567

Dynamical Systems and Group Actions • Bowen et al., Editors

This volume contains cutting-edge research from leading experts in ergodic theory, dynamical systems and group actions. A large part of the volume addresses various aspects of ergodic theory of general group actions including local entropy theory, universal minimal spaces, minimal models and rank one transformations. Other papers deal with interval exchange transformations, hyperbolic dynamics, transfer operators, amenable actions and group actions on graphs.

ISBN 978-0-8218-6922-2

9 780821 869222

CONM/567

AMS *on the* Web
www.ams.org

AMS