AN ENTROPY-PRESERVING DYE'S THEOREM FOR ERGODIC ACTIONS

DANIEL RUDOLPH

Abstract. This work shows that equality of entropy for ergodic actions of a discrete amenable group is a restricted orbit equivalence in the formal sense defined in *Restricted orbit equivalence for actions of discrete amenable groups* by Kammeyer and Rudolph [3]. An element of the full-group of such an action encodes a countable partition. A natural extension of entropy to such countable partitions is shown to be a size in the sense of [3] and hence engenders an orbit equivalence relation on the space of all such actions. The major goal achieved here is to show that this relation is precisely equality of entropy.

1. Introduction

In one of the most fundamental results in ergodic theory Dye [2] has shown that all ergodic actions of $\mathbb{Z}$ on a nonatomic space are orbit equivalent. As a significant strengthening of this fact Connes, Feldman and Weiss showed that a free and ergodic action of a discrete group is orbit equivalent to an action of $\mathbb{Z}$ if and only if the group is amenable [1]. In [3] a general approach is presented to the structure of equivalence relations which at their most basic are orbit equivalences but are subject to some further restriction. This restriction is described through a perturbation theory. A description of that theory is included here. The goal of this work is to show that the equivalence relation given by equality of entropy is a restricted orbit equivalence. The path to this result breaks into a series of steps.

First we modify the classical notion of entropy for a process (T, P) so that it applies naturally to countable partitions and in particular to the natural countable partitions generated by an element of the full-group of T. This modification is constructed for the purposes of our work here and is not intended to be of broader value. With this in hand the first major step is Theorem 2.9 where it is shown that from it one obtains a size m^e in the formal sense defined in [3]. The next step is Corollary 3.3 where it is shown that m^e is an entropy preserving size which is to say m^e-equivalence preserves entropy.

Identifying m^e-equivalence as precisely equality of entropy involves significant work much of which concerns the basic theory of restricted orbit equivalence. In [3] the notions of m-finitely determined and weakly m-finitely determined are presented with the weak version only known to imply the stronger one in the case of 3^+-sizes. Here we introduce a twist on the form of weakly m-finitely determined, calling it **reverse weakly m-finitely determined**. We show in Corollary 3.17 that so long as there exists an m-finitely determined action of a given entropy this reverse condition is equivalent to m-finitely determined. This is critical to our goal as we characterize m^e-equivalence by showing all actions are reverse weakly m^e-finitely determined. This final fact is built upon a basic piece of combinatorics. A name can

---

Date: November 22, 2003.

be modified to any other name with the same symbol counts by permuting the placement of the symbols in the name. The issue we face is to control the number of permutations necessary to change one distribution on names into another where these distributions are known to be close in “entropy and distribution”. Stated simply the answer is that one only needs an exponentially small number of permutations.

Combining this with the now classical Rokhlin lemma for discrete amenable groups of Ornstein and Weiss [4] and a copying lemma with no really new tricks in it one completes the argument that all actions of finite entropy are reverse weakly m^e-finitely determined and hence at least for finite entropy actions, entropy is a complete invariant of m^e-equivalence. We extend the general restricted orbit equivalence theory here by showing that even for infinite entropy, among the m-f.d. actions, m-entropy is a complete invariant of m-equivalence. We conclude with some open issues and questions generated by this work.

2. Making entropy into a size

Entropy has always been used in dynamics as a measure of randomness or complexity. Here we will use it in precisely that sense in the context of restricted orbit equivalence as a measure of the randomness or complexity of a rearrangement of the orbits of a system. Our goal is to study when it is possible for this complexity to be small. The precise entropy is not what is critical to us, but rather that we have a topology on rearrangments for which a notion of entropy is a metric. This gives us some freedom in our definitions.

Our first task is to establish a reasonable theory of entropy for ergodic actions and countable partitions that eliminates the enherent infinities that can arise. Let T be an ergodic action of a countable and discrete amenable group G on $(X, \mathcal{F}, \mu)$, a standard probability space. We will use the vocabulary of [3] of arrangements and rearrangements. To understand the work here in detail the reader must be familiar with the approach developed in [3]. Later we will write T^α where α is an arrangment, when we must focus on various arrangments and rearrangements of orbits.

For a countable labeling set Σ, a map $P : X \to \Sigma$ is a Σ-valued partition. If in fact the range of P is a finite subset of Σ then $h(T, P)$ is well defined in particular via the Ornstein Weiss theory [4]. Let $\mathcal{P}_\Sigma$ be the space of all Σ valued partitions which we topologize as a Polish space via a variant of the partition metric

$$d_0(P_1, P_2) = \mu(\{x : P_1(x) \neq P_2(x)\}).$$

Definition 2.1. *For P, P' two Σ-valued partitions of X set $R(P, P')$ to be the 2-set partition of X according to whether $P(x) = P'(x)$ or not and now define a generalized entropy by*

$$E(T, P) = \inf_{P' : X \to \Sigma} (h(T, P \vee R(P, P')) + d_0(P, P')).$$

If we wish to stress the measure used, or there are many measures on the space, we will write $E_\mu(T, P)$.

We now show that $E(T, P)$ and $h(T, P)$ are proportional when P is finite with the constant of proportionality depending on the cardinality of the essential range of P.

Lemma 2.2. *If Σ is a finite labeling set then for any ergodic $(X, \mathcal{F}, \mu, T)$ and $P : X \to \Sigma$,*

$$\frac{h(T,P)}{1+\log_2(\#\Sigma)} \leq E(T,P) \leq h(T,P).$$

Proof. The right inequality is trivial. For the left note that if $\alpha = \mu(\{x : P(x) \neq P'(x)\})$ then

$$h(T,P|P' \vee R(P,P')) \leq \alpha \log(\#\Sigma)$$

and $h(T,P' \vee R(P,P') + d_0(P,P'))$ is bounded below by both $h(T,P) - \alpha\log(\#\Sigma)$ and α where $\alpha \geq 0$. These two two values cross at $\alpha_0 = \dfrac{h(T,P)}{1+\log(\#\Sigma)}$ yielding the result. □

Corollary 2.3. *If $P : X \to \Sigma$ is supported on at most k symbols then*

$$\frac{h(T,P)}{(1+\log_2(k))} \geq E(T,P) \geq h(T,P).$$

Proof. Any $P' : X \to \Sigma$ can be replaced by a P'' taking on only values among the symbols in the range of P with

$$h(T,P'') + d_0(P,P') \leq h(T,P') + d_0(P,P'),$$

i.e. the inf is not changed by restricting to those P'' that are range(P)-valued □

As $E(T,P) \leq 1$ universally, it provides a bounded calculation we will use in place of entropy.

The following simple lemma lifts two basic facts about entropy to E and will be important to us in verifying the axiomatics of a size later.

Lemma 2.4. *For $P : X \to \Sigma_1$ and $Q : X \to \Sigma_2$, two countable partitions,*

$$E(T,P \vee Q) \leq E(T,P) + E(T,Q).$$

Proof. Fix an $\varepsilon > 0$ and select P' and Q' with

$$E(T,P') \geq h(T,P' \vee R(P,P')) + d_0(P,P') - \varepsilon/2 \text{ and}$$
$$E(T,Q') \geq h(T,Q' \vee R(Q,Q')) + d_0(Q,Q') - \varepsilon/2.$$

Then of course

$$\begin{aligned} E(T,P \vee Q) &\leq h(T,(P' \vee Q') \vee R(P' \vee Q', P \vee Q)) + d_0(P \vee Q, P' \vee Q') \\ &\leq h(T,(P' \vee Q') \vee R(P,P') \vee R(Q,Q')) + d_0(P,P') + d_0(Q,Q') \\ &\leq E(T,P) + E(T,Q) + \varepsilon. \end{aligned}$$

□

For Σ a countable labeling set, Σ^G with the shift action σ is a Polish G-action. Hence $\mathcal{M}_e(\Sigma^G)$, the space of σ-invariant and ergodic Borel measures on Σ^G is Polish in the weak*-topology (see [5] and [3]). Let D be a metric on $\mathcal{M}_e(\Sigma^G)$ giving the weak* topology. For the case of interest to us, where $\Sigma = G$, we give an explicit form for D later. Now Σ^G has a canonical Σ-valued partition, the "time identity" partition, which we write P_0. We now can regard E as defined on $\mathcal{M}_e(\Sigma^G)$ by writing

$$E(\mu) = E_\mu(\sigma, P_0).$$

Theorem 2.5. *The map* $E : \mathcal{M}_e(\Sigma^G) \to [0,1]$ *is upper semi-continuous, i.e. for any* $\mu_0 \in \mathcal{M}_e(\Sigma^G)$ *and* $\varepsilon > 0$ *there is a* $\delta > 0$ *so that if* $D(\mu_0, \mu) < \delta$ *then* $E(\mu) < E(\mu_0) + \varepsilon$.

Proof. For $\mu_0 \in \mathcal{M}_e(\Sigma^G)$ and $\varepsilon > 0$ we can find a Borel P' with

$$h_{\mu_0}(\sigma, P' \vee R(P_0, P')) + \mu_0(\{x : P_0(x) \neq P'(x)\}) < E(T, P) + \varepsilon/2.$$

In Σ^G, clopen partitions are d_0-dense as are finite valued partitions. Hence we can choose P' here to be a finite and clopen partition. It follows that $h_\mu(\sigma, P' \vee R(P_0, P'))$ is upper semi-continuous as a function of μ. Of course $\mu(\{x : P_0(x) \neq P'(x)\})$ is continuous as a function of μ. Hence there is a δ so that if $D(\mu_0, \mu) < \delta$ then

$$E(\mu) \leq h_\mu(\sigma, P' \vee R(P_0, P') + \mu(\{X : P_0(x) \neq P'(x)\}) < E(\mu_0) + \varepsilon.$$

□

For ϕ an element of the full group $\Gamma(T)$ of a G-action T^α, α the corresponding G-arrangement, we now associate to the rearrangement pair (α, ϕ) a notion of "entropy". Notice following the notation of Chapter 2 of [3] we can associate to (α, ϕ) a G-valued name

$$f_x^{\alpha,\phi}(g) = \alpha(T_g^\alpha(x), \phi(T_g^\alpha(x))$$

whose time-identity partition is simply

$$P^{\alpha,\phi}(x) = \alpha(x, \phi(x)),$$

a countable partition.

Definition 2.6. *Define the* **entropy** *of a rearrangement pair* (α, ϕ) *to be*

$$E(\alpha, \phi) = E(T^\alpha, P^{\alpha,\phi}).$$

Notice that in defining $E(\alpha, \phi)$ in this form one minimizes $h(T^\alpha, P' \vee R(P^{\alpha,\phi}, P')) + d_0(P^{\alpha,\phi}, P')$ where P' need not be of the form $P^{\alpha,\phi'}$. That is to say, one considers all perturbations of the partition $P^{\alpha,\phi}$ not all perturbations of the full-group element ϕ. We will find it instructive and useful to know that restricting to just those P' of the form $P^{\alpha,\phi'}$ which are finite valued leads to the same value for E. Let Γ generically represent the full-group of an action and let $\Gamma'(\alpha)$ represent those for which $P^{\alpha,\phi}$ is a finite partition . To start make the definition

$$E'(\alpha, \phi) = \inf_{\phi' \in \Gamma'(\alpha)} (h(T^\alpha, P^{\alpha,\phi'} \vee R(P^{\alpha,\phi}, P^{\alpha,\phi'})) + d_0(P^{\alpha,\phi}, P^{\alpha,\phi'})).$$

As E' restricts the partitions over which the infimum is taken, $E' \geq E$. Notice that any partition P' taking values in G defines a map $\phi^P(x) = T_{P(x)}^\alpha(x)$. This map may fail to be in Γ though as it need be neither 1-1 nor onto. It is this issue precisely that we are addressing in the following theorem.

Theorem 2.7. *For* $(T^\alpha, X, \mathcal{F}, \mu)$ *a free and ergodic action of the countably infinite and discrete amenable group* G *and for* ϕ *in the full-group of* T^α,

$$E(\alpha, \phi) = E'(\alpha, \phi).$$

Proof. Given $\varepsilon > 0$ it is sufficient to show that there is a $\phi' \in \Gamma'(\alpha)$ with

$$E(\alpha, \phi) + \varepsilon \geq h(T^\alpha, P^{\alpha,\phi'} \vee R(P^{\alpha,\phi}, P^{\alpha,\phi'})) + d_0(P^{\alpha,\phi}, P^{\alpha,\phi'}).$$

Select a finite partition P' taking values in G with

$$E(\alpha, \phi) + \varepsilon/3 \geq h(T^\alpha, P' \vee R(P^{\alpha,\phi}, P')) + d_0(P^{\alpha,\phi}, P').$$

Let $A = \{x : P^{\alpha,\phi}(x) = P'(x)\}$. Select a σ invariant factor algebra $\mathcal{H}$ with $P' \vee R(P^{\alpha,\phi}, P') \subseteq \mathcal{H}$, the action of T^α restricted to $\mathcal{H}$ is free and

$$h(T^\alpha, \mathcal{H}) < h(T^\alpha, P' \vee R(P^{\alpha,\phi}, P')) + \varepsilon/3.$$

As T^α acts freely on $\mathcal{H}$, the Ornstein-Weiss Rokhlin lemma (Theorem 5, II.2 of [4] or Theorem 3.0.5 of [3]) allows us to build Rokhlin towers that are $\mathcal{H}$ measurable. Choose ε_1 with $\varepsilon/3 > \varepsilon_1 > 0$ so small that if $a \leq \varepsilon_1$ then $H(a) < \varepsilon/3$.

We will construct a list of potential ϕ''s call $\phi_1, \phi_2, \ldots, \phi_k \in \Gamma$ with all $P^{\alpha,\phi_j} \subseteq \mathcal{H}$ and for at least one of which,

$$d_0(R(P^{\alpha,\phi}, P'), R(P^{\alpha,\phi}, P^{\alpha,\phi_j})) < \varepsilon_1.$$

Having done this we let $\phi' = \phi_j$ and obtain

$$\begin{aligned} &h(T^\alpha, P^{\alpha,\phi'} \vee R(P^{\alpha,\phi}, P^{\alpha,\phi'})) \\ \leq &h(T^\alpha, \mathcal{H}) + h(R(P^{\alpha,\phi}, P^{\alpha,\phi'}) | R(P^{\alpha,\phi}, P')) \\ \leq &h(T^\alpha, \mathcal{H}) + H(\varepsilon_1) < E(\alpha, \phi) + \varepsilon. \end{aligned}$$

It remains to construct the $\phi_1, \ldots, \phi_k$. Following the construction in Theorem 4.0.2 of [3], build a Rokhlin tower $\{A_i \times F_i\}$ covering all but $\varepsilon/10$ in measure of X. The sets A_i are the bases of the various subtowers and the $F_i \subseteq G$ are the shapes of these subtowers. If $x \in A$ and both x and $\phi(x)$ are in the same tower slice, then set $\phi_j(x) = \phi(x) = \phi^P(x)$ for all j. For all x outside of the tower image, set $\phi_j(x) = x$ for all j. Let C be the set on which the ϕ_j remain undefined. Notice that both C and the ϕ_j so far as they have been defined are $\mathcal{H}$ measurable.

If the tiling sets A_i are sufficiently large and invariant we can assume the remaining set C on which the ϕ_j are undefined satisfy $\mu(C \triangle A^c) < \varepsilon/10$. For each $x \in A_i$, let $C_i(x) \subseteq F_i$ consist of those values $c \in F_i$ for which $T^\alpha_c(x) \in C$. Furthermore, let $D_i(x) \subseteq F_i$ consist of those values $d \in F_i$ for which $T^\alpha_d(x)$ is not in the range of ϕ_j. Notice that for a.e. $x \in A_i$ the sets $C_i(x)$ and $D_i(x)$ have the same cardinality and moreover the map $Q_i : x \to (C_i(x), D_i(x))$ is $\mathcal{H}$-measurable. If for each (C, D) in the range of Q_i we choose a map $U_{C,D} : C \to D$, then we can use U to extend ϕ_j in an $\mathcal{H}$ measurable way to all of C. We construct many ϕ_j in order to ensure one satisfies

$$d_0(R(P^{\alpha,\phi}, P'), R(P^{\alpha,\phi}, P^{\alpha,\phi_j})) < \varepsilon_1.$$

Let $k = [10/\varepsilon_1] + 1$ and if the A_i are sufficiently large and invariant we can assume the tower is chosen so that each tower slice contains at least k points x in C. Note that this assumes there is a set of positive measure where $P^{\alpha,\phi}(x) \neq P'(x)$ and otherwise we already have our ϕ' in ϕ.

Define $\phi_1, \dots, \phi_k$ by choosing various maps U as described above so that for any $s \neq t$ and $x \in C$, $\phi_s(x) \neq \phi_t(x)$. This is easily arranged by choosing one map U and then cyclically permuting the range by k steps to form the other choices for U. It follows that for each $x \in C$, at most one $j \in \{1, \dots, k\}$ satisfies $\phi(x) = \phi_j(x)$. Hence for at least one choice of j,

$$\mu(\{x \in C : \phi(x) = \phi_j(x)\}) \leq \frac{1}{k} < \varepsilon_1/10.$$

Hence

$$\begin{aligned}
&d_0(R(P^{\alpha,\phi}, P'), R(P^{\alpha,\phi}, P^{\alpha,\phi_j})) = \\
&\quad \mu(\{x : P^{\alpha,\phi}(x) \neq P^{\alpha,\phi_j}(x)\} \triangle A^c\}) \leq \\
&\quad \mu(\{x : P^{\alpha,\phi}(x) \neq P^{\alpha,\phi_j}(x)\} \triangle C\}) + \varepsilon_1/10 < \\
&\quad 3\varepsilon_1/10 < \varepsilon_1.
\end{aligned}$$

□

We now turn to converting E into a size as defined in [3]. As earlier we let Γ denote the full group of $\mathcal{O}$. Also let $\mathcal{A}$ denote the set of G-arrangments α on the underlying orbit space and $\mathcal{Q}$ denote the set of rearrangements (α, ϕ).

We remind the reader of a few definitions taken directly from [3]. The reader is referred there for further discussion. To define a size we first consider three pseudometrics on the set of rearrangements. These all arise from natural topologies on functions $G \to G$, that is to say on G^G. As G is countable the only reasonable topology is the discrete one, using the discrete 0,1 valued metric. This topologizes G^G as a metrizable space with the product topology. This is the weakest topology for which the evaluations $g : f \to f(g)$ are continuous functions. Notice that $H : G^G \to G^G$ given by $H(g)(g) = f(g)gf(\mathrm{id})^{-1}$ is a continuous map from G^G to itself and the map $h \to h^{-1}$ on $\mathcal{G}$ is continuous.

We follow the notation of Chapter 2 of [3] in using the three families of functions, all measurable in x:

$$\begin{aligned}
q_z^{\alpha,\phi} &= \alpha(x, \phi(T_g^\alpha(x))), \\
h_x^{\alpha,\beta} &= \beta(x, T_g^\alpha(x)) and \\
f_x^{\alpha,\phi}(g) &= \alpha(T_g^\alpha(x), \phi(T_g^\alpha(x))).
\end{aligned}$$

Define a metric d on $\mathcal{G}$ as follows. List the elements of G as $\{g_1 = \mathrm{id}, g_2, \dots\}$ and let d_0 be the 0,1 valued metric on G. Set

$$d(h_1, h_2) = \sum_i [d_0(h_1(g_i), h_2(g_i)) + d_0(h_1^{-1}(g_i), h_2^{-1}(g_i))]2^{-(i+1)}.$$

Notice that if h_1, h_2, h_1^{-1}, and h_2^{-1} agree on $g_1, \dots, g_i$ then $d(h_1, h_2) \leq 2^{-i}$. On the other hand if $d(h_1, h_2) < 2^{-i}$ then h_1, h_2 and their inverses agree on this list of i terms.

We can use this to define an L^1 metric on arrangements:

$$\|\alpha, \beta\|_1 = \int d(h^{\alpha,\beta}, \mathrm{id})\, d\mu.$$

As $d(h_1, h_2) = d(h_2^{-1}h_1, \mathrm{id})$ and $(h_x^{\alpha,\beta})^{-1} = h_x^{\beta,\alpha}$ we see that this is a metric.

We can also define a metric similar to d on G^G itself making it a complete metric space by just taking half of the terms in d:

$$d_1(f_1, f_2) = \sum_i d_0(f_1(g_i), f_2(g_i))2^{-i}.$$

This also leads to an L^1 metric on G^G-valued functions on a measure space:

$$\|f_1, f_2\|_1 = \int d_1(f_1, f_2)\, d\mu.$$

These two L^1 distances now give us two families of L^1 distances on the full-group, one a metric the other a pseudometric, associated with an arrangement α:

$$\begin{aligned}\|\phi_1, \phi_2\|_{\mathrm{w}}^{\alpha} &= \int d(h^{\alpha,\alpha\phi_1}, h^{\alpha,\alpha\phi_2})\, d\mu \\ &= \int d(h^{\alpha\phi_1,\alpha\phi_2}, \mathrm{id})\, d\mu = \|\alpha\phi_1, \alpha\phi_2\|_1 \\ &= \int d(H(f^{\alpha,\phi_1}), H(f^{\alpha,\phi_2}))\, d\mu,\end{aligned}$$

and

$$\begin{aligned}\|\phi_1, \phi_2\|_{\mathrm{s}}^{\alpha} &= \int d_1(f^{\alpha,\phi_1}, f^{\alpha,\phi_2})\, d\mu \\ &= \|f^{\alpha,\phi_1}, f^{\alpha,\phi_2}\|_1\end{aligned}$$

The **weak** L^1 distance, $\|\cdot,\cdot\|_{\mathrm{w}}^{\alpha}$, is only a pseudometric but the **strong** L^1 distance, $\|\cdot,\cdot\|_{\mathrm{s}}^{\alpha}$, is a metric. Notice that T-invariance of μ tells us

$$\begin{aligned}\|\phi_1, \phi_2\|_{\mathrm{s}}^{\alpha} &\le 2\int d_0(f^{\alpha,\phi_1}(\mathrm{id}), f^{\alpha,\phi_2}(\mathrm{id}))\, d\mu \\ &= 2\mu(\{x : \phi_1(x) \ne \phi_2(x)\}) \le 2\|\phi_1, \phi_2\|_{\mathrm{s}}^{\alpha}.\end{aligned}$$

Thus, in fact, the topology generated by the strong L^1 distance on the full-group is independent of the arrangement α.

To describe the weak*-distance between two arrangements let $G^\star = G \cup \{\star\}$ be the one point compactification of G. Now $(G^\star)^G$ is a compact metric space and hence the Borel probability measures on $(G^\star)^G$, which we write as $\mathcal{M}_1(G^\star)$, are a compact and convex space in the weak* topology (that is to say the topology induced on Borel measures as the dual of the continuous functions).

We put an explicit metric on this space as follows. For any finite subset $F \subseteq G$ and $f \in G^G$ let f_F be the restriction of f to F. As f_F can be one of at most a countable collection of values, $f \to f_F$ partitions G^G into a countable collection of clopen sets. If two measures on $(G^\star)^G$ agree on these sets, that is to say on all cylinder sets that do not have a $\star$ in their name, then they agree. Moreover the characteristic functions of these sets are continuous. Hence if $\mu_i(f_F) \to \mu(f_F)$ the $\mu_i \to \mu$ weak*. To turn this into a metric, let F_i be an increasing sequence of finite sets that exhaust G, for example a Følner sequence. For each F_i let $P(F_i)$ be the partition of G^G according to the values f_{F_i}. These partitions refine and for any fixed F, once $F \subseteq F_i$, f_F will be $P(F_i)$-measurable. Set

$$D(\mu_1, \mu_2) = \sum_i \Bigg(\sum_{p \in P(F_i)} |\mu_1(p) - \mu_2(p)| \Bigg) 2^{-(i+1)}.$$

Notice that

$$\sum_{p \in P(F_i)} |\mu_1(p) - \mu_2(p)|/2 \leq 1$$

and so the ith term in this sum is bounded by 2^{-i}. Moreover as the partitions $P(F_i)$ refine, the values

$$\sum_{p \in P(F_i)} |\mu_1(p) - \mu_2(p)|/2 \leq 1$$

increase. It follows that for all i

$$\sum_{p \in P(F_i)} |\mu_1(p) - \mu_2(p)| + 2^{-i} \geq D(\mu_1, \mu_2).$$

We now define the distribution pseudometric between two rearrangements by

$$\|(\alpha, \phi), (\beta, \psi)\|_* = D((f^{\alpha,\phi})^*(\mu), (f^{\beta,\psi})^*(\nu)).$$

We can combine the two L^1-metrics on arrangements and the full-group to define a product metric on rearrangements in the form

$$\|(\alpha_1, \phi_1), (\alpha_2, \phi_2)\|_1 = \|\alpha_1, \alpha_2\|_1 + \mu(\{x : \phi_1(x) \neq \phi_2(x)\}).$$

A size can now be defined as a function

$$m : \mathcal{Q} \to \mathbb{R}^+$$

such that, if we write

$$m_\alpha(\phi_1, \phi_2) \underset{\text{defn}}{=} m(\alpha\phi_1, \phi_1^{-1}\phi_2),$$

then m satisfies the following three axioms.

Axiom 1. *For each* $\alpha \in \mathcal{A}$, m_α *is a pseudometric on* Γ.

Axiom 2. *For each* $\alpha \in \mathcal{A}$, *the identity map*

$$(\Gamma, m_\alpha) \overset{\text{id}}{\to} (\Gamma, \|\cdot, \cdot\|_{\text{w}}^{\alpha})$$

is uniformly continuous.

In particular this means that if $m_\alpha(\phi_1, \phi_2) = 0$ then the two arrangements $\alpha\phi_1$ and $\alpha\phi_2$ are identical.

Axiom 3. *m is upper semi-continuous with respect to the distribution metric. That is to say, for every* $\varepsilon > 0$, *there exists* $\delta = \delta(\varepsilon, \alpha, \phi)$, *such that if* $\|(\alpha, \phi), (\beta, \psi)\|_* < \delta$ *then* $m(\beta, \psi) < m(\alpha, \phi) + \varepsilon$.

This last condition tells us that if the two measures $(f^{\alpha,\phi})^*(\mu)$ and $(f^{\beta,\psi})^*(\nu)$ are the same, then $m(\alpha,\phi) = m(\beta,\psi)$. Hence the value m is well defined on those measures on G^G which arise as such an image, and we can write

$$m(\alpha,\phi) = m((f^{\alpha,\phi})^*(\mu)).$$

This completes our background discussion of sizes. We now proceed to modify E slightly to be the size we want. To begin notice that both the weak and strong distribution metrics give us sizes in the forms

$$m_0(\alpha,\phi) = \|\mathrm{i}d, , \|_{\mathrm{w}}^{\alpha}\phi$$

and

$$m_1(\alpha,\phi) = \|\mathrm{i}d, , \|_{\mathrm{s}}^{\alpha}\phi.$$

One, by Axiom 2 is the weakest possible size and as m_1-equivalence is differing by a full group element, m_1 is the strongest. We note that that sum of two sizes is also automatically a size. With this we now define the size we will work with.

Definition 2.8. *For α an arrangement and ϕ in the full group set*

$$m^e(\alpha,\phi) = E(\alpha,\phi) + m_0(\alpha,\phi).$$

Theorem 2.9. *The map m^e on rearrangement pairs is a size.*

Proof. Axiom 2 has been guaranteed by simply adding m_0 to E. Axiom 3 follows from Theorem ??1] We will verify that E itself satisfies Axioms 1 in the following lemma to complete the proof. ☐

Lemma 2.10. *For each size α, $E_\alpha(\phi_1,\phi_2) = E(\alpha\phi_1, \phi_1^{-1}\phi_2)$ is a pseudometric on the full group Γ.*

Proof. That $E_\alpha(\phi,\phi) = 0$ is trivial. Both symmetry and the triangle inequality are based on recoding rearrangements to be based at other arrangements. The following fact will be used several times. Suppose ϕ_1, ϕ_1' and ϕ_2 are three arrangements and R is some finite partition. It follows that $(T^{\alpha\phi_1}, P^{\alpha\phi_1,\phi_2} \vee R)$ and $(T^{\alpha\phi_1'}, P^{\alpha\phi_1',{\phi_1'}^{-1}\phi_1\phi_2\phi_1^{-1}\phi_1'} \vee {\phi_1'}^{-1}\phi_1(R))$ are identical in distribution. This is verified by conjugating both T^α and the partitions by $\phi_1'\phi_1^{-1}$.

To see symmetry, fix ϕ and ϕ' in the full-group where we assume $P^{\alpha,\phi'}$ is finite-valued. This is equivalent to assuming $P^{\alpha\phi',{\phi'}^{-1}}$ is finite valued. Recall that

$$E(\alpha,\phi) = \inf_{\phi'}(h(T^\alpha, P^{\alpha,\phi'} \vee R(P^{\alpha,\phi'}, P^{\alpha,\phi})) + \mu(\{x|\phi'(x) \neq \phi(x)\}).$$

It is a calculation from the above that

$$(T^{\alpha\phi'}, P^{\alpha\phi',{\phi'}^{-1}} \vee R)$$

and

$$(T^{\alpha}, P^{\alpha,{\phi'}^{-1}} \vee \phi'(R))$$

are identical in distribution. For

$$R = R(P^{\alpha\phi,\phi^{-1}}, P^{\alpha\phi',{\phi'}^{-1}}),$$

$$\phi'(R) = R(P^{\alpha,\phi^{-1}}, P^{\alpha,\phi'^{-1}})$$

as $\alpha(x,\phi'^{-1}(x)) = \alpha(\phi'^{-1}\phi(x),\phi'^{-1}(x))$ iff $\phi^{-1}(x) = \phi'^{-1}(x)$ iff $\alpha(x,\phi^{-1}(x)) = \alpha(x,\phi'^{-1}(x))$.

Hence

$$\begin{aligned} &h(T^{\alpha\phi'}, P^{\alpha\phi',\phi'^{-1}} \vee R(P^{\alpha\phi,\phi^{-1}}, P^{\alpha\phi',\phi'^{-1}})) \\ =&h(T^{\alpha}, P^{\alpha,\phi'^{-1}} \vee R(P^{\alpha,\phi^{-1}}, P^{\alpha,\phi'^{-1}})). \end{aligned}$$

We claim

$$P^{\alpha,\phi'^{-1}} \vee R(P^{\alpha,\phi'^{-1}}, P^{\alpha,\phi^{-1}}) \subseteq \bigvee_{g\in G} T^{\alpha}_{g}(P^{\alpha,\phi'} \vee R(P^{\alpha,\phi'}, P^{\alpha,\phi})).$$

This is true first as $P^{\alpha,\phi'^{-1}}(x) = g^{-1}$ where g is the unique group element with $P^{\alpha,\phi'}(T^{\alpha}_{g^{-1}}(x)) = g$, and second as

$$\begin{aligned} P^{\alpha,\phi'^{-1}}(x) =&P^{\alpha,\phi^{-1}}(x) \text{ iff} \\ P^{\alpha,\phi'}(T^{\alpha}_{g^{-1}}(x)) =&P^{\alpha,\phi}(T^{\alpha}_{g^{-1}}(x)). \end{aligned}$$

Thus

$$\begin{aligned} &h(T^{\alpha\phi'}, P^{\alpha\phi',\phi'^{-1}} \vee R(P^{\alpha\phi,\phi^{-1}}, P^{\alpha\phi',\phi'^{-1}})) \leq \\ &h(T^{\alpha}, P^{\alpha,\phi'} \vee R(P^{\alpha,\phi}, P^{\alpha,\phi'})). \end{aligned}$$

The reverse inequality is true by a symmetric argument ($\alpha \to \alpha\phi'$, $\phi' \to \phi'^{-1}$, and $\phi \to \phi^{-1}$). We conclude that $E(\alpha,\phi) = E(\alpha\phi,\phi^{-1})$ and hence symmetry of m^{e}_{α}.

To obtain the triangle inequality we follow a similar path of recoding back to the T^{α} action.

To begin we assume our partitions are finite-valued. Fix α and ϕ'_1 and $\phi'_2 \in \Gamma$ and assume both give finite valued partitions $P_1 = P^{\alpha,\phi'_1}$ and $P_2 = P^{\alpha\phi'_1,\phi'_2}$. Define an auxiliary partition $P_3 = P^{\alpha,\phi'_1\phi'_2\phi'^{-1}_1}$. We know $(T^{\alpha\phi'_1}, P_2)$ and (T^{α}, P_3) are identical processes in distribution.

We claim

$$\bigvee_{g\in G} T^{\alpha\phi'_1}_{g}(P_2) \subseteq \bigvee_{g\in G} T^{\alpha}_{g}(P_1 \vee P_3).$$

This is because $P_2(x) = P_3(\phi'_1(x)) = P_3(T^{\alpha}_{P_1(x)}(x))$. Now $P^{\alpha,\phi'_2\phi'_1} \subseteq \bigvee_{g\in G} T^{\alpha}_{g}(P_1) \vee T^{\alpha\phi'_1}_{g}(P_2)$. As a preparatory calculation for finite-valued partitions we obtain

$$h(T^{\alpha}, P^{\alpha,\phi'_2\phi'_1}) \leq h(T^{\alpha}, P_1) + h(T^{\alpha}, P_3) = h(T^{\alpha}, P_1) + h(T^{\alpha\phi'_1}, P_2).$$

Fix ϕ_1 and ϕ_2 and $\varepsilon > 0$. Now choose ϕ'_1 and ϕ'_2 so that P^{α,ϕ'_1} and $P^{\alpha\phi_1,\phi'_2}$ are finite valued and moreover

$$E(\alpha,\phi_1) \geq h(T^{\alpha}, P^{\alpha,\phi_1} \vee R(P^{\alpha,\phi'_1}, P^{\alpha\phi_1})) + \mu(\{x|\phi'_1(x) \leq \phi_i(x)\}) - \varepsilon$$

and

$$E(\alpha\phi_1,\phi_2) \geq h(T^{\alpha}\phi_1, P^{\alpha\phi_1,\phi'_2} \vee R(P^{\alpha\phi_1,\phi'_2}, P^{\alpha\phi_1,\phi_2})) + \mu(\{x|\phi'_2(x) \neq \phi_2(x)\}) - \varepsilon.$$

Remember that $(T^{\alpha\phi_1}, P^{\alpha\phi_1,\phi'_2} \vee R)$ and $(T^{\alpha\phi'_1}, P^{\alpha\phi'_1,\phi'^{-1}_1\phi_1\phi'_2\phi^{-1}_1\phi'_1} \vee \phi'^{-1}_1\phi_1(R))$ are identical in distribution. Hence $P^{\alpha\phi'_1,\phi'^{-1}_1\phi_1\phi'_2\phi^{-1}_1\phi'_1}$ is finite valued. Set $\phi''_2 = \phi'^{-1}_1\phi_1\phi'_2\phi^{-1}_1\phi'_1$.

Note also that if $\phi_1(x) = \phi'_1(x)$ and $\phi_2\phi_1(x) = \phi'_2\phi_1(x)$ then $\alpha\phi'_1(x, {\phi'_1}^{-1}\phi_1\phi_2\phi_1^{-1}\phi_1(x)) = \alpha\phi_1(x, \phi_2(x))$. That is to say

$$\mu(\{x|P^{\alpha\phi'_1,\phi''_2}(x) \neq P^{\alpha\phi_1,\phi_2}(x)\}) \leq \mu(\{x|\phi_1(x) \neq \phi'_1(x)\}) + \mu(\{x|\phi_2(x) \neq \phi'_2(x)\}).$$

We conclude that

$$\begin{aligned}E(\alpha, \phi_2\phi_1) &\leq h(T^\alpha, P^{\alpha,\phi''_2\phi'_1} \vee R(P^{\alpha,\phi''_2\phi'_1}, P^{\alpha,\phi_2\phi_1})) + \mu(\{x|\phi''_2\phi'_1(x) \neq \phi_2\phi_1(x)\}) \\ &\leq h(T^\alpha, P^{\alpha,\phi''_2\phi'_1} \vee R(P^{\alpha,\phi''_2\phi'_1}, P^{\alpha,\phi_2\phi_1})) + \mu(\{x|\phi_1(x) \neq \phi'_1(x)\}) + \mu(\{x|\phi'_2(x) \neq \phi_2(x)\}).\end{aligned}$$

Now

$$R(P^{\alpha\phi'_2\phi'_1}, P^{\alpha,\phi_2\phi_1}) \subseteq R(P^{\alpha,\phi'_1}, P^{\alpha,\phi_1}) \vee {\phi'_1}^{-1}R(P^{\alpha\phi_1,\phi_1\phi_2}, P^{\alpha\phi_1,\phi_1\phi'_2}).$$

As both P^{α,ϕ'_1} and $P^{\alpha\phi'_1,\phi"_2}$ are finite valued we can conclude from this and our earlier calculations that

$$\begin{aligned}E(\alpha, \phi_2\phi_1) \leq& h(T^\alpha, P^{\alpha,\phi'_1} \vee R(P^{\alpha,\phi'_1}, P^{\alpha,\phi_1})) + \mu(\{x|\phi'_1(x) \neq \phi_1(x)\}) \\ &+ h(T^\alpha, P^{\alpha,\phi'_1\phi_2{\phi'_1}^{-1}} \vee \phi_1^{-1}(R(P^{\alpha\phi_1,\phi_2}, P^{\alpha\phi_1,\phi'_2}))) + \mu(\{x|\phi'_2(x) \neq \phi_2(x)\}) \\ \leq& h(T^\alpha, P^{\alpha,\phi'_1} \vee R(P^{\alpha,\phi'_1}, P^{\alpha,\phi_1})) + \mu(\{x|\phi'_1(x) \neq \phi_1(x)\}) + \\ &\ h(T^{\alpha\phi_1}, P^{\alpha\phi_1,\phi'_2} \vee R(P^{\alpha\phi_1,\phi_2}, P^{\alpha\phi_1,\phi'_2})) + \mu(\{x|\phi'_2(x) \neq \phi_2(x)\}) \\ \leq& E(\alpha, \phi_1) + E(\alpha\phi_1, \phi_2) + 2\varepsilon\end{aligned}$$

Letting ε tend to zero completes the triangle inequality and lemma .

□

3. m-EQUIVALENCE, m-ENTROPY AND m-FINITELY DETERMINED

Our terminal goal in this work is to show that two actions are m^e-equivalent iff they have the same entropy. We do this via the equivalence theory of [3]. To accomplish this we will remind the reader of certain basic ideas there and moreover expand on those ideas. As this machinery is still rather new we first describe in general the how a size m leads to an equivalence relation first among arrangments of a single orbit space and second among measure preserving actions. This will also allow us to establish the basic vocabulary we will need as we continue our investigation of m^e. These results all come from Chapter 2 of [3].

3.1. **m-equivalence and m-entropy.** For a size m and arrangement α we first set Γ_α to be the equivalence classes of elements ϕ of the full group Γ at an m_α distance zero and set $(\hat{\Gamma}_\alpha, m_\alpha)$ to be the m_α completion of Γ_α. That is to say $\hat{\Gamma}_\alpha$ consists of equivalence classes $\langle\phi_i\rangle$ of m_α-Cauchy sequences of full-groups elements, modulo being m_α asymptotically equal. For fixed α and ϕ the map $\phi \to \alpha\phi$ from Γ to $\mathcal{A}$ is well-defined as a map from $\Gamma_\alpha \to \mathcal{A}$ and extends to a uniformly continuous map from $(\hat{\Gamma}_\alpha, m_\alpha) \to (\mathcal{A}, \|\cdot,\cdot\|_1)$. We call this map $P_{m,\alpha}$. The set $P_{m,\alpha}(\hat{\Gamma}_\alpha)$ is our first cut at an equivalence class of arrangements. We know that if $\langle\phi_i\rangle$ is an m_α-Cauchy sequence then there is an arrangement β with $\alpha\phi_i$ converging to β in L^1.

If $P_{m,\alpha}(\hat{\Gamma}_\alpha)$ were to be an equivalence class, then for $\beta \in P_{m,\alpha}(\hat{\Gamma}_\alpha)$ we would have to obtain the same set when we formed $P_{m,\beta}(\hat{\Gamma}_\beta)$. We need not. The size m^e provides the first real example of this phenomenon. Suppose that α is chosen so that T^α is a zero entropy transformation. This means that $m^e(\alpha,\phi) = m^0(\alpha,\phi)$ as $E(\alpha,\phi) = 0$ universally. But now, as described in Example 2 in [3], the set $P_{m^0,\alpha}(\hat{\Gamma}_\alpha)$ consists of all arrangements. In particular it will include β for which T^β is of positive entropy. It should appear reasonable at this point, and we will wee later that if $\alpha_1 \in P_{m^e,\alpha_2}(\hat{\Gamma}_{\alpha_2})$ then $h(T^{\alpha_2}) \geq h(T^{\alpha_1})$. This means it would be impossible for α to be in $P_{m^e,\beta}(\hat{\Gamma}_\beta)$.

To obtain an equivalence relation we consider the subset $\langle\phi_i\rangle$ of elements in $\hat{\Gamma}_\alpha$ where if $\alpha\phi_i \to \beta$ then we find that $\langle\phi_i^{-1}\rangle$ is in $\hat{\Gamma}_\beta$. Of necessity $\beta\phi_i^{-1} \to \alpha$. One can show (see [3]) this condition does not depend on the choice of representative and that in fact the two spaces $(\hat{\Gamma}_\alpha, m_\alpha)$ and $(\hat{\Gamma}_\beta, m_\beta)$ are canonically isometric. Hence we define the subset $\hat{E}_m(\alpha) \subseteq \hat{\Gamma}_\alpha$ of classes $\langle\phi_i\rangle$ such that $\alpha\phi_i \to \beta$ and $\langle\phi_i^{-1}\langle$ is in $\hat{\Gamma}_\beta$. The images $P_{m,\alpha}(\hat{E}_m(\alpha))$ partition $\mathcal{A}$, the set of all arrangements. This is one way to define the notion of m-equivalence, as the equivalence relation on arrangements given by this partition. On the level of measure preserving and ergodic actions U and V, not necessarilly on the same space or with the same orbits, we can say they are m-equivalent if there are m-equivalent arrangements α and β on the same orbit space with U and V conjugate to T^α and T^β. We use the same notation $\alpha \overset{m}{\sim} \beta$ and $U \overset{m}{\sim} V$ for both of these relations without confusion as one is a relation on arrangements and the other on actions. One should keep in mind that $T^\alpha \overset{m}{\sim} T^\beta$ without having $\alpha \overset{m}{\sim} T^\beta$.

One associates an "entropy" with an m-equivalence by defining the m-entropy of T^α to be

$$h_m(T^\alpha) = h_m(\alpha) = \inf(\{h(T^\beta) | \alpha \overset{m}{\sim} \beta\}).$$

The core fact that follows is that sizes come in one of two sorts, those where $h_m(T) = h(T)$ for all ergodic T, in which case m is called *entropy-preserving* and those for which $h_m(T) = 0$ universally, in which case m is called *entropy-free.* We now show that m^e is entropy-preserving. We will see m^e can be regarded as the weakest of entropy-preserving sizes as we will see entropy is a complete invariant.

Lemma 3.1. *For any finite partition Q with k sets, arrangement α and full-group element ϕ,*

$$h(T^\alpha, Q) \leq h(T^{\alpha\phi}, Q) + E(\alpha,\phi)(1 + \ln k).$$

Proof. Fix $\delta > 0$ and use Theorem 2.7 to choose ϕ' so that

$$E(\alpha,\phi) + \delta \geq h(T^\alpha, P^{\alpha,\phi'} \vee R(P^{\alpha,\phi'}, P^{\alpha,\phi})) + \mu(\{x | \phi(x) \neq \phi'(x)\}).$$

Now $(T^{\alpha\phi}, Q)$ is identical in distribution to $(T^\alpha, \phi^{-1}(Q))$ so we will relate $h(T^\alpha, Q)$ and $(T^\alpha, \phi^{-1}(Q))$. Let $B = \{x | \phi(x) \neq \phi'(x)\}$ so $R(P^{\alpha,\phi'}, P^{\alpha,\phi}) = \{B, B^c\}$. Also define a partition $\hat{Q} = \{Q\big|_B, B^c\}$. With this established we observe that

$$Q \subseteq \bigvee_{i=-\infty}^{\infty} (T^\alpha)^{-1}(\phi^{-1}(Q) \vee P^{\alpha,\phi'} \vee \hat{Q})$$

and hence

$$\begin{aligned} h(T^\alpha, Q) &\leq h(T^\alpha, \phi^{-1}(Q) \vee P^{\alpha,\phi'} \vee \hat{Q}) \\ &\leq h(T^\alpha, \phi^{-1}(Q)) + h(T^\alpha, P^{\alpha,\phi'} \vee R(P^{\alpha,\phi}, P^{\alpha,\phi'}) \vee \hat{Q}) \\ &\leq h(T^\alpha, \phi^{-1}(Q)) + h(T^\alpha, P^{\alpha,\phi'} \vee R(P^{\alpha,\phi}, P^{\alpha,\phi'})) + h(T^\alpha, \hat{Q}|(B, B^c)) \\ &\leq h(T^{\alpha,\phi}, Q) + h(T^\alpha, P^{\alpha,\phi'} \vee R(P^{\alpha,\phi}, P^{\alpha,\phi'})) + \mu(B)\ln(k) \\ &\leq h(T^{\alpha,\phi}, Q) + (E(\alpha,\phi) + \delta)(1 + \ln(k)). \end{aligned}$$

Now let $\delta \to 0$. ☐

Theorem 3.2. *Suppose α is an arrangement and $\langle \phi_i \rangle \in \hat{\Gamma}_\alpha$ with $\alpha\phi_i \to \beta$. We conclude that*

$$h(T^\alpha) \leq h(T^\beta).$$

Proof. Let Q be a finite partition with k sets. Fix $\delta > 0$ and choose i_0 so that for all $j \geq i_0$

$$E(\alpha\phi_{i_0}, \phi_{i_0}^{-1}\phi_j) \leq \delta/(\ln(k) + 1).$$

We now compute that

$$\begin{aligned} h(T^\alpha, Q) &= h(T^{\alpha\phi_{i_0}}, \phi_{i_0}^{-1}(Q)) \\ &\leq h(T^{\alpha\phi_j}, \phi_{i_o}^{-1}(Q)) + E(\alpha\phi_{i_0}, \phi_{i_0}^{-1}\phi_j)(1 + \ln(k)) \\ &\leq h(T^{\alpha\phi_j}, \phi_{i_0}^{-1}(Q)) + \delta. \end{aligned}$$

As $(T^{\alpha\phi_j}, \phi_{i_0}^{-1}(Q))$ converges in distribution to $(T^\beta, \phi_{i_0}^{-1}(Q))$,

$$\overline{\lim}_{j\to\infty} h(T^{\alpha\phi_j}, \phi_{i_0}^{-1}(Q)) \leq h(T^\beta, \phi_{i_0}^{-1}(Q)) \leq h(T^\beta).$$

Hence $h(T^\alpha, Q) \leq h(T^\beta)$ and taking the supremum over Q we are done. ☐

Corollary 3.3. *If $\alpha \overset{m^e}{\sim} \beta$ the $h(T^\alpha) = h(T^\beta)$ and hence m^e is an entropy-preserving size.*

3.2. **Variations on m-finitely determined.** In [3] a notion of m-joinings of actions is given for any size m. We give here a somewhat simplified version of that definition that will be sufficient for our purposes. We will then describe and extend the basic theory of m-finitely determined actions. We will introduce a new condition we call **reverse weakly m-finitely determined** and will show that, so long as an m-f.d. action exists, it is equivalent to being m-f.d. This will play a critical role in the characterization of m^e-equivalence.

Definition 3.4. *Suppose $T_1^{\alpha_1}$ and $T_2^{\alpha_1}$ are two ergodic actions. We define the space of m-joinings (in the weak sense) of $J(T_1^{\alpha_1}, T_2^{\alpha_1})$ to be the set of all quadruples $(\hat{\alpha}_1, \pi_1, \hat{\alpha}_1, \pi_2)$ where $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are m-equivalent arrangements on some common ergodic orbit space $(\hat{X}, \hat{\mathcal{F}}, \hat{\mu}, \hat{\mathcal{O}})$ and π_i projects $\hat{T}^{\hat{\alpha}_i}$ to T^{α_i}.*

In [3] much more structure is added to this definition in the form of a specific Cauchy-sequence of rearrangements giving the m-equivalence of $\hat{\alpha}_1$ and $\hat{\alpha}_2$ and these are topologized as a space of measures on a canonical symbolic space encoding all these rearrangements. That is why we call these joinings in the weak sense. This full level of complexity is not necessary for our purposes as we will not be reproving the equivalence theorem. Any joining in the weak sense will lift to a joining in the more elaborated sense and certainly any such elaborated joining will project to what we have defined here. We will omit saying in the weak sense as we continue.

Earlier we worked extensively with countably infinite partitions. That work is finished but we will need to work now with finite partitions. To remind the reader, for us a **finite partition** of a space X will be a measurable map $P : X \to \Sigma$ where Σ is some finite labeling set. A **process** is a pair (T^α, P) where P is a finite partition and T^α is an ergodic action. We generally will suppress the state variables $(X, \mathcal{F}, \mu)$ in such expressions. There are three important metrics on processes. The first is the **partition metric** given earlier, $d_0(P, P') = \mu(\{x : P(x) \neq P'(x)\})$, which we can also write as $\mu(P \triangle P')$. Notice that $d_0(P, P')$ will be the density along any T^α orbit of those points where the two names differ. The second **distribution metric** $\|(T_1^{\alpha_1}, P_1), T_2^{\alpha_2}, P_2)\|_*$ is a metric giving the weak* topology on shift invariant measures on Σ^G where we project these processes to their representations as such measures on sequences of symbols. The third metric is directly related to the size m.

Definition 3.5. *For two ergodic and Σ-valued processes $(T_1^{\alpha_1}, P_1)$ and $(T_2^{\alpha_2}, P_2)$ we define the $\overline{m}$-distance between them by*

$$\overline{m}((T_1^{\alpha_1}, P_1), (T_2^{\alpha_2}, P_2)) = \inf_{(\hat{\alpha}_1, \pi_1, \hat{\alpha}_2, \pi_2) \in J_m(T_1^{\alpha_1}, T_2^{\alpha_2})} [\mu(\pi_1^{-1}(P_1) \triangle \pi_2^{-1}(P_2))) + m_{\alpha_1}(\alpha_1, \alpha_2)]$$

That is to say, we seek the closest possible m-joining of the two processes, close in that we want both m_{α_1} to be small and the two partitions to agree as often as we can.

We now begin a discussion of the notion of m-finitely determined, which characterizes the m-Bernoulli classes of the equivalence relation. We will extend some basic conclusions of [3] in preparation for showing all actions are m^e-finitely determined.

Definition 3.6. *We say a process (T^α, P) is m-finitely determined (abbreviated m-f.d.) if for every $\varepsilon > 0$ there is a δ so that for any other ergodic process $(T_1^{\alpha_1}, P_1)$ satisfying*

(i) $\|(T^\alpha, P), (T_1^{\alpha_1}, P_1)\|_* < \delta$ *and*
(ii) $h_m(T_1^{\alpha_1}, P_1) \geq h_m(T^\alpha, P) - \delta$ *then*
(iii) $\overline{m}((T^\alpha, P), (T_1^{\alpha_1}, P_1)) < \varepsilon$.

This definition is useful from the perspective of the m-equivalence theorem but from the perspective of verifying a system is m-f.d. it is unwieldy since to verify (iii) one must construct an m-joining. In [3] a more constructive condition is also defined.

Definition 3.7. *We say a process (T^α, P) is weakly m-finitely determined if for all $\varepsilon > 0$ there is a δ so that if we have an ergodic $(T_0^{\alpha_0}, P_0)$ satisfying:*

(i) $\|(T^\alpha, P), (T_0^{\alpha_0}, P_0\|_* < \delta$ *and*
(ii) $h_m(T_0^{\alpha_0}, P_0) \geq h_m(T^\alpha, P) - \delta$

then the following holds. For all $\delta' > 0$ there is an ergodic lift $T_1^{\alpha_1}$ of $T_0^{\alpha_0}$ (i.e. there is a π factoring the $T_1^{\alpha_1}$ action onto $T_0^{\alpha_0}$) and a ϕ_1 in the full group of $\mathcal{O}(T_1^{\alpha_1})$ and a Σ-valued partition P' with:

(a) $\mu_1(\pi^{-1}(P_0) \triangle P_1) < \varepsilon$ *and*
(b) $m(\alpha_1, \phi_1) < \varepsilon$ *and moreover*

(i)′ $\|(T_1^{\alpha_1\phi_1}, P_1), (T^\alpha, P)\|_* < \delta'$ *and*
(ii)′ $h_m(T_1^{\alpha_1}, P_1) \geq h_m(T^\alpha, P) - \delta'$.

This definition is obviously meant to generate an induction. We note there is a typo in the definition as given in [3], a ϕ_1 is missing is statement (ii)'.

We now give a variant of the weak m-f.d. condition that switches which action is perturbed. As we will soon show, this condition is actually more well-suited to the theory than the notion of weakly m-finitely determined of [3]. Although it appears to not allow an induction to a Sinai's Theorem we will see it actually allows an induction to proceed when both actions under consideration satisfy it.

Definition 3.8. *We say a process (T^α, P) is reverse-weakly m-finitely determined (r.w. m-f.d.) if for all $\varepsilon > 0$ there is a δ so that if we have an ergodic $(T_0^{\alpha_0}, P_0)$ satisfying:*

(i) $\|(T^\alpha, P), (T_0^{\alpha_0}, P_0)\|_* < \delta$ *and*
(ii) $h_m(T_0^{\alpha_0}, P_0) \geq h_m(T^\alpha, P) - \delta$

then the following holds. For all $\delta' > 0$ there is an ergodic lift $T_1^{\alpha_1}$ of T^α (i.e. there is a π factoring the $T_1^{\alpha_1}$ action onto T^α) and a ϕ_1 in the full group of $\mathcal{O}(T^\alpha)$ and a Σ-valued partition P_1 with:

(a) $\mu_1(\pi^{-1}(P)\triangle P_1) < \varepsilon$ *and*
(b) $m(\alpha, \phi_1) < \varepsilon$ *and moreover*
(i)′ $\|(T_1^{\alpha_1\phi_1}, P_1), (T_0^\alpha, P_0)\|_* < \delta'$ *and*
(ii)′ $h_m(T_1^{\alpha_1}, P_1) \geq h_m(T_0^\alpha, P_0) - \delta'$.

Definition 3.9. *We say an ergodic action T^α is m-finitely determined, (weakly m-finitely determined or reverse-weakly m-finitely determined) if it is for all finite partitions.*

It is shown in [3] that m-finitely determined processes must be weakly m-finitely determined and moveover if m is a 3^+ size in that the upper-semicontinuity of Axiom 3 is actually continuity, then the converse holds as well. The proof of the first part of this involves examining an m-joining that almost achieves the $\overline{m}$-distance between (T^α, P) and $(T_0^{\alpha_0}, P_0)$. This argument works equally well for reverse-weakly m-f.d. and implies that any m-f.d. process will be r.w. m-f.d. The proof of the second half, that for 3^+ sizes the converse holds, does not appear to lift to the reverse-weakly m-f.d. property. We will see though that so long as their exists an m-f.d. action, any r.w. m-f.d. action will be m-f.d. We obtain this by showing that any two r.w. m-f.d. actions of the same m-entropy must be m-equivalent. As the m-f.d. property is an m-equivalence invariant the result will follow.

One might naturally ask why reverse weakly m-f.d. is not taken as a better definition of the most random classes rather than m-f.d. There are two reasons. First, the m-f.d. condition as given provides a good topological context, the space of m-joinings, for the equivalence theorem. Second, to establish our work here one will in fact need that machinery. In the final analysis we conjecture the two notions are the same in general and agree with weak m-f.d. The only potential gap would be among sizes that are not 3^+ and are entropy free. We know of no examples of this.

We now proceed to show that reverse-weakly m-f.d. actions of the same entropy are m-equivalent. We accomplish this by showing that r.w. m-f.d. processes satisfy the m-f.d.

condition where one restricts the target process $(T_1^{\alpha_1}, P_1)$ to also be r.w. m-f.d. and of the same m-entropy. From there the conclusion will follow almost immediately. Our first step will be to show that the only need for the lift of T^α to $T_1^{\alpha_1}$ in the r.w. m-f.d. condition (or the weak m-f.d. condition for that matter but we do not need this case) is a possible lack of sufficient entropy in T^α. We will present this development in detail for entropy-preserving sizes. The entropy-free case is easier as there is no need for any entropy conditions.

Definition 3.10. *We say a process (T^α, P) is extra-weakly m-finitely determined if for all $\varepsilon > 0$ there is a $\delta > 0$ so that for any ergodic $(T_0^{\alpha_0}, P_0)$ with*

(i) $\|(T^\alpha, P), (T_0^{\alpha_0}, P_0\|_* < \delta$ *and*
(ii) $h(T_0^{\alpha_0}, P_0) \geq h(T^\alpha, P) - \delta$ *and*
(iii) $h(T^\alpha) \geq h(T_0^{\alpha_0}, P_0)$

then the following holds. For all $\delta' > 0$ there is a ϕ_1 in the full-group of $\mathcal{O}(T^\alpha)$ and a Σ-valued partition P_1 so that

(a) $\mu(P \triangle P_1) < \varepsilon$ *and*
(b) $m(\alpha, \phi_1) < \varepsilon$ *and moreover*
(i)$'$ $\|(T_0^{\alpha_0}, P_0), (T^{\alpha\phi_1}, P_1\|_* < \delta'$ *and*
(ii)$'$ $h_m(T^{\alpha\phi_1}, P_1) \geq h_m(T_0^{\alpha_0}, P_0) - \delta'$

The difference between reverse-weakly m-f.d. and extra-weakly m-f.d. is simply that no lift of T^α is required so long as T^α has at least as much entropy as $(T_0^{\alpha_0}, P)$. We now show that processes which are reverse-weakly m-f.d. are in fact extra-weakly m-f.d. What is needed to prove this is a copying lemma that allows us to copy both P_1 and ϕ_1 from the extended system $T_1^{\alpha_1}$ down into the factor action T^α. What we need is given in [3] as the first part of the proof of Lemma 4.0.12. We remind the reader that for a rearrangment pair (α, ϕ) the function $g_{(\alpha,\phi)}$ is defined to be $\alpha(x, \phi(x))$. Moreover we say a rearrangement is "finite-valued" if this function takes on only finitely many values and hence generates a finite partition.

Lemma 3.11. *Suppose $T_1^{\alpha_1}$ is a free and ergodic action and P_1 and Q are finite partitions. Suppose $\mathcal{H}$ is a $T_1^{\alpha_1}$ invariant sub σ-algebra and Q is $\mathcal{H}$ measurable. Suppose also that ϕ_1 is a finite valued element in the full-group of $\mathcal{O}(T_1^{\alpha_1})$ and for some ε_0*

$$h(T_1^{\alpha_1}, Q) \geq h(T_1^{\alpha_1\phi_1}, P) - \varepsilon_0/2.$$

Then for any $\delta_5 > 0$ there exists a full-group element ϕ_1' and partition P_1', both of which are $\mathcal{H}$-measurable with:

(1) $\|(T_1^{\alpha_1}, P_1 \vee Q \vee g_{(\alpha_1,\phi_1)}), (T_1^{\alpha_1}, P_1' \vee Q \vee g_{(\alpha_1,\phi_1')})\|_* < \delta_5$ *and*
(2) $h(T_1^{\alpha_1\phi_1'}, P_1') \geq h(T_1^{\alpha_1\phi_1}, P_1) - \varepsilon_0$.

Note that if δ_5 is sufficiently small then (1) will imply

$$\|(T_1^{\alpha_1\phi_1}, P_1 \vee Q), (T_1^{\alpha_1\phi_1'}, P_1' \vee Q)\|_* < \varepsilon_0.$$

Proof. This result is not stated explicitely in [3] but is already obtained in the beginning of the proof of Lemma 4.0.12. There one is copying a finite sequence of ϕ_i's where the terminal one (ϕ_{I+1} and our ϕ_1) is copied first obtaining

$$\|(T_1^{\alpha_1}, P_1 \vee Q \vee g_{(\alpha,\phi_1)}), (T_1^{\alpha_1}, P_1' \vee Q \vee g_{(\alpha,\phi_1')})\|_* < \delta_5$$

where δ_5 can be chosen arbitrarily small. (This is why we refered to this quantity in our lemma as δ_5.) Statement (2) is in the conclusion of Lemma 4.0.12. □

Theorem 3.12. *If (T^α, P) is reverse-weakly m-f.d. then it must be extra-weakly m-f.d. As stated earlier we assume m is entropy preserving.*

Assume (T^α, P) is reverse-weakly m-f.d. and fix $\varepsilon > 0$, the value relative to which we intend to prove extra-weakly m-f.d. Choose $\varepsilon_1 < \varepsilon/3$ so small that for any $T_0^{\alpha_0}$ and P_0, if

$$\|(T_0^{\alpha_0}, P_0), (T^\alpha, P)\|_* < \varepsilon_1 \quad \text{then}$$

$$h(T_0^{\alpha_0)}) < h(T^\alpha, P) + \varepsilon/3.$$

Use ε_1 in the definition of reverse-weakly m-f.d. for (T^α, P) to obtain a δ and now suppose we have a $(T_0^{\alpha_0}, P_0)$ with:

(i) $\|(T^\alpha, P), (T_0^{\alpha_0}, P_0)\|_* < \delta$ and
(ii) $h(T_0^{\alpha_0}, P_0) \geq h(T^\alpha, P) - \delta$.

For all $\delta_1 > 0$ we will then obtain a lift $T_1^{\alpha_1}$ of T^α and a ϕ_1 in the full-group of $\mathcal{O}(T_1^{\alpha_1})$ and a partition P_1 so that

(a) $m(\alpha_1, \phi_1) < \varepsilon_1$ and
(b) $\mu_1(P \triangle P_1) < \varepsilon_1$ and moreover
(i)′ $\|(T_1^{\alpha_1\phi_1}, P_1), (T_0^{\alpha}, P_0)\|_* < \delta_1$ and
(ii)′ $h(T_1^{\alpha_1\phi_1}, P_1) \geq h(T_0^{\alpha}, P_0) - \delta_1$.

By Theorem 4.0.2 of [3] we can assume here w.l.o.g. that ϕ_1 is finite valued. Assume now the added condition that

$$h(T^\alpha) \geq h(T_0^{\alpha_0}, P_0).$$

Now fix a value $\delta' > 0$ and we attempt to verify the conclusions of extra-weakly m-f.d. for this value. Choose $\delta_1 < \delta'/2$ so small that (i)′ implies

$$h(T_1^{\alpha_1\phi_1}, P_1) < h(T_0^{\alpha_0}, P_0) + \delta'/4 < h(T^{\alpha_0}) + \delta'/2.$$

Let $\mathcal{H}$ be the sub σ-algebra of the factor action T^α of the lift $T_1^{\alpha_1}$. Choose a partition Q in $\mathcal{H}$ so that

$$h(T_1^{\alpha_1}, Q) > h(T_1^{\alpha_1}, P) - \delta'/4.$$

The lemma now tells us that there are ϕ_1' and P_1' both $\mathcal{H}$-measurable so that

(1) $\|(T_1^{\alpha_1\phi_1}, P_1 \vee Q), (T_1^{\alpha_1\phi_1'}, P_1' \vee Q)\|_* < \delta'/2$ and
(2) $h(T_1^{\alpha_1\phi_1'}, P_1') \geq h(T_1^{\alpha_1\phi_1}, P_1) - \delta'/2 > h(T^\alpha, P) - \delta'/2 + \delta_2 > h(T^\alpha, P) - \delta'$.

and as ϕ_1' and P_1' are $\mathcal{H}$-measurable, i.e. in the sub σ-algebra of the T^α factor we are done

We now need a small idea about Polish spaces or more particularly G_δ subsets of a complete metric space.

Lemma 3.13. *Suppose* (X, d) *is a complete metric space and* $A \subseteq X$ *is a* G_δ *subset. There are then continuous functions* $r_i : A \to \mathbb{R}^+$ *so that for any sequence* $\{x_i\} \in A^{\mathbb{N}}$ *satisfying* $d(x_i, x_{i+1}) < r_i(x_i)$*, one has*

$$\lim_{i\to\infty} x_i = x \in A.$$

Proof. Without loss of generality we can assume that X has diameter 1. For $A = \bigcap_{i=1}^{\infty} \mathcal{O}_i$, where the $\mathcal{O}_i$ are open and nested, set

$$r_i(x) = d(x, \mathcal{O}_i^c)/2^{i+1} \leq 1/2^{i+1}.$$

If $d(x_i, x_{i+1}) < r_i(x_i) \leq 1/2^{i+1}$ then the x_i are Cauchy and converge to some $x \in X$. Our goal then is to show $x \in A$. Suppose not, i.e. $x \notin \mathcal{O}_i$ for some i. Then $d(x_j, \mathcal{O}_i^c) \to 0$ so select j_0 so that for all $j \geq j_0$, $d(x_j, \mathcal{O}_i^c) \leq d(x_{j_0}, \mathcal{O}_i^c)$. Now for $j > j_0$,

$$\begin{aligned} d(x_j, \mathcal{O}_i^c) &\geq d(x_{j_0}, \mathcal{O}_i^c) - \sum_{k=j_0}^{j} d(x_k, x_{k+1}) \\ &\geq d(x_{j_0}, \mathcal{O}_i^c) - \sum_{k=j_0}^{j} r_k(x_k) \\ &\geq d(x_{j_0}, \mathcal{O}_i^c) - \sum_{k=j_0}^{j} \frac{1}{2^{k+1}} d(x_k, \mathcal{O}_i^c) \\ &\geq d(x_{j_0}, \mathcal{O}_i^c)/2 \end{aligned}$$

which is a conflict and implies $x \in A$. □

We develop an extension of the equivalence theorem of Section 7.3 of [3]. The idea is to prove a restricted version of this theorem subject to all actions considered lying in some conjugacy invariant subset. We first modify the definition of m-f.d. to be given relative to such a subset.

Definition 3.14. *Suppose* $\mathcal{C}$ *is a set of actions closed under conjugation and* m *is a size. We say this* $\mathcal{C}$ *is self-*m*-f.d. if for any* $T^\alpha \in \mathcal{C}$ *and finite partition* P *the process* (T^α, P) *satisfies the definition of* m*-f.d. subject to the requirement that the target process* $(T_1^{\alpha_1}, P_1)$ *comes from an action* $T_1^{\alpha_1}$ *also in* $\mathcal{C}$*. In this case we say* T^α *is* $\mathcal{C}$*-*m*-finitely determined.*

Theorem 3.15. *Suppose* $\mathcal{C}$ *is a set of actions that is self-*m*-f.d. Then any two* m*-f.d. actions in* $\mathcal{C}$ *of the same* m*-entropy are* m*-equivalent.*

Proof. The development in section 7.3 of [3] can be followed in complete detail. Remember the goal there is to show that when T^α is m-f.d. the set of m-joinings of (T^α, P) and $T_1^{\alpha_1}$ where some large finite number of full-group elements $\phi_1, \ldots, \phi_I$ in the m-joining and the partition P are almost measurable with respect to the $T_1^{\alpha_1}$-system form an open and dense subset of the Polish space of m-joinings. This is the content of Theorem 7.3.3. If we modify the statement of Theorem 7.3.3 to ask that T^α only be $\mathcal{C}$-m-f.d. but also ask that $T_1^{\alpha_1}$ should be in $\mathcal{C}$ the argument follows in full detail. Remember that the actions $T_1^{\alpha_1 \phi_i'}$ will be conjugate to $T_1^{\alpha_1}$ and hence still in $\mathcal{C}$. The result now follows. □

Theorem 3.16. *For m an entropy-preserving size, set $\mathcal{C}_m(e)$ to be the collection of reverse-weakly m-finitely determined actions of entropy e. Each of the sets $\mathcal{C}_m(e)$ is self-m-finitely determined.*

Proof. Fix a value $e \geq 0$ and suppose $T^\alpha \in \mathcal{C}_m(e)$. In particular for all partitions P, (T^α, P) is reverse-weakly m-f.d. and hence by Theorem 3.12 extra-weakly m-f.d. Our goal is to show that (T^α, P) is $\mathcal{C}_m(e)$ m-finitely determined. Toward this end select a value $\varepsilon > 0$ and we seek a δ. Sketching the argument, to establish the $\mathcal{C}_m(e)$-m-finitely determined property we must construct an m-joining. We will do so by constructing inductively two sequences of full-group elements, one moving (T^α, P) and the other moving $(T_0^{\alpha_0}, P_0)$ so that both sequences give m-equivalences and they reach a common process.

Recall that the set $E_m(\alpha)$, the set of arrangements m-equivalent to α, is a G_δ subset of the m_α completion $\hat{\Gamma}_\alpha$ and hence there are continuous functions $r_i : E_m(\alpha) \to \mathbb{R}^+$ from Lemma 3.13. We can assume w.l.o.g. that $r_i < \varepsilon/10 \times 2^i$. Using $r_1(\alpha)$ (as ε) in the extra-weakly m-f.d. condition of (T^α, P) we obtain a $\delta > 0$ and this is the δ we seek.

Now suppose that $T_0^{\alpha_0} \in \mathcal{C}_m(e)$ and P_0 satisfy

(i) $\|(T^\alpha, P), (T_0^{\alpha_0}, P_0)\|_* < \delta$ and
(ii) $h(T_0^{\alpha_0}, P_0) \geq h(T^\alpha, P) - \delta$. We automatically also have
(iii) $h(T^\alpha) = e \geq h(T_0^{\alpha_0}, P_0)$.

We are now in a position to apply the e.w. m-f.d. condition on (T^α, P). We must choose the δ' we seek. To do this note that $E_m(\alpha_0)$ is also a G_δ subset of its m_{α_0} completion and hence Lemma 3.13 gives us a sequence of functions we call $\hat{r}_i : E_m(\alpha_0) \to \mathbb{R}^+$. Once more, w.l.o.g. we can assume $\hat{r}_i < \frac{\varepsilon}{10\times 2^i}$. Using $\hat{r}_1(\alpha_0)$ (as ε) in the e.w. m-f.d. condition on $(T_0^{\alpha_0}, P_0)$ we obtain a "δ" which we call δ_0 and use as δ' to obtain ϕ_1 and P_1 satisfying:

(a) $m(\alpha, \phi_1) < r_1(\alpha)$,
(b) $\mu(P \triangle P_1) < r_1(\alpha)$ and
(i) $\|(T^{\alpha\phi_1}, P_1), (T_0^{\alpha_0}, P_0)\|_* < \delta_0$,
(ii) $h(T^{\alpha\phi_1}, P_1) > h(T_0^{\alpha_0}, P_0) - \delta_0$ and
(iii) $h(T_0^{\alpha_0}) = e \geq h(T^{\alpha\phi_1}, P_1)$.

This completes the first step of the induction we now describe by creating $\alpha_1 = \alpha\phi_1$ and P_1. In the induction at even steps we will modify α_0 and P_0 and along the odd steps we will modify α and P. We have just completed an odd step, i.e. step 1. We now describe the inductive even step in detail. Suppose we have constructed arrangements

$$\begin{aligned} &\alpha, \alpha_1 = \alpha\phi_1, \ldots, \alpha_{2k-1} = \alpha\phi_{k+1}, \\ &\alpha_0, \alpha_2 = \alpha_2\hat{\phi}_1, \ldots, \alpha_{2k} = \alpha_0\hat{\phi}_k \quad \text{and partitions} \\ &P, P_1, P_3, \ldots, P_{2k-1} \quad \text{and} \\ &P_0, P_2, \ldots, P_{2k} \quad \text{satisfying} \end{aligned}$$

(1-odd) $m_\alpha(\alpha_{2i-3}, \alpha_{2i-1}) = m(\alpha_{2i-3}, \phi_i^{-1}\phi_{i+1}) < r_i(\alpha_{2i-3})$
(1-even) $m_{\alpha_0}(\alpha_{2i-2}, \alpha_{2i}) = m(\alpha_{2i-2}, \hat{\phi}_{i-1}^{-1}\hat{\phi}_i) < \hat{r}_i(\alpha_{2i-2})$.
(2-odd) $\mu(P_{2i-3} \triangle P_{2i-1}) < r_i(\alpha_{2i-3}) \leq \frac{2}{10\times 2^i}$ and

(2-even) $\mu_0(P_{2i-2}\triangle P_{2i}) < \hat{r}_i(\alpha_{2i-2}) \leq \frac{1}{10\times 2^i}$.

Furthermore

(i) $\|(T^{\alpha_{2k-1}}, P_{2k-1}), (T_0^{\alpha_{2k}}, P_{2k})\|_* < \delta_{2k}$
(ii) $h(T^{\alpha_{2k-1}}, P_{2k-1}) > h(T_0^{\alpha_{2k}}, P_{2k}) - \delta_{2k}$ and
(iii) $h(T_0^{\alpha_2 k}) = e \geq h(T^{\alpha_{2k-1}}, P_{2k-1})$

where δ_k is obtained from the following considerations. As $T_0^{\alpha_{2k}}$ is conjugate to $T_0^{\alpha_0}$ the process $(T_0^{\alpha_{2k}}, P_{2k})$ is e.w. m-f.d. and hence using $\hat{r}_{k+1}(\alpha_{2k})$ (as ε) we obtain a δ which is the value taken for δ_{2k}. We now wish to set the value of δ' in this e.w. m-f.d. condition and read of the conclusions. As $(T^{\alpha_{2k-1}}, P_{2k-1})$ is e.w. m-f.d., using $r_{k+1}(\alpha_{2k-1})$ (as ε) we obtain a δ which we take as δ' in the e.w. m-f.d. condition on $(T_0^{\alpha_{2k}}, P_{2k})$. We define δ_{2k+1} to be this δ'. Reading off the conclusions, we will have a $\hat{\phi}$ in the full group of T_0^{2k} and a partition P_{2k+2} with

(1-even) $m(\alpha_{2k}, \hat{\phi}) = m_{\alpha_0}(\alpha_{2k}, \alpha_{2k}\hat{\phi}) < \hat{r}_k(\alpha_{2k})$
(2-even) $\mu_0(P_{2k}\triangle P_{2k+2}) < \hat{r}_k(\alpha_{2k})$.

Furthermore for $\alpha_{2k+2} = \alpha_{2k}\hat{\phi}$ and $\hat{\phi}_{k+1} = \hat{\phi}_k^{-1}\hat{\phi}$

(i) $\|(T_0^{\alpha_{2k_2}}, P_{2k+2}), (T^{\alpha_{2k-1}}, P_{2k-1})\|_* < \delta_{2k+1}$
(ii) $h(T_0^{\alpha_{2k+2}}, P_{2k+2}) > h(T^{\alpha_{2k-1}}, P_{2k-1}) - \delta_{2k+1}$ and
(iii) $h(T^{\alpha_{2k-1}}) = e \geq h(T_0^{\alpha_{2k+2}}, P_{2k+2})$.

Replacing the previous even-term statements (1-even) and (2-even) with these new ones and statements (i), (ii) and (iii) with these new ones and now switching the roles of the odd and even terms we are ready to proceed with the induction on the T^α side. In the limit then we obtain two sequences of arrangements $\{\alpha = \alpha_{-1}, \alpha_1, \alpha_3, \dots\}$ and $\{\alpha_0, \alpha_2, \alpha_4, \dots\}$ both of which must converge by statements (1) to arrangements β and β_0 with $\alpha \overset{m}{\sim} \beta$ and $\alpha_0 \overset{m}{\sim} \beta_0$. Moreover

$$m_\alpha(\alpha, \beta) \leq \sum_i r_i(\alpha_{2i-3}) < \varepsilon/10 \qquad \text{and} \qquad m_{\alpha_0}(\alpha_0, \beta_0) < \sum_i \hat{r}_i(\alpha_{2i-2}) < \varepsilon/10.$$

We also obtain sequences of partitions $\{P = P_{-1}, P_1, P_3, \dots\}$ and $\{P_0, P_2, P_4, \dots\}$ which by similar estimates must converge in d_0 to partitions Q and Q_0 with

$$\mu(P\triangle Q) < \varepsilon/10 \qquad \text{and} \qquad \mu_0(P_0\triangle Q_0) < \varepsilon/10.$$

The sequence of statements of type (i) tell us that

$$\|(T^\beta, Q), (T_0^{\beta_0}, Q_0)\|_* = 0$$

which is to say these two processes are identical. Let $\overline{T}^{\overline{\beta}}$ be the relatively independent joining of T^β and $T_0^{\beta_0}$ over this common factor and let $\overline{Q}$ be the common lift of Q and Q_0. Let π_1 and π_2 be the factor maps to T^β and $T_0^{\beta_0}$. The arrangements α_{2k-1} lift via π_1 to arrangements $\overline{\alpha}_{2k-1}$. Ramifying the σ-algebra does not change the fact that $\overline{\alpha}_{-1} \overset{m}{\sim} \overline{\beta}$. Similarly lifting the even sequence via π_2 we will have $\overline{\alpha}_0 \overset{m}{\sim} \overline{\beta}$. We conclude $(\overline{\alpha}_{-1}, \pi_1, \overline{\alpha}_0, \pi_2)$ is an m-joining of

T^α and $T_0^{\alpha_0}$ and moveover

$$\begin{aligned} m(\overline{\alpha}_{-1},\overline{\alpha}_0) &\leq m(\overline{\alpha}_{-1},\overline{\beta}) + m(\overline{\alpha}_0,\overline{\beta}) \\ &= m(\alpha,\beta) + m(\alpha_0,\beta_0) < \varepsilon/5. \end{aligned}$$

We conclude by a similar calculation comparing with $\overline{Q}$ that

$$\overline{\mu}(\pi_1^{-1}(P) \triangle \pi_2^{-1}(P_0)) < \varepsilon/5.$$

We conclude that

$$\overline{m}((T^\alpha, P), (T_0^{\alpha_0}, P_0)) < \varepsilon$$

and hence that T^α is $\mathcal{C}_m(e)$-m-f.d. and hence $\mathcal{C}_m(e)$ is self-m-f.d. □

Corollary 3.17. *If there exists an m-f.d. action of m-entropy e then all reverse-weakly m-f.d. actions of m-entropy e are m-f.d. This is always true if $e > 0$.*

Proof. First note that in all results in this section one can replace all occurences of h with h_m. For entropy-preserving sizes this is obvious. For entropy-free sizes it amounts to deleting all conditions and conclusions involving entropy. All arguments given proceed under these circumstances. We have just shown that in $\mathcal{C}_m(e)$ any two r.w. m-f.d. actions will be m-equivalent. If there exists an m-f.d. action of m-entropy e, it will be such an action and hence any r.w. m-f.d. action would be m-equivalent to it. As [3] proves that m-f.d. is an m-equivalence invariant, the result follows. For $e > 0$ a Bernoulli shift of entropy e will be m-f.d. (note that m must be entropy preserving for e to be > 0). □

4. Combinatorial background

Using the developments of the previous subsections we can now aim to understand m^e-equivalence. Our goal will be to see that all actions are r.w. m^e-f.d. and hence conclude that two actions are m^e equivalent iff they have the same entropy. At the core of such a result must be a basic combinatorial fact about the number of permutations needed to rearrange one set of names into another. We now present that background. We first giving a simplified version, then the one we really need.

Fix a finite symbol set Σ and C a finite subset of G. Take a vector of values $\vec{t} = (t_s : s \in \Sigma)$ with $t_s \in \mathbb{N}$ and $\sum t_s = \#C$. Now let $\Omega = \Omega(G)$ be the set of all maps $\omega : C \to \Sigma$ and $\#\omega^{-1}(s) = t_s$. This is the space of all Σ, C-names where the symbol s occurs t_s times. We know that $\#\Omega = \frac{n!}{\Pi t_s!}$ where $\#C = n$. This is of the form $2^{n(h+\varepsilon_n)}$ where $h = -\sum (t_s/n)\log_2(t_s/n)$ i.e. the entropy of the probability vector $\{t_s/n\}$, and $\varepsilon_n \to 0$. We consider Ω to be a probability space where each name has equal probability $1/\#\Omega$.

Consider now certain other probability measures m on Ω. What we require is that the mass of any name $m(\omega)$ is of the form $\frac{c(\omega)}{\#\Omega}$ where $c(\omega)$ is a nonnegative integer. Any such measure can be represented as follows. Let $\ell : \{1, 2, \ldots, \#\Omega\} \to \Omega$ be some not-necessarilly 1-1 map and let λ be normalized counting measure on $\{1, \ldots, \#\Omega\}$. The measures we are considering are of the form $\ell^*(\lambda)$ where $c(\omega)$ is simply $\#\ell^{-1}(\omega)$. Of course any such measure will have $\#\Omega!$ such representations. For a map ℓ we define $c(\ell) = \sup_\omega \#\ell^{-1}(\omega)$, that is to say ℓ is at most $c(\ell)$ to one. We say a measure that can take the form $m = \ell^*(\lambda)$ is **adjusted** and we write $c(m) = c(\ell)$ which now is $\#\Omega$ times the largest measure of any element ω.

For $p \in S(C)$, the symmetric group of all permutations of C we can regard p as taking $\Omega \to \Omega$ by setting $p(\omega)(s) = \omega(p^{-1}(s))$, i.e. we permute the placement of the symbols in the name ω by p. As $S(C)$ acts transitively on Ω for any two maps ℓ_1 and ℓ_2 we can find a vector of permutations $\vec{p} = \{p_1, p_2, \dots, p_{\#\Omega}\}$ so that $p_i \ell_1(i) = \ell_2(i)$. We will regard such vectors $\vec{p}$ as acting on maps ℓ in this way $(\vec{p}\ell)(i) = p_i \ell(i)$. Suppose ℓ_0 is some 1-1 map, i.e. $\ell_0^*(\lambda)$ is uniform measure. Writing $R(\vec{p})$ for the "range" of $\vec{p}$, i.e. the collection of permutations that appear as coordinates in the vector, suppose $\#R(\vec{p}) \leq c$. Then for $\ell = \vec{p}\ell_0$ we have $c(\ell) \leq c$ since for any $i \neq j$ $\ell(i) = \ell(j)$ only if $p_i \neq p_j$. We seek a converse to this. That is to say we seek a vector $\vec{p}$ that involves as few permutations as possible so that $\vec{\ell}$ is nearly 1-1. Adequate control here will be very easy to obtain via a classical covering argument.

Theorem 4.1. *Suppose* $\ell : \{1, \dots, \#\Omega\} \to \Omega$. *For all* $\varepsilon > 0$ *there is a* $\vec{p}$ *consisting of at most* $\frac{c(\ell)}{\varepsilon^2}$ *permutations so that* $\ell_1 = \vec{p}\ell$ *is 1-1 on* $(1-\varepsilon)\#\Omega$ *of its domain.*

Proof. We will attempt to define $\vec{p}$ on as large a range as possible satisfying the following conditions. Suppose $\vec{p}$ has been defined on some subset I of indices so that $\vec{p}\ell$ is 1-1 where it is defined and moreover for all permutations p that appear in $\vec{p}$,

$$\#\{i : p_i = p\} \geq \frac{\#\Omega\varepsilon^2}{c(\ell)}.$$

In this case $R(\vec{p})$ consists of at most $c(\ell)/\varepsilon^2$ permutations. Choose I to be maximal under containment for these condition. If $\#I \geq (1-\varepsilon)\#\Omega$ we are done. Suppose not, i.e. $\#I^c > \varepsilon\#\Omega$. In I^c choose a maximal collection of indices $i_1, i_2, \dots, i_t$ so that $\ell(i_1), \dots, \ell(i_t)$ are all distinct. We can do this with $t \geq \#I^c/c(\ell) > \#\Omega\varepsilon/c(\ell)$ from the definition of $c(\ell)$.

Now notice that for each i_j the set $\{p\ell(i_j)\}_{p \in S(n)}$ covers Ω uniformly $\Pi_i t_i!$ times. Hence among the $tn!$ pairs (p, i_j) at least $\varepsilon tn!$ of them put $p\ell(i_j)$ outside of $\{p_i\ell(i)\}_{i \in I}$ where $\vec{p}\ell$ has already been defined. We can partition this space of pairs by the choice of value p. For at least one of these values then, call it p', at least $\varepsilon t \geq \#\Omega\varepsilon^2/c(\ell)$ of the necessarilly distinct names $p'\ell(i_j)$ must lie outside of $\{p_i\ell(i)\}_{i \in I}$. For these values i_j we extend $\vec{p}$ by setting $p_{i_j} = p'$. With this extension $\vec{p}\ell$ remains 1-1 and hence the conditions have been preserved, conflicting with the maximality of I and completing the result. ☐

We translate this into a result on adjusted measures.

Theorem 4.2. *Suppose* m_1 *and* m_2 *are to adjusted measures on* $\Omega = \Omega(F)$. *For all* $\varepsilon > 0$ *there are* ℓ_1 *and* $\ell_2 : \{1, \dots, \#\Omega\} \to \Omega$ *and a* $\vec{p}$ *so that*

(i) $\ell_i^*(\lambda) = m_i$ *for* $i = 1, 2$,
(ii) $\#\{i : \vec{p}\ell_1(i) \neq \ell_2(i)\} < \varepsilon\#\Omega$ *and*
(iii) $\#R(\vec{p}) \leq \frac{c(m_1)c(m_2)}{16\varepsilon^4}$.

Proof. Using the Theorem 4.1, find $\ell_1, \ell_2, \vec{p}_1$ and $\vec{p}_2$ with

(i) $\ell_i^*(\lambda) = m_i$,
(ii) $\#R(\vec{p}_i) \leq \frac{c(\ell_i)}{(2\varepsilon)^2}$ and
(iii) the $\vec{p}_i\ell_i$ are each 1-1 on $1 - \varepsilon/2)\#\Omega$ indices.

Let Ω_0 be the collection of $(1-\varepsilon)\#\Omega$ values ω on which both the $(\vec{p}_i\ell_i)^{-1}$ are 1-1. W.l.o.g. we can assume that for $\omega \in \Omega_0$ $(\vec{p}_1\ell_1)^{-1}(\omega) = (\vec{p}_2\ell_2)^{-1}(\omega)$ be reordering the domain of $\vec{p}_1\ell_1$. Setting $I_0 = (\vec{p}_1\ell_1)^{-1}(\Omega_0$ and setting $\vec{p} = \vec{p}_2^{-1}\vec{p}_1$ we have

(iv) $\#\{i : \vec{p}\ell_1(i) \neq \ell_2(i)\} \leq \#I_0^c \leq \varepsilon\#\Omega$ and

(v) $\#R(\vec{p}) \leq \#R(p_1)\#R(p_2) \leq \frac{c(\ell_1)}{4\varepsilon^2}\frac{c(\ell_2)}{4\varepsilon^2}$.

□

We now give a more general version of this combinatorics that is no more difficult to verify. Again assume $C \subseteq G$ is a finite set, partitioned now into a finite list of sets $C_1, C_2, \dots, C_k$. Suppose we also partition the symbol space Σ into $\Sigma_1, \dots, \Sigma_k$. We now assume we have k vectors of integers $\{t_s^i\}_{s\in\Sigma_i}$ with $\sum_{s\in\Sigma_i} t_s^i = \#C_i$. We now set $\Omega = \Omega(C_i, t_s^i)$ to be the set of maps $\omega : C \to \Sigma$ so that whenever $s \in \Sigma_i$ then $\omega^{-1}(s) \subseteq C_i$ and $\#\omega^{-1}(s) = t_s^i$. We know

$$\#\Omega = \Pi_{i=1}^k \frac{\#C_i!}{\Pi_{s\in\Sigma_i} t_s^i!}.$$

We again assume Ω is a probability space with each name being equally likely. We consider once more adjusted measures on Ω, those for which all masses are multiples of a basic size $1/\#\Omega$ and again note that such are the images of a uniform measure under a map $\ell : \{1, \dots, \#\Omega\} \to \Omega$. Now notice that all permutations in $S(C)$ no longer preserve the names in Ω. Let $H \subseteq S(C)$ be those permuations which leave the C_i invariant. This is a subgroup of $\Pi\#C_i!$ elements and it acts transitively on Ω. We can now reformulate the covering argument in terms of this more complex structure. The goal is the same, to bound the number of permutations needed to modify one adjusted measure approximately into another.

Theorem 4.3. *Suppose $\ell : \{1, \dots, \#\Omega\} \to \Omega$. For all $\varepsilon > 0$ there is a $\vec{p}$ consisting of at most $\frac{c(\ell)}{\varepsilon^2}$ permutations of H so that $\ell_1 = \vec{p}\ell$ is 1-1 on $(1-\varepsilon)\#\Omega$ of its domain.*

Proof. The same argument as in Theorem 4.1 works once one notes that the group H acting on any single name in ω covers Ω uniformly. □

We once more translate this into a result on adjusted measures. The proof is identical to that of Theorem 4.2

Theorem 4.4. *Suppose m_1 and m_2 are to adjusted measures on $\Omega = \Omega(C_1, \dots, C_k)$. For all $\varepsilon > 0$ there are ℓ_1 and $\ell_2 : \{1, \dots, \#\Omega\} \to \Omega$ and a $\vec{p}$ so that*

(i) $\ell_i^*(\lambda) = m_i$ *for* $i = 1, 2$,

(ii) $\#\{i : \vec{p}\ell_1(i) \neq \ell_2(i)\} < \varepsilon\#\Omega$ *and*

(iii) $\#R(\vec{p}) \leq \frac{16c(m_1)c(m_2)}{\varepsilon^4}$.

We now translate this into the realm of probability spaces (A, μ) by considering a measure m on Ω as coming from a map η that assigns Σ, C-names to points $x \in A$. If $m = \eta^*(\mu)$ is adjusted and A is nonatomic then we can partition each element $\eta^{-1}(\omega)$ into sets of size $1/\#\Omega$ and regard maps of the form ℓ as ordering these small sets. The construction of $\vec{p}$ in the previous theorem can now be regarded as assigning a permutation of its name to each of the small sets in this partition. That is to say if $p : A \to \Omega$ then we can write $p\eta$ for the map $(p\eta)(x) = p(\eta(x))$. We now can regard p as acting on the measure m by $pm(\omega) = \mu((p\eta)^{-1}(\omega))$. We now translate our combinatorial work as a corollary.

Corollary 4.5. *Suppose (A_1, μ_1) and (A_2, μ_2) are two nonatomic probability spaces and $\eta_i : A_i \to \Omega = \Omega(C_i, t^i_s)$. Let $m_i = \eta^*_i(\mu_i)$ and assume the m_i are adjusted. For all $\varepsilon > 0$ there is a map $p : A_1 \to H$ so that*

(i) $\sum_{\omega\in\Omega} |pm_1(\omega) - m_2(\omega)| \le \varepsilon$ *and*
(ii) $\#R(p) \le \frac{16c(m_1)c(m_2)}{\varepsilon^4}$.

We finish this section by removing the need for the m_i to be adjusted.

Theorem 4.6. *Suppose (A_i, μ_i), $i = 1, 2$ are two nonatomic probability spaces and the maps $\eta_i : A_i \to \Omega = \Omega(C_1, t^i_s)$ as above and let $m_i = \eta^*_i(\mu_i)$. Set $d_i = \#Range(\eta_i)$ and $M_i = \max_{\omega\in\Omega} m_i(\omega)$. For all $\varepsilon > 0$ there is a map $p : A_1 \to H$ so that*

(i) $\sum_{\omega\in\Omega} |pm_1(\omega) - m_2(\omega)| < \varepsilon$ *and*
(ii) $\#Range(p) \le \left(\frac{20{,}736}{\varepsilon^8}\right) M_1 M_2 \#\Omega^2$

Proof. We work on A_1 and will have a symmetric argument on A_2. Regard A_1 as the interval $[0, 1)$ cut into d_1 intervals $I_1, \dots, I_{d_1}$ of lengths $\mu_1(\eta_1^{-1}(\omega))$, $\omega \in \text{range}(\eta_1)$. Set $N_1 = \lfloor \frac{2d_1}{\varepsilon\#\Omega} \rfloor + 1$ and cut $A_1 = [0, 1)$ into $N\#\Omega$ intervals of equal length. Now at most $d_1 - 1$ of these intervals are cut by the I_j. Modify η_1 to $\hat\eta_1$ by changing it only on these $d_1 - 1$ intervals so that $\hat\eta_1$ is a constant on all of these $N\#\Omega$ intervals. This means

$$\mu_1(\{x : \eta(x) \ne \hat\eta(x)\}) < (d_1 - 1)\frac{1}{N\#\Omega} < \varepsilon/2.$$

Now cut $[0, 1)$ into N pieces of equal length, each consisting of precisely $\#\Omega$ of the little intervals. Renormalizing μ_1 to each of these pieces, and with regard to $\hat\eta_1$, we have a collection of N_1 adjusted measures on Ω which we write as $m^1_1, \dots, m^{N_1}_1$. We can do the same on (A_2, μ_2) constructing $\hat\eta_2$ and getting $N_2 = \lfloor \frac{2d_2}{\varepsilon\#\Omega} \rfloor + 1$ adjusted measures $m^1_2, \dots, m^{N_2}_2$ on subintervals. We calculate that each $c(m^j_i) \le M_i N_i \#\Omega + 1 < 2M_i N_i \#\Omega$. Cut each m^i_1 into N_2 identical copies and each m^i_2 into N_1 identical copies and renormalize the measures. Now pair up the $N_1 N_2$ adjusted measures into which we have cut A_1 and A_2. Apply Corollary 4.5, using $\varepsilon/2$ to each pair to construct a map p. For each pair the range of p will have cardinality at most

$$\frac{64c(m^i_1)c(m^j_2)}{\varepsilon^4} < \frac{256 M_1 N_1 M_2 N_2 \#\Omega^2}{\varepsilon^4}.$$

The map p is now defined on all of A_1 as the combination of all these maps on subsets and hence

$$\#\text{Range}(p) \le N_1 N_2 \frac{256 M_1 N_1 M_2 N_2 \#\Omega^2}{\varepsilon^4}.$$

Noting that $N_i \le \frac{2d_i+\varepsilon\#\Omega}{\varepsilon\#\Omega} \le \frac{3}{\varepsilon}$ (as $d_i \le \#\Omega$) gives the bound (ii). For (i) just note that we have now matched the $\hat\eta_i$ names on all but $\varepsilon/2$ of the measures and removing the ˆ 's adds a further $\varepsilon/2$ to this error. ☐

To show the reader where this estimate is leading, note that in our later work we will work in systems of some entropy h and on some large Følner set F. The value $\#\Omega$ will be bounded above by something like $2^{(h+\delta)\#F}$ and the M_i will be bounded above by something like $2^{-(h-\delta)\#F}$. In this situation $\#\mathrm{Range}(p) \le \frac{20{,}736}{\varepsilon^8}2^{2\delta\#F}$, and will have "small entropy".

We now turn to the final construction using the fundamentals laid out by Ornstein and Weiss [4], [3] for ergodic actions of amenable groups, in particular the the Rokhlin Lemma, ergodic theorem and Shannon-McMillan theorem. This will lead to our final conclucion, that all actions of equal entropy are m^e-equivalent. It should be clear that the permutations $\vec{p}$ found in the above work will be used to generate the elements ϕ to show that all actions are r.w. m^e-f.d. and that the control on the $\#\mathrm{Range}(\vec{p})$ will control $E(T^\alpha, P^{\alpha,\phi})$. To control $m_0(\alpha,\phi)$ we will need to use the Ornstein-Weiss Rokhlin Lemma machinery.

5. Identifying m^e-equivalence

In the following discussion (T^α, P) and $(T_0^{\alpha_0}, P_0)$ will be two Σ-valued processes, where both are free ergodic G-actions of the same entropy $h < \infty$. These will be the two processes involved in the statement of r.w. m^e-f.d. We restate this definition here for reference.

Definition 5.1. *We say a process (T^α, P) is reverse-weakly m-finitely determined (r.w. m-f.d.) if for all $\varepsilon > 0$ there is a δ so that if we have an ergodic $(T_0^{\alpha_0}, P_0)$ satisfying:*

(i) $\|(T^\alpha, P), (T_0^{\alpha_0}, P_0)\|_* < \delta$ *and*
(ii) $h_m(T_0^{\alpha_0}, P_0) \ge h_m(T^\alpha, P) - \delta$

then the following holds. For all $\delta' > 0$ there is an ergodic lift $T_1^{\alpha_1}$ of T^α (i.e. there is a π factoring the $T_1^{\alpha_1}$ action onto T^α) and a ϕ_1 in the full group of $\mathcal{O}(T^\alpha)$ and a Σ-valued partition P_1 with:

(a) $\mu_1(\pi^{-1}(P) \triangle P_1) < \varepsilon$ *and*
(b) $m(\alpha, \phi_1) < \varepsilon$ *and moreover*
(i)′ $\|(T_1^{\alpha_1\phi_1}, P_1), (T_0^\alpha, P_0)\|_* < \delta'$ *and*
(ii)′ $h_m(T_1^{\alpha_1}, P_1) \ge h_m(T_0^\alpha, P_0) - \delta'$.

Many facts we state will be true of both processes in which case we abbreviate them as $(T_*^{\alpha_*}, P_*)$. That is to say $*$ is either no symbol or 0. (We will have a few words to say about the infinite entropy case at the end.) For a finite subset $F \subseteq G$ we write Σ^F for the set of Σ, F-names. We write $P_*^{-1}(\eta)$ for the set of points in X_* whose $T_*^{\alpha_*}, P_*, F$-name is η etc. We assume a value $\varepsilon > 0$ is now fixed and we will work toward a choice for δ in the definition.

In terms of quantifiers a value δ_0 will be determined by (T^α, P) and ε and will help determine determine δ. We will then assume and fix the process $(T_0^{\alpha_0}, P_0)$ and then we will fix $\delta' > 0$ which will determine for us a value δ_1 with which we will carry out the construction of the lift of $T_1^{\alpha_1}$ of T^α and the full-group element ϕ and partition P_1 giving the conclusions of the definition of r.w. m^e-f.d.

We first require of $\delta_0 > 0$ and a finite subset F^0 of G sufficiently invariant that whenever

$$\mu(\{x : T^\alpha_{,g}(\phi(x)) = \phi(T^\alpha_g(x)) \text{ for all } g \in F^0\}) > 1 - \delta_0$$

then $m^0(\alpha, \phi) < \varepsilon/10$.

It will be extremely convenient to use auxiliary "cookie cutter" processes that will lay out for us on orbit names the regions that we will be permutating. For this we take a free Bernoulli G-action of $(Z, \mathcal{G}, \nu, U^\gamma)$ with $h(U^\gamma) < \delta_0/10$. Our two applications of the Rokhlin Lemma will be in two independant copies of this auxiliary action. We will make our basic construction on the lift $T_1^{\alpha_1} = T^\alpha \times U^\gamma \times U^\gamma$ acting on $X_1 = X \times Z \times Z$ with product measure. At this point we only add the first copy of Z. The second will come later. To continue the construction, by the Rokhlin Lemma given explicitely in Theorem 3.0.3 of [3] we can find subsets $A_1, \ldots, A_k \subseteq Z$ and finite subsets $F_1, \ldots, F_k \subseteq G$ so that if we define $V : \bigcup_i (A_i \times F_i) \to Z$ by $V(z, g) = U_g^\gamma(z)$ then

(a) $\frac{\#(\cap_{g \in F^0}(F_i g))}{\#F_i} < \delta_0/10$ for $i = 1, \ldots, k$,
(b) V is 1-1 and $\nu(\text{Range}(V)) > 1 - \delta_0/10$.

Notice that as we need no control over the number of sets k here we can disjointify the usual quasi-towers of the Rokhlin Lemma, as shown in Theorem 3.0.3. What statement (a) adds says is that the F_i are sufficiently invariant for later purposes. We control the entropy of the tower process by putting this tower in Z. We assume at this point that the process (T^α, P) is fixed but δ is not yet chosen and hence $(T_0^{\alpha_0}, P_0)$ is not yet known.

Using the Shannon-McMillan Theorem (Corollary 3.0.10 of [3]) we can also require that the F_i are sufficiently invariant that for each i there are subsets of names $G_i \subseteq \Sigma^{F_i}$ so that

(a) for all $\eta \in G_i$, $\mu(G_i) > 1 - \delta_0/10$ and
(b) $\mu(\eta) = 2^{-h(T^\alpha, P) \pm \delta_0/10)\#F_i}$.

We now require that δ be so small that (i) and (ii) also imply that $\mu_0(G_i) > 1 - \delta_0/10$ and for $\eta \in G_i$ we get the same bound

$$\mu_0(\eta) = 2^{-(h(T_0^{\alpha_0}, P_0) \pm \delta_0/10)\#F_i}.$$

As an explicit bound we ask that δ be so small that (i) gives us

$$\sum_{\eta \in \Sigma^{F_i}} |\mu_1(\eta) - \mu_2(\eta)| < \delta_0/10.$$

At this point δ_0 and δ have been specified and the process $(T_0^{\alpha_0}, P_0)$ satisfying (i) and (ii) of r.w. m^e-f.d. has been fixed. We now fix a $\delta' > 0$ and proceed.

For any finite valued partition Q we will write $\text{dist}_\mu(Q)$ for the measure $Q^*(\mu)$ on the symbol space of Q and similarly define conditional $\text{dist}_\mu(\quad |A)$ given a subset A as the image of the measure normalized on A. It will also be convenient, if $A \subseteq X_*$ to also regard A as the subset $A \times Z$ of $X_* \times Z$. In these terms we get from the independence of $\mathcal{F}$ and $\mathcal{G}$ that for all $B \subseteq Z$,

$$\text{dist}_{\mu \times \nu}(\bigvee_{g \in F_i} T_{1,g^{-1}}^{\alpha_1}(P)|B) = \text{dist}_\mu(\bigvee_{g \in F_i} T_{1,g^{-1}}^{\alpha_1}(P)).$$

First a proposition to prepare for the next level of construction.

Proposition 5.2. *For δ_0 and δ already chosen, for any choice of $\delta_1 > 0$, for any sufficiently invariant $\overline{F}$ in G we will have a subset $B_2 = B_2(F) \subseteq Z$, $\nu(B_2) < \frac{\delta_1}{10}$ and for $z \notin B_2$, subsets $C_i = C_i(z) \subseteq G$, $i = 1, \dots, k$ in $\overline{F}$ with $U_c^\gamma(z) \in A_i$ for all $c \in C_i(z)$ with*

(i) *for $c \in C_i$, $F_i c \subseteq F$,*

(ii) *for $i = 1, \dots, k$, $\left| \dfrac{\# C_i}{\# F} - \nu(A_i) \right| < \delta_0/10$,*

(iii) *$\#\big(\cup_{i=1}^k \cup_{c \in C_i} F_i c\big) > (1 - \delta_0/10)\#F$ and the union is disjoint of course.*
In addition from the Shannon-McMillan Theorem applied to U^γ

(iv) *the number of distinct choices for the list of sets $\{C_i(z)\}$ is at most $2^{\frac{\delta_0}{10}\#\bar{F}}$.*

If in addition we are given $(T_0^{\alpha_0}, P_0)$ satisfying (i) and (ii) of r.w. m^e-f.d., then for all but a subset $B_* = B_*(F) \subseteq X_*$ of measure $< \delta_1/10$, if $x_* \notin B_*$

(v) *for each i and any name $\eta \in \Sigma^{F_i}$*

$$\left| \frac{\#\{c \in C_i : T_{*,c}^{\alpha_*}(x_*) \in \eta\}}{\# C_i} - \mu_j(\eta) \right| < \delta_1/10$$

and from the Shannon-McMillan Theorem applied to $T_*^{\alpha_*}$,

(vi) *For $\eta_{\bar{F}}(x_*) \in \Sigma^{\bar{F}}$, the $T_*^{\alpha_*}, P_*, \bar{F}$-name of x_*, we have*

$$\mu_*(\eta_{\bar{F}}(x_*)) = 2^{-(h(T_*^{\alpha_*}, P_*) \pm \delta_1)\#\bar{F}}.$$

Proof. Statements (i),(ii), (iii), and (v) follow directly from the mean Ergodic Theorem applied to appropriate processes and from choosing $\bar{F}$ sufficiently invariant. Statements (iv), as is said, follows from the Shannon-McMillan Theorem applied to the process U^γ which is assumed of entropy less than $\delta_0/10$. Statement (vi) follows from the Shannon McMillan Theorem applied to $(T_*^{\alpha_*}, P_*)$. □

We take a second copy of our cookie cutter process U^γ now and in it take Rokhlin towers $(\bar{A}_j, \bar{F}_j)$, $j = 1, \dots, \bar{k}$ all assumed sufficiently invariant to play the role of $\bar{F}$ in the previous proposition and covering all but $\delta_1/10$ of Z. As we continue we will require greater invariance of the $\bar{F}_j$.

Partition each $B_2(\bar{F}_j)^c \times \bar{A}_j$ in $Z \times Z$ according to the values of $C(z) = \{C_i(z)\}$ of Proposition 5.2 and write this partition of $\cup_j B_*(\bar{F}_j) \times \overline{A}_j$ as $\{D_1, D_2, \dots, D_m\}$ and note that

$$m \leq \sum_{j=1}^{\bar{k}} 2^{-\frac{\delta_0}{10}\#\bar{F}_j}.$$

It is important to remember that the sets D_t are $Z \times Z$ measurable.

Our goal now is to describe the construction of ϕ and the modification of the partition P to P_1. Fix one of these sets D_t to work on. In keeping with our use of "cookie cutter" we will refer to the slices through the towers labeled by a subset of G as "cookies". Here then we see a large "cookie" $\bar{F}_j \subset G$ with many smaller disjoint "cookies" marked out on it at $F_i c$, $c \in C_i$.

For these choices we are ready to establish the combinatorial situation discussed earlier. Set $\bar{\Sigma}_i$ to be those $\eta \in G_i$ that is to say satisfying

$$\mu(\eta) = 2^{-(h(T^\alpha, P) \pm \delta_0)\#F_i}$$

along with a dummy symbol $\star_i$. Let $\bar\Sigma = \cup\bar\Sigma_i$. Let $C = \cup C_i$. For non-$\star_i$ elements $\eta \in \bar\Sigma_i$ choose the integer value $t_\eta = \lfloor \mu(\eta)(1-\delta_1/10)\#C_i \rfloor$ and set $t_{\star_i}$ so that $\sum_{s\in\bar\Sigma_i} t_x = \#C_i$. With this we can define Ω_t to be all $\bar\Sigma, C$-names subject to the constraint that names in C_i must come from Σ_i and the number of occurrences of name s must be t_s. We now partition $D_t \times B$ according to the Σ^{F_i} name of $T^{\alpha_1}_{1,c}(x)$ for each $c \in C_i$ and each $i = 1, \ldots, k$. Call this partition N, the partition into names across the smaller cookies. By lumping together those Σ^{F_i} names which are not in Σ_i to $\star_i$ we get a map $\bar\eta$ from D_t to $C^{\bar\Sigma}$. We have restricted the X coordinate to B^c and so it consists of points with good averages along the C_i. Any $x \notin B$ will have at least t_η occurrences of the name $\eta \in \Sigma_i$ among the $T^{\alpha_1}_1, P, F_i$-name of $T^{\alpha_1}_{1,c}(x)$, $c \in C_i$. Modify $\bar\eta$ at such an x by mapping any excess occurences beyond t_η of such an $\eta \in \Sigma_i$ to the dummy symbol $\star_i$. Extend $\bar\eta$ to the remainder of $D_t \times B$ by simply assigning some allowed name $\eta \in \Omega_t$.

The names $\bar\eta$ can be expanded to be $\Sigma \cup \{\star_i\}$ names on the smaller cookies. We can extend these to agree with the P name on all other indices. By painting these expanded $\bar\eta$ names on the towers over the $\bar A_j$ we can regard this construction as modifying the partition P to a new partition $\bar P$ of the lift space $X \times Z \times Z$with values in $\Sigma \cup \{\star_i\}$ and can consider the map $\bar\eta$ as mapping $\cup_j B^c(\bar F_j) \times \overline{A}_j$ to the $\overline{P}$ names that occur accross the smaller cookies. In this painting we do not modify the P names outside of these smaller cookies. A completely symmetric construction can be carried out in $(T^{\alpha_0}_0, P_0)$ gives a map $\bar\eta_0$ on $\cup_j \overline{A}_j$ but now in $X_0 \times Z \times Z$. Let $\bar\mu_* = \mu_* \times \nu \times \nu$. Our next step is to apply our basic combinatorics to describe how to permute these names. To prepare for that we collect the current status of this construction into the following lemma:

Lemma 5.3. *From the construction just described, if the $\bar F_j$ are sufficiently invariant then we can conclude the following:*

(a) *The tower images of the larger towers $(B_2(\bar F_j)^c \times A_j, \bar F_j)$ have total measure given by*

$$\sum_j \bar\mu_*(B^c_*(\bar F_j) \times A_j)\#\bar F_j > 1 - \delta_1/5.$$

(b) *The measure of the subset of these tower images that lies inside the "smaller cookies" can be written as*

$$\sum_{t=1}^{m} \bar\mu_*(D_t)\Big(\sum_{C_i \in C(D_t)} \#C_i \#F_i\Big) > 1 - \delta_0/10 - \delta_1/5.$$

(c) $\bar\mu_*(P_* \triangle \bar P_*) < \delta_0/10 + 7\delta_1/5$.

(d) *The number of sets D_t in each $B^c_*(\bar F_j) \times A_j$ is at most $2^{-\frac{\delta_0}{10}\#\bar F_j}$.*

(e) *For each t, where $D_t \subseteq \bar A_j$,*

$$\#\Omega_t < 2^{(h(T^{\alpha_*}_*, P_*) + \frac{2\delta_0}{5})\#\bar F_j}.$$

(f) *For each t with $D_t \subseteq \bar A_j$ and $\omega \in \Omega_t$, if $\omega \in$ range$(\bar\eta_*)$ then for all but a set of $\bar\eta_*^{-1}(\omega)$ of relative measure less than δ_1 in D_t,*

$$2^{-(h(T^{\alpha_*}_*, P_*) - \log_2(\#\Sigma)\delta_0)\#\bar F_j} > \frac{\bar\mu_*(\bar\eta_*^{-1}(\omega))}{\bar\mu_*(D_t)} > 2^{-(h(T^{\alpha_*}_*, P_*) + \delta_1)\#\bar F_j}$$

Proof. For (a) note that the original tower images were chosen to cover all but $\delta_1/10$ of the spaces and that the sets B_2 have measures at most $\delta_1/10$. For (b) one observes that the small cookies cover all but $\delta_0/10$ of each of the larger towers. For (c) note that the only place where the partitions P_* and $\bar{P}_*$ differ is inside smaller cookies. The changed set consists of three pieces. The first is where the name on the smaller cookie is not in G_i. The second comes from the excess names above t_η in number. As the number of occurrences of η among all the $c \in C_i$ is at most $\mu_*(\eta)(1+\delta_1/10)$ the set of points mapped now to $\star_i$ in this second case has measure at most $3\delta_1/10$. The third are all the cookies over B_*. The first set is at most $\delta_0/10$ of the entire space. For the second we will assume that the $\#\bar{F}_j$ are large enough that $t_\eta > \mu_*(\eta)(1-\delta_1/5)\#C_i$. The third piece has measure bouned by that of B_*, i.e. $\delta_1/10$. We already described the bound (d) as coming from the Shannon-McMillan theorem on U^γ. For (e) note that the number of possible names η coming from G_i is at least $2^{(h(T_*^{\alpha*},P_*)-\delta_0/10)\#F_i}$. Each such name covers $\#F_i$ indices in a smaller cookie. These sets are disjoint and cover a fraction $1-\delta_1/5-\delta_0/10$ of $\bar{F}_j$. Knowing $\delta_1 \leq \delta_0$ we get the estimate.

For the right half of the last estimate (f) note that the $\bar{\eta}_*^{-1}(\omega)$ is a union of complete $T_*^{\alpha_*}, P_*, \bar{F}_j$-names. If the $\bar{F}_j$ are sufficiently invariant the Shannon-McMillan Theorem will tell us that for all but a set of $\delta_1/2$ in measure of such names, they are at least as large as the right hand side of the estimate.

For the left hand estimate we must think a bit more. To begin, modify P_* by replacing names across a tower image of (F_i, A_i) that are not in G_i with the symbol $\star_i$. Also replace all replace all symbols that lie outside of these towers by $\tilde{\star}$. This produces a partition $\tilde{P}_*$ of $X_* \times Z$ whose entropy we can estimate by:

$$h((T_*^{\alpha_*} \times U^\gamma), \tilde{P}_*) \geq h(T_*^{\alpha_*}, P_*) - \delta_0/10 - \log_2(\#\Sigma)\delta_0/5.$$

In this estimate the $\delta_0/10$ is from the entropy of U^γ which determines where the towers sit and the $\log_2(\#\Sigma)\delta_0/5$ is a bound on the entropy lost in erasing the names in G_i and outside of the towers as these cover at most a fraction $\delta_0/5$ of large names. A name $\bar{\eta}_*^{-1}(\omega)$ will be a union of at most $(\#\Sigma)^{\delta_1\#\bar{F}_j}$ atoms of the $(T_*^{\alpha_*} \times U^\gamma, \tilde{P}_*)$ process obtained by removing the smaller cookies that overlap the boundary of $\bar{F}_j$ and by placing $\star_i$ on the excess occurrences of names in the remaining smaller cookies. In each D_t if the $\bar{F}_j$ is sufficiently invariant, for all but $\delta_1/2$ of the $\bar{\eta}_*^{-1}(\omega)$, (those which are a union of this many such atoms, where at least half in measure have size controlled by the Shannon-McMillan Theorem) will have their relative measure bounded by

$$\frac{\bar{\mu}_*(\bar{\eta}_*^{-1}(\omega))}{\bar{\mu}_*(D_t)} < 2 \cdot 2^{-(h(T_*^{\alpha*},P_*)-\delta_0/10-\log_2(\#\Sigma)(\delta_0/5+3\delta_1/5))\#\bar{F}_j} < 2^{-(h(T_*^{\alpha*},P_*)-\log_2(\#\Sigma)\delta_0)\#\bar{F}_j}.$$

□

As the penultimate step in the construction of ϕ and P_1 we will define maps $\bar{p}$ and $\bar{\rho}$ on $\cup_c D_t$. these maps will be the union of maps $\bar{p}_t$ and $\bar{\rho}_t$ each defined on D_t. The map $\bar{p}_t$ will be to permutations of $\bar{F}_j$ and $\bar{\rho}_t$ will be to names in $\Sigma^{\bar{F}_j}$. On each set D_t the set of centers $C(t)$ gives a restricted subgroup of allowed permutations $H(t)$ that fixes each $C_i \in C(t)$. Let $I(t) = \cup_{C_j \in C(t)} \cup_{c \in C_i} F_i c$, that is to say the union of the smaller cookies given by $C(t)$ inside of $\bar{F}_j$. Any $p \in H(t)$ lifts to a permutation $\bar{p}_p$ of $\bar{F}_j$ by setting

(i) $\bar{p}_p(g) = g$ for $g \notin I_t$ and
(ii) $\bar{p}_p(gc) = gp(c)$ for $gc \in F_i c$, $c \in C_i \in C(t)$.

The range of $\bar{p}$ gives a subgroup $\bar{H}(t)$ of permutations of $\bar{F}_j$.

In a similar manner notice that any name $\omega \in \Omega_t$ lifts to a name $\rho(\omega) \in \Sigma^{I_t}$ by expanding $\omega(c) \in \Sigma^{F_i}$ to be the $\rho(\omega)$ name on $F_i c$. Notice that restricted to I_t we have

$$\rho(p(\omega)) = \bar{p}_{p(\omega)}(\rho(\omega))$$

as all we have done is to "expand" the names across the smaller cookies which are then permuted by a rigid right translation.

In the construction of $\bar{\rho}_t$ and $\bar{p}_t$ we want that $\bar{\rho}_t(x)$ should be close in $\bar{d}$ to the $T^\alpha, P, \bar{F}_j$-name of x and that

$$\operatorname{dist}(\bar{p}_t\bar{\rho}_t|D_t) = \operatorname{dist}(\vee_{g\in\bar{F}_j} T^{\alpha_1}_{1,g^{-1}}(P_1)).$$

Our first step toward this is to further modify the partitions $\bar{P}$ and $\bar{P}_0$ to $\widehat{P}$ and $\widehat{P}_0$ by replacing the names assigned on the sets of relative measure δ_1 which fail the bounds given in (f) of Proposition 5.3 by copies of the distributions on their compliments in D_t. We set $\hat{\eta}_*$ to be the corresponding map to Ω_t names we obtain from these partitions. This gives us new versions of (c) and (f) for these partitions and names.

(c) $\bar{\mu}_*(P_*\triangle\widehat{P}_*) < \delta_0/10 + 8\delta_1/15$.
(f) For each t with $D_t \subseteq A_j$ and $\omega \in \Omega_t$, if $\omega \in \operatorname{range}(\hat{\eta}_*)$ then

$$2^{-(h(T_*^{\alpha*},P_*)-\log_2(\#\Sigma)\delta_0-\delta_1)\#\bar{F}_j} > \frac{\bar{\mu}_*(\hat{\eta}_*^{-1}(\omega))}{\bar{\mu}_*(D_t)} > 2^{-(h(T_*^{\alpha*},P_*)+\delta_1)\#\bar{F}_j}.$$

We are now ready to apply our combinatorial work.

Theorem 5.4. *Using the construction described so far there exist maps $p_t : D_t \to H(t)$ so that setting $\widehat{m}_* = \hat{\eta}_*^*(\bar{\mu}_*)/\bar{\mu}_*(D_t)$ then*

(i) $\sum_{\omega\in\Omega_t} |p_t(\widehat{m})(\omega) - \widehat{m}_0(\omega)| \le \delta_1$. *If $\#(\bar{F}_j)$ is large enough and $\delta_0 < \frac{\varepsilon}{10\log_2(\#\Sigma)}$ and $\delta_1 < \delta_0/20$ then*
(ii) *$\#Range(p_t) \le 2^{\varepsilon\#\bar{F}_j/3}$.*

Proof. Theorem 4.6 constructs p_t satisfying (i) and with

$$\begin{aligned}\#\mathrm{range}(p_t) \le& \left(\frac{20,736}{\delta_1^8}\right) 2^{-(h(T^\alpha,P)-\log_2(\#\Sigma)\delta_0-\delta_1)\#\bar{F}_j} 2^{(h(T^\alpha,P)+\frac{2\delta_0}{5})\#\bar{F}_j} \times \\ & 2^{-(h(T_1^{\alpha_1},P_1)-\log_2(\#\Sigma)\delta_0-\delta_1)\#\bar{F}_j} 2^{(h(T_1^{\alpha_1},P_1)+\frac{2\delta_0}{5})\#\bar{F}_j} \\ \le& \left(\frac{20,736}{\delta_1^8}\right) 2^{(2\log_2(\#\Sigma)\delta_0+\frac{4}{5}\delta_0+2\delta_1)\#\bar{F}_j}\end{aligned}$$

If $\#\bar{F}_j$ is large enough then this is

$$\le 2^{3\log_2(\#\Sigma)\delta_0\#\bar{F}_j} < 2^{\varepsilon\#\bar{F}_j}.$$

□

We can now define $\bar{p}_t$ to be the lift to $\bar{H}(t)$ of p_t. Statement (i) above now allows us to construct measure preserving and invertible maps $q_t : D_t \times X \to D_t \times Y$ so that

$$\bar{\mu}(\{\bar{x} \in D_t \times X : p_t(\bar{x})\hat{\eta}(x) \neq \hat{\eta}_1(q_t(\bar{x}))\}) < \delta_1 \bar{\mu}(D_t).$$

We are also ready to define $\bar{\rho}_t$. For $\bar{x} \in D_t \times X$ let $\eta(\bar{x})$ be the $T^{\alpha}, P, \bar{F}_j$-name of x and for $\bar{y} \in D_t \times Y$ let $\eta_0(x)$ be the $T_0^{\alpha_0}, P_0, \bar{F}_j$-name of y. Define $\bar{\rho}_t(x) = (\bar{p}_t(\bar{x}))^{-1}\eta_0(q_t(\bar{x}))$. That is to say, we take the P_0 name of the q_t image point and "un-permute" it by $(\bar{p}_t)^{-1}$.

Theorem 5.5.

$$\int_{D_t} \bar{d}_{\bar{F}_j}(\bar{p}_t(\bar{x})\eta(\bar{x}, \eta_0(q_t(\bar{x}))\, d\mu_1 = \int_{D_t} \bar{d}_{\bar{F}_j}(\eta(\bar{x}), \bar{\rho}_t(\bar{x})) < \left(\frac{3\delta_0}{10} + \frac{17\delta_1}{10}\right)\bar{\mu}(D_t).$$

If in addition one requires $\delta_0 < \frac{1}{2}\left(\frac{\varepsilon}{10}\right)^2$ *then for each set* D_t*, for all but a fraction* $\left(\frac{\varepsilon}{10}\right)$ *in measure of the* $\bar{x} \in D_t$ *we will have*

$$\bar{d}_{\bar{F}_j}(\eta(\bar{x}), \bar{\rho}_t(\bar{x})) < \frac{\varepsilon}{10}.$$

Proof.

$$\begin{aligned}
\int_{D_t} \bar{d}_{\bar{F}_j}(\eta(\bar{x}), \bar{\rho}_t(\bar{x})) \leq & \bar{\mu}(D_t)\left(1 - \frac{\#I_t}{\#\bar{F}_j}\right) \\
& + \int_{D_t} \bar{d}_{I_t}(\eta(\bar{x}), \rho(\bar{\eta}(barx))) + \int_{D_t} \bar{d}_{I_t}(\rho(\bar{\eta}(\bar{x})), \rho(\hat{\eta}(\bar{x}))) \\
& + \int_{D_t} \bar{d}_{I_t}(\rho(\hat{\eta}(\bar{(}x)), \rho(\hat{\eta}_0(q_t(\bar{x})))) + \int_{D_t} \bar{d}_{I_t}(\rho(\hat{\eta}_0(q_t(\bar{x}))), \eta_1(q_t(\bar{x}))) \\
\leq & \left(\frac{\delta_0}{10} + 2\left(\frac{\delta_0}{10} + \frac{7\delta_1}{10}\right) + 2\delta_1\right)\bar{\mu}(D_t) < 2\delta_0\bar{\mu}(D_t).
\end{aligned}$$

□

The last step in showing (T^{α}, P) is r.w. m^e-f.d. is a copying lemma. The pieces of this lemma are rather standard, catching good entropy and distribution, and hence we will only outline the construction. First a brief sketch. The map q_t assigns to each $T^{\alpha}, P, \bar{F}_j$-name occuring in D_t a collection of $T_1^{\alpha_1}, P_1, \bar{F}_j$-names, those which intersect its q_t-image. We wish to select various of these to paint onto the $\bar{F}_j$ tower over this name to obtain the δ'-closeness we require in both entropy and distribution. We have the added necessity of carrying along the permutations $\bar{p}_t$ which we will use to construct ϕ. Here is more detail.

As we are allowed to lift the action T^{α} after being given $(T_0^{\alpha_0}, P_0)$ we can assume w.l.o.g. that $h(T^{\alpha}) \geq h(T_0^{\alpha_0}, P_0)$. We select a partition R that refines P with $h(T^{\alpha}, R) \geq h(T_0^{\alpha_0}, P_0)$.

Theorem 5.6. *Assuming the construction up to this point, for all* $\delta' > 0$*, if* δ_1 *is sufficiently small and* $\bar{F}_j$ *are sufficiently invariant then within each atom* $a \in \vee_{g \in \bar{F}_j} \bar{T}^{\bar{\alpha}}_{g^{-1}}(R)\big|_{D_t}$ *one can selest a point* $\bar{x}_a \in a$ *so that if* $\eta_0(q_t(\bar{x}_a))$ *(the* $\bar{T}_0^{\alpha_0}, P_0, \bar{F}_j$*-name of* $q_t(\bar{x}_a)$*) is painted on the* $\bar{F}_j$ *tower over the set a then the partition* P_2 *so obtained will satisfy*

(a) $\|(\bar{T}^{\bar{\alpha}}, P_2), (T_0^{\alpha_0}, P_0)\|_* < \delta'$ *and*
(b) $h(\bar{T}^{\bar{\alpha}}, P_2) > h(T_0^{\alpha_0}, P_0) - \delta'$.

Further, if $\bar{p}_t(\bar{x}_a) = \bar{p}_t(\bar{x}_a)\eta_0(q_t(\bar{x}_a))$ *is painted on this same tower image then the partition* P_1 *so obtained will satisfy*

(c) $\mu_1(P \triangle P_1) < \varepsilon$.

Proof. Statement (a) is the most standard. As we are copying from the full distribution of $T_0^{\alpha_0}, P_0, \bar{F}_j$-names, if δ_1 is small enough and the $\bar{F}_j$ are sufficiently invariant, then there is a subset $G_1 \subseteq Y$ with $\mu_1(G_1) > 1-\delta_1$ that consists of points sufficiently "good in distribution" so that if all but a fraction $\delta'/10$ of the sets a have $q_t(\bar{x}_1)$ in G_1 then (a) will hold. This goodness in distribution of most points is just a consequence of the ergodic theorem.

Next we need to catch entropy. From the Shannon-McMillan theorem we will have a subset G_2 of all but a fraction δ_1 of the sets a which have "controlled size" in that they have measure $2^{-(h(\bar{T}^{\bar{\alpha}},R)\pm\delta_1)\#\bar{F}_j}$. Similarly for a set G_3 of all but δ_1 in measure of the atoms $b \in \vee_{g\in\bar{F}_j} T_{0,g^{-1}}^{\alpha_0}(P_0)$ we have controlled size in that

$$\mu_0(b) = 2^{-(h(T_0^{\alpha_0},P_0)\pm\delta_1)\#\bar{F}_j}.$$

Cut each atom b into $M = \lfloor 2^{(h(\bar{T}^{\bar{\alpha}},R)-h(T_0^{\alpha_0},P_0)+4\delta_1)\#F_j} \rfloor$ pieces labeled $(b,1),(b,2),\ldots,(b,M)$ all of equal size. For $b \subseteq G_3$ this size is at least a fraction $2^{-\delta_1\#\bar{F}_j/10}$ smaller than the size of any atom $a \subseteq G_2$. Assume any atom a "knows" any atom (b,j) if $q_t(D_t \times a) \cap (D_t \times b_0) \neq \emptyset$.

We now seek to assign unique atoms (b,j) to atoms a which know them. Do this in the following order. First attempt to match a's with controlled size and good distribution to atoms (b,j) which also have controlled size but where in addition their is an $\bar{x}_a \in a$ so that $q_t(\bar{x}_a) \in b$ and $\bar{d}_{\bar{F}_j}(\bar{p}_t(\bar{x}_a)\eta_0(q_t(\bar{x}_a),\eta(\bar{x}_a)) < \varepsilon/3$. Each name (b,j) so assigned will have measure only a fraction $2^{-\delta_1\#\bar{F}_j}$ of the measure of the set a to which it has been assigned. It follows that under these conditions we can cover all but a fraction $\varepsilon/3 + 3\delta_1 + 2^{-\delta_1\#\bar{F}_j} < \varepsilon/3$ of D_t by sets a given a uniquely assigned b. Now remove the requirement that the permuted P_0-name of b should be within $\varepsilon/3$ in $\bar{d}$ of the P-name of a and continue to assign unique b's to a's, picking representative points x_a. When finished at most a fraction $3\delta_1 + 2^{-\delta_1\#\bar{F}_j} < 4\delta_1$ of the a in D_t will remain unassigned. On this set make any assignment of a point $\bar{x}_a$.

Paint the (b,j)-name onto the tower over the set a to which it is assigned and call this partition Q. Let P_2 be the name obtained by restricting to the b component of the name and let P_1 be this name permuted by $\bar{p}_t(\bar{x}_a)$. Extend both P_1 and P_2 outside of the towers in some way, but in the same way for both. Our earlier remarks and computations guarantee (a) and (c). As Q almost catches R given the tower process and R is independent of the tower process, if δ_1 is small enough then

$$h(T_1^{\alpha_1}, Q) \geq h(T^{\alpha}, R) - \delta'/2.$$

On each set D_t each set b is cut into at most $2^{(h(\bar{T}^{\bar{\alpha}},R)-h(T_0^{\alpha_0},P_0)+4\delta_1)\#\bar{F}_j}$ sets. This implies

$$h(T_1^{\alpha_1}, Q|P_2) \leq (h(T^{\alpha}, R) - h(T_0^{\alpha_0}, P_0)) + 4\delta_1.$$

We conclude that

$$h(T_1^{\alpha_1}, P_2) \geq h(T_0^{\alpha_0}, P_0) - \delta'/2 - 4\delta_1 > h(T_0^{\alpha_0}, P_0) - \delta'$$

which is (b). □

We are now prepared to construct the ϕ in the full-group of T^{α_1}. For each set a the two names painted on the tower over a differ by the permutation $\bar{p}_t(\bar{x}_a)$. Instead of painting two names we can instead permute the orbit points to move one partition to the other. Define ϕ to permute points in the tower over a by the permutations $(\bar{p}_t(\bar{x}_a))^{-1}$ and fix all other points. Stated formally:

1. For all $\bar{x}_0 \in a$, $a \subseteq D_t$ an atom of $\vee_{g \in \bar{F}_j} T^{\alpha_1}_{1,g^{-1}}(R)\big|_{D_t}$ and all $g \in \bar{F}_j$ set
$$\phi(T^{\alpha_1}_{1,g}(\bar{x}_0)) = T^{\alpha_1}_{1,(\bar{p}_t(\bar{x}_a))^{-1}(g)}(\bar{x}_0).$$
2. For all other $\bar{x} \in X \times Z \times Z$ set $\phi(\bar{x}) = \bar{x}$.

From this definition it is an easy check that $P_2 = \phi(P_1)$ and hence the processes $(T^{\alpha_1}_1, P_2)$ and $(T^{\alpha_1 \phi}_1, P_1)$ are identical. The following lemma will now conclude the proof that (T^α, P) is reverse-weak m^e-finitely determined.

Lemma 5.7. *Both* $h(T^{\alpha_1}_1, P^{\alpha_1,\phi}) \leq \frac{8\varepsilon}{15}$ *and* $m^0(\alpha_1, \phi) < \frac{\varepsilon}{5}$ *and hence* $m^e(\alpha_1, \phi) < \varepsilon$.

Proof. First to estimate $h(T^{\alpha_1}_1, P^{\alpha_1,\phi})$ note that the $(\bar{A}_j, \bar{F}_j)$ towers and their partition into the sets D_t are all contained in the $U^\gamma \times U^\gamma$ factor and hence have entropy at most $\delta_0/5 \leq \varepsilon/5$. The permutations $\bar{p}_t(a)$ are chosen from the range of p_t, a set of cardinality bounded by $2^{\varepsilon \# \bar{F}_j/3}$. Hence the entropy of $P^{\bar{\alpha},\phi}$ conditioned on knowing the towers and the sets D_t is at most $\varepsilon/3$. This gives us the first estimate. For the second note that so long as $\bar{x}$ and $T^{\alpha_1}_{1,g}(x)$ lie in the same small cookie in a tower that
$$T^{\alpha_1}_{1,g(\phi(\bar{x}))} = \phi(T^{\alpha_1}_{1,g(x)})$$
and so
$$\begin{aligned}\mu_1(\{\bar{x} :& T^{\alpha_1}_{1,g(\phi(\bar{x}))} = \phi(T^{\alpha_1}_{1,g(\bar{x})}) \text{ for all } g \in F^0\}) \\ &\geq \sum_{A_i} \bar{\mu}(A_i)(\# \cap_{g \in F^0} gF_i) \geq \sum_{A_i} \bar{\mu}(A_i) \# F_i \left(1 - \frac{\delta_0}{10}\right) \\ &\geq \left(1 - \frac{\delta_0}{10}\right)\left(1 - \frac{\delta_0}{10}\right) > 1 - \frac{\delta_0}{5} > 1 - \frac{\varepsilon}{5}.\end{aligned}$$

□

5.1. **The infinite entropy case.** No explicit discussion is made in [3] of the case of infinite entropy. We give a very brief discussion here to fill in this gap. Infinite entropy needs treatment in two places. The first is in the discussion of m-entropy and the argument that all sizes are either entropy-free or entropy-preserving. The second is in the proof of the equivalence theorem for an entropy preserving size. Here nothing is missing but one must make note that the result still holds for infinite entropy.

First the situation of entropy-free sizes. Two facts need to be verified. One is that if a size is entropy-free when regarded on finite entropy actions, then it remains so on those of infinite entropy. The second is to see that if entropy can be lowered on some infinite entropy action, then it can be on some finite entropy action and hence is entropy free.

Note that for a countable and discrete amenable group, any action of infinite entropy will have factor actions which remain free but have finite entropy. This can be seen in a variety of ways. Perhaps the quickest is to note that in any action of positive entropy there will be Bernoulli factors of all smaller entropies. Suppose now that T^α has infinite entropy and m is entropy-free when considered on actions of finite entropy. This means that for any α_0 with $h(T^{\alpha_0}) < \infty$ we know that for a residual set of $\beta_0 \overset{m}{\sim} \alpha_0$, $h(T^{\beta_0}) = 0$. But now for any partition P, P sits in a finite entropy factor of T^α on which the action is free. Hence there must exist a $\beta \overset{m}{\sim} \alpha$ with $h(T^\beta, P) = 0$ with $m(\alpha, \beta)$ as small as we like. Hence those β for which $h(T^\beta, P) = 0$ are a residual subset of the equivalence class. Intersecting over a countable dense family of partitions, those β for which $h(T^\beta) = 0$ is also residual.

For the second issue, we state a Lemma.

Lemma 5.8. *suppose we have an infinite entropy action T^α for which there is an $\alpha_1 \overset{m}{\sim} \alpha$ with $h(T^{\alpha_1}) < \infty$. Then m must be entropy-free.*

Proof. Let $\mathcal{H}$ be some T^α invariant sub-σ-algebra with $\infty > h(T^\alpha\big|_{\mathcal{H}}) > h(T^{\alpha_1}) - e$ for some $e > 0$. We can assume the action of T^α on $\mathcal{H}$ is Bernoulli and following the notation of Lemma 5.0.6 of [3] write it as (U^β, P) where U^β is the factor action and P is an i.i.d. generator. The proof of that lemma shows how to do the following. For any $\varepsilon > 0$ one can find a ϕ that is $\mathcal{H}$ measurable so that both $m(\beta, \phi) < \varepsilon$ and $h(U^{\beta\phi}, P) < h(U^\beta, P) - e$. This now tells us, using Lemma 5.0.7 of [3], that

$$h_m(U^\beta) < h(U^\beta) - e$$

and hence m is entropy free. □

For the second area under discussion, the equivalence theorem itself, the proof of the applies to infinite entropy without any change. One must simply make note of this at one particular point. This discussion only concerns entropy-preserving sizes of course. In [3], Theorem 7.3.3 it is shown that for T^α acting on $(X, \mathcal{F}, \mu)$ an m-f.d. action and $T_1^{\alpha_1}$ acting on $(X_1, \mathcal{F}_1, \mu_1)$ some other free and ergodic action with $h_m(T_1^{\alpha_1}) \geq h_m(T^\alpha)$, then certain open sets $\mathcal{O}(I, P, \varepsilon)$ are dense in the Polish space of m-joinings of these two actions. In this notation, $I \in \mathbb{N}$, P is a finite partition and $\varepsilon > 0$. In this result either or both of the actions can have infinite entropy and the proof will hold without change. The intersection of these sets over I, a countable dense collection of partitions and for $\varepsilon = 1/n$ consists of precisely those joinings where $T_1^{\alpha_1}$ can be taken as the covering space. That is to say both $\mathcal{F}$ and the m-Cauchy sequence of full-group elements giving the m-equivalence are a.s. $\mathcal{F}_1$-measurable. Hence the collection of such m-joinings is residual in this Polish space. If both actions were m-f.d. and of infinite entropy then, as is standard, intersecting the two residual sets gives the equivalence theorem.

6. CONCLUSIONS

We conclude with a discussion of three issues deserving of further study. To begin, one piece missing in the development presented here is an easily checked criterion for when two G-arrangements α and β on the same orbit space are m^e-equivalent. How would one effectively look for an m^e-Cauchy sequence of full-group elements leading from one to the other? Is there some evaluation that can be made directly in terms of the pair, for example some calculation made on the function $h_x^{\alpha,\beta}(g) = \beta(x, T_g^\alpha(x))$ that would determine whether the arrangements are m^e-equivalent? There are numerous ways one might evaluate the "complexity" of this family of bijections of G. For example, take a Følner sequence F_n. For each n the possible values of $h_x^{\alpha,\beta}\big|_{F_n}$ give a countable partition of X. One can calculate its entropy and take the $\liminf$ as $n \to \infty$. As another possibility, rather than use the full function $h_x^{\alpha,\beta}\big|_{F_n}$ one could partition by just the value of $\mathrm{Range}(h_x^{\alpha,\beta}\big|_{F_n})$, and once more calculate its entropy and take the $\liminf$. Perhaps one should take a conditional entropy of one level $h_x^{\alpha,\beta}\big|_{F_N}$ given another $h_x^{\alpha,\beta}\big|_{F_n}$ then let N go to infinity and then n. This latter would give a more properly asymptotic notion of the complexity. The question is though what natural calculation of this form can be directly associated to the m^e-equivalence of α and β.

A second issue that would be extremely interesting to investigate is to attempt to make the development here group independant. Is there a Connes, Feldman, Weiss [1] version of the result we have given here? As currently configured the restricted orbit equivalence theory is built on the full-group, which seemingly cannot be used to change the group which acts. If one can find a good answer to the previous issue, finding a good characterization of m^e equivalence of arrangements that does not depend on the restricted orbit equivalence theory then perhaps one can find such a characterization for pairs of arrangements that are not necessarilly of the same group. This would give a definition of entropy then over the whole class of discrete amenable group actions in terms of the complexity necessary to change one action into the other.

As a last possible direction and perhaps the most conjectural, can one use some notion of complexity of the change in orbit structure to give a robust and natural theory of entropy for non-singular and recurrent actions. Once more one has moved beyond the applicability of the restricted orbit equivalence machinery as it is not cast. In order to extend that machinery for the entropy-preserving case one needs an entropy so to hope to build the r.o.e. theory first would be putting the cart before the horse. Perhaps though one can focus on the size m^e precisely and attempt to give a non-singular version of the theory for just such a size.

REFERENCES

1. J. Feldman A. Connes and B. Weiss, *An amenable equivalence relation is generated by a single transformation*, Erg. Th. and Dyn. Sys. **1** (1981), 431–450.
2. H. A. Dye, *On groups of measure preserving transformations i*, Amer. J. Math. **81** (1959), 119–159.
3. J. W. Kammeyer and D. J. Rudolph, *Restricted orbit equivalence for actions of discrete amenable groups*, Cambridge University Press, 2001.
4. D. Ornstein and B. Weiss, *Entropy and isomorphism theorems for actions of amenable groups*, J. Analyse Math. **48** (1987), 1–141.
5. V.S. Varadarajan, *Measures on topological spaces*, Amer. Math. Soc. Transl., Ser. 2 **48** (1965), 161–228.