ISRAEL JOURNAL OF MATHEMATICS **137** (2003), 35–60

ENTROPY AND σ-ALGEBRA EQUIVALENCE OF CERTAIN RANDOM WALKS ON RANDOM SCENERIES

BY

KAREN BALL*

Department of Mathematics, University of Maryland, College Park
College Park, MD 20742, USA
e-mail: kareball@indiana.edu

ABSTRACT

Let $X = X_0, X_1, \ldots$ be a stationary sequence of random variables defining a sequence space Σ with shift map σ and let (T_t, Ω) be an ergodic flow. Then the endomorphism $\hat{T}_X(x, \omega) = (\sigma(x), T_{x_0}(\omega))$ is known as a random walk on a random scenery. In [4], Heicklen, Hoffman and Rudolph proved that within the class of random walks on random sceneries where X is an i.i.d. sequence of Bernoulli-$(\frac{1}{2}, \frac{1}{2})$ random variables, the entropy of T_t is an isomorphism invariant. This paper extends this result to a more general class of random walks, which proves the existence of an uncountable family of smooth maps on a single manifold, no two of which are measurably isomorphic.

1. Introduction

Let T_t be the geodesic flow on the unit tangent space to a compact surface of negative curvature, T^1M. For $\alpha > 0$, define $\widehat{T}(\alpha)$: $[0, 1] \times T^1M \to [0, 1] \times T^1M$ by

$$\widehat{T}(\alpha)(x, \omega) = (2x \bmod 1, T_{\alpha \sin(2\pi x)}(\omega)).$$

If ν is the Liouville measure on T^1M, T_t is known to be ergodic with positive entropy. When λ is Lebesgue measure on [0,1], each $\widehat{T}(\alpha)$ is ergodic with respect to the product measure, $\lambda \times \nu$. It is not difficult to see that the entropy of

---

* This research was sustained in part by fellowship support from the National Physical Science Consortium and the National Security Agency.
Received April 8, 2002 and in revised form August 4, 2002

$\widehat{T}(\alpha)$ is $\log 2$ for all α. However, in this paper we prove that if $\alpha \neq \beta$, then $\widehat{T}(\alpha)$ and $\widehat{T}(\beta)$ are not isomorphic as measurable dynamical systems. Thus, $\{\widehat{T}(\alpha)\}_{\alpha>0}$ is an uncountable family of smooth dynamical systems on the same four-dimensional manifold, all with the same entropy but with no two isomorphic. No such collection of invertible maps is known.

This example motivates the more abstract result proved in this paper. The maps $\widehat{T}(\alpha)$ are examples of random walks on random sceneries, which we define as follows. Let $X = X_0, X_1, \ldots$ or $X = \ldots, X_{-1}, X_0, X_1, \ldots$ be a stationary sequence of real-valued random variables, which define a sequence space Σ with measure m and shift map σ. Then if T_t is any measure-preserving, invertible flow on a space $(\Omega, \mathcal{F}, \nu)$, we can define a map $\hat{T}_X$ on $(\Sigma \times \Omega, \mu = m \times \nu)$ by

$$\hat{T}_X(x, \omega) = (\sigma(x), T_{x_o}(\omega)).$$

In the special case where X is a one-sided sequence of independent random variables with Bernoulli-$(\frac{1}{2}, \frac{1}{2})$ distribution taking values ± 1 and $T = T_1$ is the time-one map for T_t, such random walks on random sceneries are known as T, T^{-1} endomorphisms. T, T^{-1} endomorphisms and their invertible counterparts have been studied in various papers, such as [5], [3], [4].

Random walks on random sceneries are so named because the second coordinate of $\hat{T}_X^n$ is $T_{S_n(x)}(\omega)$ where $S_n = \sum_0^{n-1} X_i$. The first coordinate gives the random walk and the second is the scenery. In this notation, our maps $\widehat{T}(\alpha)$ are trivially isomorphic to $(\widehat{T_\alpha})_X$ where the random walk X is given by $X_n = \sin(2\pi S^n(x))$ where $S(x) = 2x \bmod 1$ is the expanding map on the unit interval and the random scenery is $(T_\alpha)_t = T_{\alpha \cdot t}$, the geodesic flow on $T^1 M$ at speed α.

In investigating when non-invertible random walks on random sceneries cannot be isomorphic, it turns out to be useful to consider a weaker form of isomorphism, which we call σ-algebra isomorphism. For any endomorphism S on a probability space $(Y, \mathcal{H}, \mu)$, we can define a decreasing sequence of σ-algebras $\mathcal{H}_i = \sigma(S^i, S^{i+1}, \ldots)$. We say that two such endomorphisms S_1 and S_2 are σ-algebra isomorphic if there is a 1-1, measure-preserving map from Y_1 to Y_2 which takes $\mathcal{H}_i(S_1)$ to $\mathcal{H}_i(S_2)$.

Note that any measurable isomorphism gives a σ-algebra isomorphism, so that σ-algebra isomorphism is a weaker notion than isomorphism of dynamical systems. If S is invertible, then $\mathcal{H}_i |\, \mathcal{H}_{i+1}$ is always a point mass, so any two invertible dynamical systems are σ-algebra isomorphic. However, the situation is much richer for non-invertible maps. Heicklen, Hoffman and Rudolph prove in [4] that if two endomorphisms T, T^{-1} and S, S^{-1} are σ-algebra isomorphic, then the entropy of T and S must be the same. This immediately implies that the entropy of T is an isomorphism invariant within the class of T, T^{-1} endomorphisms.

In this paper, we show that the basic idea of the proof in [4] can be adapted to show that the entropy of the scenery process is an invariant of σ-algebra isomorphism within a more general class of random walks on random sceneries, which includes the transformations $\widehat{T}(\alpha)$. As mentioned above, the scenery of $\widehat{T}(\alpha)$ is given by T_α, so the scenery entropy of $\widehat{T}(\alpha)$ is $\alpha \cdot h(T_1)$. Since $h(T_\alpha) \neq h(T_\beta)$ when $\alpha \neq \beta$, we conclude that $\widehat{T}(\alpha)$ and $\widehat{T}(\beta)$ are not isomorphic, or even σ-algebra isomorphic.

The more general class of random walks on random sceneries we consider is as follows. Let f: $[0,1] \to \mathbb{R}$ be a Hölder continuous function with $\int f = 0$ and f not a coboundary. For example, in the maps $\widehat{T}(\alpha)$ we take $f(x) = \sin(2\pi x)$. Let S be the expanding map on the unit interval defined by $S(x) = 2x \bmod 1$ (so S is isomorphic to the one-sided Bernoulli 2-shift). Define a sequence of random variables $X_f = X_0, X_1, \ldots$ where $X_n = f \circ S^n$. Then X_f is a weakly dependent sequence of random variables taking values in $\mathbb{R}$. The scenery flow can be any invertible, ergodic flow T_t. We will use the notation $\hat{T}_f$ for $\hat{T}_{X_f}$ and will consider it as a map $\hat{T}_f$: $[0,1] \times \Omega \to [0,1] \times \Omega$ given by $\hat{T}_f(x, \omega) = (Sx, T_{f(x)}(\omega))$. Clearly, the maps $\widehat{T}(\alpha)$ are of this type.

For any $a > 0$, let $f' = a^{-1} f$ and $S_t = T_{a \cdot t}$ and use these to define $\hat{S}_{f'}$. Then $\hat{T}_f$ and $\hat{S}_{f'}$ are isomorphic via the isomorphism $(x, \omega) \mapsto (a^{-1}x, \omega)$. Thus for any a, $\hat{T}_f$ is isomorphic to a random walk on a random scenery, $\hat{S}_{f'}$, with $h(S) = a$. It is therefore necessary to normalize f to eliminate these trivial isomorphisms. Let $S_n = X_0 + \cdots + X_{n-1}$ and let σ_n be the standard deviation of S_n. Because f is chosen not to be a coboundary, it follows from the work of Rio in [8] that $0 < \lim_{n\to\infty} \sigma_n/\sqrt{n} = a < \infty$. We choose to normalize so that $a = 1$.

The main result of this paper then states that if f and g are functions as above which are appropriately normalized and if T_t and S_t are invertible, measure-preserving transformations such that $\hat{T}_f$ and $\hat{S}_g$ are σ-algebra isomorphic, then $h(T) = h(S)$. As an immediate corollary, we see that if $h(T) \neq h(S)$, then $\hat{T}_f$ and $\hat{S}_g$ are not measurably isomorphic dynamical systems.

2. Outline of the proof

The proofs of both the Heicklen, Hoffman and Rudolph result and the one in this paper depend on the tree structure associated with the nth inverse images of any point under the endomorphism. T, T^{-1} endomorphisms are similar to those considered here in that the maps are all uniformly two-to-one: $\mathcal{H}_i | \mathcal{H}_{i+1}$ has two point fibers of equal mass. Therefore for any $\hat{T}_f$, the tree $\mathcal{T}_{n,(x,\omega)}$ of nth inverse images of a point (x,ω) is a binary n-tree and the conditional measure on the tree, $\mu_{n,(x,\omega)} = \mu(\bullet|\, \hat{T}_f^n = (x,\omega))$, puts equal mass on each of the 2^n branches. The key feature of a σ-algebra isomorphism is that it maps nth trees of inverses to nth trees of inverses.

We define a metric on the domain of $\hat{T}_f$, $[0,1]\times\Omega$, making use of the tree structure. If P is a finite partition of $[0,1]\times\Omega$, then we can define the nth mean Hamming distance between (x,ω) and (x',ω') by

$$d_n^P((x,\omega),(x',\omega')) = \mu_{n,(x,\omega)}(\{\text{branches } b \text{ s.t. the } P\text{-labels of } b \text{ in } \mathcal{T}_{n,(x,\omega)} \text{ and } \mathcal{T}_{n,(x',\omega')} \text{ disagree}\}).$$

The Hamming distance allows us to define the Vershik distance between two points (x,ω) and (x',ω'), which is easily seen to be a pseudometric.

Definition 2.1: Let $\mathcal{A}_n$ be the set of all tree automorphisms of the binary n-tree. Then the Vershik distance v_n^P on $[0,1]\times\Omega$ is defined to be

$$v_n^P((x,\omega),(x',\omega')) = \inf_{a\in\mathcal{A}_n} d_n^P(a(\mathcal{T}_{n,(x,\omega)}),\mathcal{T}_{n,(x',\omega')}).$$

We sometimes write v_n for v_n^P when the partition P is understood.

For technical reasons, we choose to control separately the degree to which the random walk and scenery coordinates of points are matched when calculating the v_n distance.

Definition 2.2: Let Q be a finite partition of Ω. Two points (x,ω) and (x',ω') are said to be (δ,m)-close in v_n^Q if $v_n^Q((x,\omega),(x',\omega')) < \delta$ and there is a tree automorphism, a, with $d_n^P(a(\mathcal{T}_{n,(x,\omega)}),\mathcal{T}_{n,(x',\omega')}) < \delta$ and which acts as the identity up to height m.

We will also make use of the $\bar{d}$ metric and Feldman's $\bar{f}$ metric [2]. For any $w,w'\in\{0,\dots,l\}^{\mathbb{Z}}$ and $m<n\in\mathbb{Z}$,

$$\bar{d}_{[m,n]}(w,w') = 1-\frac{k}{n-m+1}$$

where k is the number of indices $i, m\le i\le n$, such that $w(i)=w'(i)$. Also,

$$\bar{f}_{[m,n]}(w,w') = 1-\frac{k}{n-m+1}$$

where k this time is the maximal integer for which there are subsequences of integers $m\le i_1<i_2<\cdots<i_k\le n$ and $m\le j_1<j_2<\cdots<j_k\le n$ such

that $w(i_s) = w(j_s)$ for all $s = 1, \dots k$. Given a dynamical system (T_t, Ω) and a finite partition Q on Ω, $\bar{d}^Q_{[m,n]}(\omega, \omega')$ is the $\bar{d}$ distance between $(Q(T_i\omega))_{i\in\mathbb{Z}}$ and $(Q(T_i\omega'))_{i\in\mathbb{Z}}$ and similarly for $\bar{f}$.

To prove the main result, we first prove that for certain partitions Q of Ω and up to sets of small measure, the projections of most small v_n^Q neighborhoods in $[0,1] \times \Omega$ onto the second coordinate are contained in small $\bar{f}$ neighborhoods in Ω. This is done by looking carefully at the properties of the tree automorphisms a which achieve small v_n^Q distances.

Now suppose that $(\hat{S}_g, [0,1] \times \Omega_S)$ and $(\hat{T}_f, [0,1] \times \Omega_T)$ are σ-algebra isomorphic. By a finite code argument, we can use the σ-algebra isomorphism and the relation between v_n and $\bar{f}$ to map cylinder sets of a partition P on Ω_S to $\bar{f}$-neighborhoods for Q in Ω_T. Because the exponential size of both cylinder sets and $\bar{f}$-neighborhoods are given by the entropy of the transformation, this is sufficient to show that T_1 and S_1 must have the same entropy.

The central technical result used in the proof of the main theorem of the paper is thus the one which relates the v_n-distance between two points to the $\bar{f}$-distance between the Ω coordinates of the points. The following statement gives the idea of this theorem and shows the roles of the various parameters we will use in the proof. The actual theorem we will prove is Theorem 6.7.

IDEA FOR RELATION BETWEEN v_n AND $\bar{f}$ METRICS: *Let $(\hat{T}_f, [0,1] \times \Omega)$ be a random walk on a random scenery as above. Then there is a finite partition Q of Ω, such that for any large c there is a good set G with $\mu(G) > 1 - 1/c$, a choice of $\delta > 0$ and n_0 such that the following holds. If n is chosen sufficiently large and (x, ω), (x', ω') are in G and are $(\delta, 2n_0)$-close in v_n^Q, then*

$$\bar{f}^Q_{[-c\sigma_n, c\sigma_n]}(\omega, \omega') < 1/c.$$

In particular, this will imply that the number of v_n neighborhoods for this partition grows at an exponential rate in $\sigma_n \sim \sqrt{n}$. We use as a starting point for the proof Theorem 3.2 below, proved in [1], which says that the number of such v_n neighborhoods is superpolynomial in n.

The basic structure of the proof is very similar to that used by Heicklen, Hoffman and Rudolph to prove the analogous result for T, T^{-1} maps. The main additional difficulties that arise come from the lack of uniformity in the n-trees, $\mathcal{T}_{n,(x,\omega)}$. The steps taken by the random walk are no longer independent, though the dependence is weak over the course of many steps. Also, the fact that the scenery space now consists of continuous time flows requires some new constructions, which are given in Section 3. Section 4 describes the set of large measure on which the relationship between v_n and $\bar{f}$ holds. The relationship itself is established in Sections 5 and 6. Finally, the finite coding argument is given in Section 7.

3. Preliminaries

We first establish some terminology. Define an **n-tree** to be the binary tree of height n, which we can associate with $\mathcal{T}_{n,(x,\omega)}$ for each point (x,ω). An **n-branch** b is an element of $\{0,1\}^n$. In the context of $\mathcal{T}_{n,(x,\omega)}$, we can also think of b as the point (y,ψ) with $\hat{T}_f^n(y,\psi) = (x,\omega)$ and $(y_0,\ldots,y_{n-1}) = b$ where $(y_0,y_1,\ldots)$ is the binary expansion of y. The branch $b = (y,\psi)$ passes through r at height m if $\sum_{i=m+1}^{n}(f \circ S^i(y)) = r$. The **distance between two n-branches** $b = (y,\psi)$ and $b' = (y',\psi')$ in $\mathcal{T}_{n,(x,\omega)}$ at height m is $\sum_{i=m+1}^{n}(f \circ S^i(y) - f \circ S^i(y'))$. We say that $\mathcal{T}_{m,(y,\psi)}$ is an **m-subtree** in the n-tree $\mathcal{T}_{n,(x,\omega)}$ if $\hat{T}_f^{n-m}(y,\psi) = (x,\omega)$. The **distance between two m-subtrees** in an n-tree is the distance at height m between any representative branches from the subtrees. The **c-middle** of an n-tree is the interval $[-c\sigma_n, c\sigma_n]$ at height 0 along the base of the tree.

The paths traced out by the trees of inverse images of points are not as rigid for $\hat{T}_f$ as they are for T,T^{-1} endomorphisms. If we define the **shape of a tree** $\mathcal{T}_{n,(x,\omega)}$ to be the 2^n n-step paths traced out as we walk along each branch of the tree, then the T,T^{-1} map has only one tree shape and the distribution of the branches along the base of the tree is binomial. By contrast, since we are no longer constrained to take steps only of length ± 1, the shapes of any two n-trees of inverse images for $\hat{T}_f$ may differ. The requirement that f be Hölder continuous does assure us that there is a constant B such that the shapes of any two infinite trees of inverses for $\hat{T}_f$ can differ by at most B. For a proof of this simple fact, see [1].

The following result gives an estimate for the distribution of branches at the base of any n-tree over a point in $[0,1] \times \Omega$. For an interval I, we define $\mu_{n,(x,\omega)}(I) := \mu_{n,(x,\omega)}(-S_n \in I)$.

THEOREM 3.1 ([1]): *If f satisfies a Hölder condition of order $\alpha > 0$ and if $\int f = 0$, then there exists a constant A such that for each $(x,\omega) \in [0,1] \times \Omega$,*

$$\sup_{y \in \mathbb{R}} |\mu_{n,(x,\omega)}((-\infty, \sigma_n y]) - \gamma(-\infty, y]| \le A n^{-1/2}$$

where γ is the standard Gaussian measure.

We now specify the class of partitions for which we will prove the relation between the v_n and $\bar{f}$ metrics. Though the scenery processes we are considering are continuous in time, we discretize to some extent by taking our partition Q to be defined by a flow under a function. Problems may occur in our argument wherever the T_t-orbit of a point intersects the base in this construction. Such points are where transitions between different partition elements occur and will be called cut points for the partition, or Q-cuts.

In order to manage the cut points, we consider a partition in which these transitions are widely spaced compared to the variation we see in the trees. The following constants depend on f:

- A is as in Theorem 3.1. It gives the scale on which the measures $\mu_{n,(x,\omega)}$ are approximated by the binomial distribution.
- B is the constant bounding the differences in the shapes of any two infinite trees of inverses for $\hat{T}_f$.
- $s = ||f||_\infty$.

We use these numbers to determine the spacing of our partition Q:

- Let $d > 2B + 2s + 3Ae^8 + 1$. This d will be the smallest scale on which we consider the distribution of branches of an n-tree.
- δ_0 will be chosen as specified by Theorem 3.2 and Lemma 4.2.
- a_1 and a_2 are then chosen so that $a_1 > a_2 > 64d/\delta_0$, a_1/a_2 is irrational, and $a_1/a_2 < 3/2$. Then a_1 and a_2 will be the values of the function in our flow under a function construction.

We construct a positive-entropy Bernoulli flow as a flow under a function over an i.i.d. base. Let $(\bar{T}, \bar{Q}, \bar{\mu})$ be an i.i.d. process on sequences $\{\bar{q_1}, \bar{q_2}\}^{\mathbb{Z}}$ with $\bar{\mu}(\bar{q_1}) = \bar{\mu}(\bar{q_2}) = 1/2$. Take h to be the function with height a_1 over $\bar{q_1}$ and height a_2 over $\bar{q_2}$. Let S_t be the flow under h and choose a_1, a_2 large enough so that the entropy of S_t is less than the entropy of T_t. Then S_t is a factor of T_t by Ornstein's isomorphism theory for flows [6]. Let $Q = Q(a_1, a_2)$ be a full-height refinement of the partition $\{q_1, q_2\}$ of Ω, defined as the partition into sets lying under h and over $\{\bar{q_1}, \bar{q_2}\}$. If $h(T) < \infty$, we may choose Q so that it achieves the full entropy of T_t.

We can now apply a theorem from [1] to the partition Q defined above to show that the number of v_n^Q-names is growing superpolynomially for a_1, a_2 sufficiently large, even when we condition on the location of a certain Q-cut.

THEOREM 3.2 ([1]): *Let $\hat{T}_f$ be a random walk on a random scenery such that f is Hölder continuous and $h(T_t) > 0$. There exists $\delta_0 > 0$ and $a = a(X)$ such that for any partition $Q = Q(a_1, a_2)$ with $a_1/a_2 < 3/2$ irrational and $a_1 > a_2 > a$, there exists m_0 such that the following property holds. For p any polynomial with positive leading coefficient, n sufficiently large, $0 < u \le a_1$, and any $\omega \in \Omega$,* *if*

$$\Theta(n,\epsilon,\omega)=\{\omega'|\ \exists x,x'\in[0,1] \textit{ and } s\in[-1,1] \textit{ such that}$$
$$(x,\omega),(x',T_s(\omega')) \textit{ are } (\delta_0,m_0)\textit{-close in } v_n^Q\},$$
$$\gamma_{n,u}=\{\omega'|\ u=\textit{smallest } t>0 \textit{ with } T_{-(sn+t)}(\omega') \textit{ centered over a } Q\textit{-cut}\},$$

then $\nu(\Theta(n,\epsilon,\omega)|\ \gamma_{n,u})<1/p(n)$.

4. Nondegeneracy

We wish to establish properties of the tree automorphisms that generate small v_n distances between points in our space. In particular, we wish to establish a relationship between the v_n distance between two points (x,ω) and (x',ω') and the $\bar{f}$ distance between the Q-names of ω and ω'.

Following Heicklen, Hoffman, and Rudolph, we will first define a notion of degeneracy of points in $[0,1]\times\Omega$ and will restrict our attention to nondegenerate points, which will form the good set G (see the **Idea for relation between** v_n **and** $\bar{f}$ in Section 2). Our notion of nondegeneracy will depend on the four parameters, δ_0,δ,n_0, and n:

1. δ_0 is chosen small as in Theorem 3.2 above and Lemma 4.2 below. The partition Q of Ω depends only on δ_0 and X_f.
2. $\delta\ll\delta_0$ is chosen next. δ is determined by the degree of closeness we wish to achieve in the $\bar{f}$-matching, and is given as a function of c.
3. n_0 is chosen large, depending on δ_0 and δ. All tree automorphisms considered in our argument will act as the identity at least up to height n_0.
4. n depends on all of the above. This number is the full height of a tree.

The roles of these parameters can be seen in the **Idea for relation between** v_n **and** $\bar{f}$ in Section 2.

Properties 1–4 below together define the set G of nondegenerate points. G will be the set of points (x,ω) for which most branches in $\mathcal{T}_{n,(x,\omega)}$ satisfy all of these properties. Properties 1 and 2 correspond to the notions of nondegeneracy defined in [4]. The last two properties are necessitated by the nonuniformity of the tree structures for the skew products we are using. Let (x,ω) lie in $[0,1]\times\Omega$.

Property 1: No small shift of ω is $\bar{d}$-close to ω. A trivial example of what we want to avoid is the Q-word consisting of all q_1. Formally, we define

$$B(n_0,\delta_0)=\{\omega|\ \exists t, a_1<|t|<3n_0 \text{ with } \bar{d}^Q_{[-\sigma_{n_0},\sigma_{n_0}]}(\omega,T_t(\omega))<\delta_0/4\},$$
$$\overline{B}(n_0,\delta_0)=\bigcup_{t\in[-sn_0,sn_0]} T_t(B(n_0,\delta,\delta_0)),$$

and require that $\omega \notin \overline{B}(n_0, \delta_0)$.

Property 2: No mid-sized shift of ω is v_k-close to ω for any $k > n_0$. If the shifted tree overlaps with the original tree, a good match is likely. If we are allowed to shift as far as we want, we will eventually get a good match because the flow is ergodic. Between these two extremes, we get our definition. Let

$$\begin{aligned}\gamma_k &= \{\omega \mid \exists t, 2sk < |t| < k^5, \text{ and } x, x' \in [0,1] \text{ such that} \\ &\qquad (x,\omega), (x', T_t(\omega)) \text{ are } (\delta_0, n_0)\text{-close in } v_k^Q\}, \\ \Gamma_k &= \bigcup_{t\in[-sk,sk]} T_t(\gamma_k),\end{aligned}$$

and take $\omega \notin \bigcup_{k=n_0}^{\infty} \Gamma_k$.

Property 3: Because of the non-uniformity in the trees, problems arise when a branch lands too close to a Q-cut. Such situations will certainly arise, but they should happen nearly as infrequently as possible and be spread rather evenly through the tree. To this end, we take (x,ω) so that the fraction of branches in the n_0-tree over $\hat{T}_f^{n_0}(x,\omega)$ which land within $2d$ of a cut is at most $\delta_0/4$.

Property 4: The final property is called n_0-regularity. A point (x,ω) has the n_0-regularity property if for each $h \geq n_0$, then for more than half of the $h' \in (h^{1/4}, h^{1/2})$ we have $|S^f_{h'}(x) = X_0(x) + \cdots + X_{h'-1}(x)| < 102\sqrt{h'}$. In particular, we will use the fact that for any two n_0-regular points, there is an h' with both $|S^f_{h'}(x)|, |S^f_{h'}(x')| < 102\sqrt{h'}$.

Definition 4.1: A point $(x,\omega) \in [0,1]\times\Omega$ is $(n, n_0, \delta, \delta_0)$-nondegenerate if the fraction of branches in $\mathcal{T}_{n,(x,\omega)}$ which satisfy Properties 1–4 is at least $1-\delta$. The set of these good nondegenerate points is called G.

First we need to show that once we have chosen δ_0, δ, n_0 appropriately, the set of $(n, n_0, \delta, \delta_0)$-nondegenerate points is large for large n. The following lemmas show that the set of points which do not have Properties 1–4 has small measure.

LEMMA 4.2: *There exists δ_0 such that for any Q constructed as in Section 3 and any δ, if n_0 is taken sufficiently large then $\nu(\overline{B}(n_0,\delta_0)) < \delta^2/4$.*

Proof: Let $\omega \in \Omega$. For any $t \in [-1,1]$,

$$\bar{d}^Q_{[-\sigma_{n_0},\sigma_{n_0}]}(\omega, T_t(\omega)) \leq 1/a_2 \leq \delta_0/64d < \delta_0/4.$$

Thus, if there exists t' such that $\bar{d}^Q_{[-\sigma_{n_0},\sigma_{n_0}]}(\omega, T_{t'}(\omega)) < \delta_0/4$, then for an integer $j \in [t'-1, t'+1]$, ω and $T_j(\omega)$ are within $\delta_0/2$ in $\bar{d}$. We may therefore restrict our attention to shifts by integer amounts.

We have partitioned Ω so that (T_t, Q) is a refinement of a flow under a function over an i.i.d. base. We wish to use the independence of the base process $(\bar{T}, \bar{Q})$ to estimate probabilities. In order to do so effectively, we need to understand how the probabilities behave when we condition on full past algebra, which includes positioning information as well as the symbolic information of the independent base process. Let $\mathcal{P}$ be the past σ-algebra for the flow T_t generated by the partition $\{q_1, q_2\}$. It is not difficult to show that $1/2 \le \nu(T_{t'}(q_1)|\,\mathcal{P})(\omega) \le 3/4$ for all ω when $t' > a_1$.

Fix an integer $a_1 < |j| \le 3n_0$. We now have the tools to estimate the probability that ω and $T_j(\omega)$ are within $\delta_0/2$ in $\bar{d}$. Suppose they are. Set $\bar{a} = \lceil a_1 \rceil$. We divide the integers in the interval $[-\sigma_{n_0}, \sigma_{n_0}]$ into sets $J_0, \ldots, J_{\bar{a}-1}$ according to the value of the integer modulo $\bar{a}$. The $\bar{d}$ distance between ω and $T_j(\omega)$ is the average of the quantities $\bar{d}_{J_k}(\omega, T_j(\omega))$. There must then exist a k such that $\bar{d}_{J_k}(\omega, T_j(\omega)) < \delta_0/2$. This implies that at least a fraction $1 - \delta_0/2$ of the q_1, q_2-coordinates in J_k must match with their images when shifted by j. Since consecutive coordinates are at least $\bar{a}$ apart and the shift is by $j > a_1$, the dependence of any matching probability is governed by the estimates on the conditional measures for the past algebra $\mathcal{P}$ and we get the following estimate:

$$\nu(\{\omega|\ \bar{d}_{J_k}(\omega, T_j(\omega)) < \delta_0/2\}) \le \binom{2\bar{a}^{-1}\sigma_{n_0}}{(2\bar{a}^{-1}\sigma_{n_0})\delta_0/2}\left(\frac{3}{4}\right)^{(1-\delta_0/2)2\bar{a}^{-1}\sigma_{n_0}}.$$

We use Stirling's formula to approximate the binomial, sum over j and k, and shift the resulting sets by integer amounts up to sn_0 to get

$$\nu(\overline{B}(n_0, \delta_0)) \le 6\bar{a}sn_0^2 2^{h(\delta_0/2)\cdot 2\bar{a}^{-1}\sigma_{n_0}}\left(\frac{3}{4}\right)^{(1-\delta_0/2)2\bar{a}^{-1}\sigma_{n_0}}$$

where $h(x) = -x\log x - (1-x)\log(1-x)$. If δ_0 is chosen sufficiently small, then $2^{h(\delta_0/2)}(3/4)^{1-\delta_0/2} < 1$ and then, once δ is fixed, we can choose n_0 large enough so that $\nu(\overline{B}(n_0, \delta_0)) < \delta^2/4$. ∎

LEMMA 4.3: *For any n_0 sufficiently large, $\nu(\bigcup_{k=n_0}^{\infty}\Gamma_k) < \delta^2/4$.*

Proof: Choose δ_0 as in Theorem 3.2 and apply the theorem with $p(k) = k^{11}$. For any $\omega \in \Omega$ and any t with $|t| > 2sk$, if x and x' are any two points in $[0,1]$, then the k-trees over (x, ω) and $(x', T_t(\omega))$ do not overlap. Therefore, for such t it follows from Theorem 3.2 that if we set

$$G_{t,k} = \{\omega|\ \exists x, x' \in [0,1] \text{ such that } \exists t' \in [-1,1] \text{ such that } (x,\omega), (x', T_{t+t'}(\omega)) \text{ are } (\delta_0, m_0)\text{-close}\},$$

then $\nu(G_{t,k}) < 1/k^{11}$ for $k > n_0$ sufficiently large. If $n_0 > m_0$, $T_{t'}(\gamma_k) \subseteq \bigcup_{|i|=2sk}^{k^5} G_{i,k}$ whenever $t' \in [-1,1]$. Thus, $\nu(\Gamma_k) \leq 2sk \cdot k^{-6}$ when $k \geq n_0$. If we choose n_0 large enough, we get

$$\nu\left(\bigcup_{k=n_0}^{\infty} \Gamma_k\right) \leq 2s \sum_{k=n_0}^{\infty} k^{-5} < \delta^2/4. \qquad \blacksquare$$

LEMMA 4.4: *If n_0 is chosen to be sufficiently large, then the measure of the set of points not satisfying Property 3 is less than $\delta^2/4$.*

Proof: The measure of points in $[0,1] \times \Omega$ which lie within $2d$ of a Q-cut is $8d/(a_1 + a_2) < \delta_0/4$. The fraction of branches in $\mathcal{T}_{n_0,(x,\omega)}$ which land within $2d$ of a cut is a reverse martingale, so it converges almost surely to $8d/(a_1 + a_2)$ as $n_0 \to \infty$. We choose n_0 large enough so that the measure of the set of points (x, ω) with more than $\delta_0/4$ of the branches of $\mathcal{T}_{n_0,(x,\omega)}$ landing within $2d$ of a cut is less than $\delta^2/4$. Since $\hat{T}_f$ is measure preserving, $\hat{T}_f^{-n_0}$ of this set has the same measure. $\blacksquare$

LEMMA 4.5: *If n_0 is chosen to be sufficiently large, then the measure of the set of points without the n_0-regularity property (Property 4) is less than $\delta^2/4$.*

Proof: Heicklen, Hoffman and Rudolph proved in [4] that if Y_n is a sequence of i.i.d. Bernoulli $(1/2, 1/2)$ with $S_n = \sum_0^{n-1} Y_i$, then there is a subset R_{n_0} of the sequence space Σ with measure less than $4/\log n_0$ such that for any $\chi \in R_{n_0}$ and any $h \geq n_0$, more than half of the $h' \in (h^{1/4}, h^{1/2})$ satisfy $|S_{h'}(\chi)| < 101\sqrt{h'}$. By Philipp and Stout's invariance principle on page 80 of [7], we may redefine the process $S_n^f = \sum_0^{n-1} (X_f)_i$ on a richer probability space together with S_n in such a way that

$$\frac{S_n^f - S_n}{\sqrt{n}} \to 0 \text{ almost surely as } n \to \infty.$$

The conditions of the invariance principle in the case of S^f are checked by Rudolph in [9]. Take n_0 large enough so that $4/\log n_0 < \delta^2/8$ and the

$$P(|S_n^f - S_n| < \sqrt{n}, \forall n \geq n_0^{1/4})$$

is at least $\delta^2/8$. Then we can use R_{n_0} to define a set $R'_{n_0} \subseteq [0,1]$ of points with the n_0-regularity property, such that the measure of R'_{n_0} is less than $\delta^2/4$. $\blacksquare$

Now that we know that it is rare for a point not to have all of the four non-degeneracy properties, it is easy to prove that most points are in the good set G of $(n, n_0, \delta, \delta_0)$-nondegenerate points.

LEMMA 4.6: *For any delta sufficiently small and $n_0 = n_0(\delta)$ large enough with fixed $n > n_0$, the set of all $(x, \omega) \in [0,1] \times \Omega$ which are $(n, n_0, \delta, \delta_0)$-degenerate has measure less than δ.*

Proof: By the previous four lemmas, the probability that a point does not have Properties 1–4 is at most δ^2. By Chebyshev's inequality, the probability that a tree will contain more than a fraction δ of such branches is at most δ. ∎

5. Tree automorphisms

Take (x, ω) and (x', ω') in the nondegenerate set G, which are $(\delta, 2n_0)$-close in v_n^Q. Let a be a tree automorphism realizing this distance. The goal of this section is to prove that if we restrict our attention to a large set of branches which behave well for a, then any two branches passing close together in $\mathcal{T}_{n,(x,\omega)}$ must have images under a which pass close together in $\mathcal{T}_{n,(x',\omega')}$. This mapping on the branches will form the basis for the $\bar{f}$-matching between ω and ω'.

The first lemma is a technical result which allows us to manipulate with subtrees.

LEMMA 5.1 (Matching Lemma): *Let t_1, t_2 be two h-subtrees of an n-tree $\mathcal{T}$ and take $C > 0$. Suppose that the distance between t_1 and t_2 at height h is $C\sigma_h$. Further, suppose we have a collection of branches $\mathcal{C}$ which contains more than a fraction*

$$1 - \frac{1}{2}\left(\frac{1}{\sqrt{2\pi}}\int_{C/2}^{\infty} e^{x^2/2}dx - \frac{2A}{\sqrt{h/2}}\right)$$

of the branches of t_1 and t_2. Then there exist branches $\beta_1 \in t_1$ and $\beta_2 \in t_2$ both in $\mathcal{C}$ which remain within a distance d of each other from height 0 to $h/2$.

Proof: Let F_1 be the set of branches in t_1 which land closer to the center of t_2 than t_1 at height $h/2$. Because the distance between the two trees is at most $C\sigma_h$,

$$P(F_1) \geq \frac{1}{\sqrt{2\pi}}\int_{C/2}^{\infty} e^{-x^2/2}dx - \frac{2A}{\sqrt{h/2}}.$$

If F_2 is the set of branches in t_2 which land closer to the center of t_1 than t_2, then the same estimate works for the measure of F_2.

Let b_1 be any branch in F_1 and b_2 any branch in F_2. By the definitions of F_1 and F_2, there is some height between $h/2$ and h at which the two branches cross. If they cross between height h' and $h'+1$, then they pass within $2s$ of each other at height h'. Choose b_1 and b_2 so that h' is as large as possible. Match the two

h'-subtrees containing b_1 and b_2 via the identity map. Then every set of paired branches passes within $2s + 2B < d$ of each other at every height between 0 and $h/2$. By continuing to match branches in this way in order of decreasing h' until we run out of unmatched branches in F_1 or F_2, we are able to match at least a fraction $(1/\sqrt{2\pi}) \int_{C/2}^{\infty} e^{-x^2/2}dx - 2A/\sqrt{h/2}$ of the branches of each subtree. More than half of the matched branches in each tree must lie in $\mathcal{C}$, so we are forced to have matched two branches β_1 and $\beta_2 \in \mathcal{C}$. ∎

The following definitions give a collection of branches where a behaves well.

Definition 5.2: A branch b is **good** for the automorphism a if the following conditions hold:

1. The Q-label for b in the n-tree over (x, ω) is the same as the Q-label for $a(b)$ in the n-tree over (x', ω'),
2. The branches b and $a(b)$ both satisfy the four nondegeneracy properties, Properties 1–4.

Definition 5.3: A branch b is **very good** for a if for all $j \leq n$, the j-subtree containing b has at least $2^j(1 - C_1)$ good branches, where

$$C_1 = \frac{\delta_0}{16} \frac{2(e^{-1/2}d - 3A)}{d\sqrt{2\pi}} \int_{102}^{\infty} e^{-x^2/2}dx.$$

LEMMA 5.4: *If (x, ω) and (x', ω') are $(n, n_0, \delta, \delta_0)$-nondegenerate and if (x, ω), (x', ω') are (δ, n_0) close in v_n, then the fraction of good branches in $\mathcal{T}_{n,(x,\omega)}$ is greater than $1 - 3\delta$. The fraction of very good branches is greater than $1 - 3\delta/C_1$.*

Proof: Since $v_n((x, \omega), (x'\omega')) < \delta$, the fraction of branches with matching Q-labels under a is greater than $1 - \delta$. The nondegeneracy of (x, ω) and (x', ω') implies that the fraction of branches b satisfying the nondegeneracy properties for both b and $a(b)$ is greater than $1 - 2\delta$. The fraction of good branches is therefore greater than $1 - 3\delta$.

Now, suppose branch b passes though a subtree in which the fraction of good branches is less than $1 - C_1$. Then it must pass though a maximal such subtree. Each such maximal subtree of height n' contains at least $2^{n'}C_1$ branches which are not very good. Summing over all such maximal subtrees and dividing through by C_1, we get that the fraction of branches passing through such maximal trees is less than $3\delta/C_1$. ∎

As a first step to establishing the desired mapping properties of very good branches, we first prove that if two very good branches pass close together in $\mathcal{T}_{n,(x,\omega)}$, then their images under a cannot pass too far apart in $\mathcal{T}_{n,(x',\omega')}$.

LEMMA 5.5: *Assume (x,ω) and (x',ω') are as above. Given b, b' any two very good branches for a and $h \geq n_0$, if b and b' land a distance at most d apart at height h, then $a(b)$ and $a(b')$ land a distance at most $2sh$ apart at height h.*

Proof: Assume there exist very good branches b and b' in the n tree over (x,ω) and a height h such that b and b' pass within d of each other at height h, but $a(b)$ and $a(b')$ are in h-subtrees that are more than a distance $2sh$ apart. We will use the notation $t_{b,h}$ for the h-subtree containing b.

First we prove that $v_h(t_{a(b),h}, t_{a(b'),h}) < \delta_0$. Let t be the distance between b and b' at height h. By assumption, $t \leq d$. The branches are very good, so the fraction of branches in $t_{b,h}, t_{b',h}$ landing within $2d$ of a Q-cut is less than $\delta_0/4 + C_1 \leq \delta_0/2$. If we map $t_{b,h}$ to $t_{b',h}$ via the identity tree automorphism, corresponding branches will lie within $t + B < 2d$ of each other since the shapes of the two trees are B-similar. If a branch of $t_{b,h}$ does not lie within $2d$ of a Q-cut, the branch and its corresponding branch in $t_{b',h}$ will have the same label. Thus, $v_h(t_{b,h}, t_{b',h}) < \delta_0/2$.

Since b and b' are very good, at least a fraction $1-C_1 > 1-\delta_0/4$ of the branches of each h-tree have a Q label which agrees with the label on the image branch under tree automorphism a. Thus, $v_h(t_{b,h}, t_{a(b),h}) < \delta_0/4$ and $v_h(t_{b',h}, t_{a(b'),h}) < \delta_0/4$, so by the triangle inequality, $v_h(t_{a(b),h}, t_{a(b'),h}) < \delta_0$. Each of the tree automorphisms used in establishing these distances acts as the identity up to height n_0, so the two trees are (δ_0, n_0)-close in v_h.

The definition of a good branch implies that $a(b)$ and $a(b')$ cannot land in Γ_h and by assumption they do not land within $2sh$ of each other at height h. Hence the distance g between $a(b)$ and $a(b')$ at height h must be at least h^5. By the n_0-regularity property, we can find $h', g^{1/4} < h' < g^{1/2}$ such that at height h', the branches b and b' pass within $204\sqrt{h'}$ of each other.

By the definition of a very good branch, $t_{b,h'}$ and $t_{b',h'}$ each contain at most a fraction

$$C_1 = \frac{\delta_0}{16}\Big(\frac{1}{\sqrt{2\pi}}\int_{102}^{\infty} e^{-x^2/2}dx\Big)$$

of branches which are not good. Let $\mathcal{C}$ be the set of branches passing through $h'/4$-subtrees with at least a fraction $1-\delta_0/4$ good branches. Then the fraction of branches in $t_{b,h'}$ and $t_{b',h'}$ which are not in $\mathcal{C}$ is less than $\frac{1}{4}((1/\sqrt{2\pi})\int_{102}^{\infty} e^{-x^2/2}dx)$. For $h' > n_0$ sufficiently large, then the requirements for the Matching Lemma (5.1) are satisfied and there must exist branches β in $t_{b,h'}$ and β' in $t_{b',h'}$ which are in $\mathcal{C}$ and which pass within d of each other at height $h'/4$.

By the way in which h' was chosen, the distance between the two trees $t_{a(\beta),h'/4}$

and $t_{a(\beta'),h'/4}$ at height $h'/4$ is between $2s(h'/4)$ and $(h'/4)^5$. Finally, we note that the trees $t_{a(\beta),h'/4}$ and $t_{a(\beta'),h'/4}$ are (δ_0, n_0)-close in $v_{h'/4}$ and so all of the branches of both trees land in $\Gamma_{h'/4}$, violating the second of the nondegeneracy properties. This contradicts the way in which β and β' were chosen and the proof is complete. ∎

Using the previous lemma and the first nondegeneracy property, we can now prove that if two very good branches pass within a distance d of each other at height n_0, then their images under a must pass within $2a_1$ of each other at height n_0.

LEMMA 5.6: *Let (x,ω) and (x',ω') be (δ, n_0) close in v_n and $(n, n_0, \delta, \delta_0)$-nondegenerate. Let b_1 and b_2 be two very good branches in the n-tree over (x,ω) passing through the points r_1 and r_2 at height n_0 with $0 \le r = r_1 - r_2 \le d$. Then the branches $a(b_1)$ and $a(b_2)$ in the n tree over (x',ω') pass through points s_1 and s_2 at height n_0 with $|s_1 - s_2| < r + a_1$.*

Proof: If we can show that

$$\bar{d}_{[-\sigma_{n_0},\sigma_{n_0}]}(T_{r_1}(\omega), T_{s_1}(\omega')) < \delta_0/8, \tag{1}$$

$$\bar{d}_{[-\sigma_{n_0},\sigma_{n_0}]}(T_{r_2+r}(\omega), T_{s_2+r}(\omega')) < \delta_0/8, \tag{2}$$

then since $r_1 = r_2 + r$, we will have proven that

$$\bar{d}_{[-\sigma_{n_0},\sigma_{n_0}]}(T_{s_1}(\omega'), T_{s_2+r}(\omega')) < \delta_0/4.$$

By Lemma 5.5, the distance between the two branches $a(b_1)$ and $a(b_2)$ at height n_0 is at most $2sn_0$, so $|s_1 - (s_2 + r)| < 3n_0$. Since $a(b_1)$ cannot land in $\overline{B}(n_0, \delta_0)$ by the first nondegeneracy property, this implies that $a(b_1)$ and $a(b_2)$ pass within $a_1 + r$ at height n_0.

We now prove inequality 1. The proof of inequality 2 is similar. Let J be the set of integers j between $-\sigma_{n_0}$ and σ_{n_0} such that

1. $r_1 + j \in [r_1 + kd, r_1 + (k+1)d)$ for some k with a good branch from t_{n_0,b_1} landing in $[r_1 + kd, r_1 + (k+1)d)$ at height 0,
2. $r_1 + j$ is not within d of a Q-cut for ω,
3. $s_1 + j$ is not within $2d$ of a Q-cut for ω'.

The number of branches from the t_{n_0,b_1} landing in any interval $[kd, (k+1)d) \subset [-\sigma_{n_0}, \sigma_{n_0}]$ is at least $2^{n_0}(e^{-1/2} - 3A)/\sigma_{n_0}$. Since b_1 is very good, t_{n_0,b_1} contains at most $2^{n_0}C_1$ branches which are not good. Therefore, there can be at most $C_1\sigma_{n_0}/(e^{-1/2} - 3A)$ intervals with no good branches. There are $2\sigma_{n_0}/d$ such intervals in the 1-middle of the n_0-subtree containing b, so the fraction of intervals with no good branches is less than $C_1 d/2(e^{-1/2}d - 3A) < \delta_0/16$. The fraction of integers j not within $2d$ of a Q-cut for either $T_{r_1}(\omega)$ or $T_{s_1}(\omega')$ is greater than $1 - \delta_0/16$ because $a_2 > 64d/\delta_0$.

If $j \in J$, then there is a good branch β landing within d of $r_1 + j$ at height 0. The tree automorphism a acts as the identity on the subtrees of height n_0, so the branch $a(\beta)$ lands within $d + B < 2d$ of $s_1 + j$. Since β is a good branch, the Q-labels on β and $a(\beta)$ agree. We chose j so that there can be no Q-cuts in the way and so $Q(T_{r_1+j}(\omega)) = Q(T_{s_1+j}(\omega'))$. Hence, the set of indices on which the two Q-words disagree is contained in the complement of J and the $\bar{d}$ distance between them is less than $\delta_0/8$. ∎

We now know that two very good branches lying close together at height n_0 will map under a to branches which are close at height n_0. However, the scale on which this lemma operates is very small. Consider an interval $I \subset \mathbb{R}$. Our estimates for $\mu_{n,(x,\omega)}(I)$ given by Theorem 3.1 are rather nearsighted in the following sense. Near the center of the distribution, the error will be manageable when the length of I is on the order of A, where A is as in Theorem 3.1. If I is near $\pm c\sqrt{n}$, however, the length of I must be on the order of $e^{c^2/2}A$ in order to manage the error. For this reason, we prove a rigidity result that will operate on the scale appropriate to our choice of c.

Select two nondegenerate points (x, ω) and (x', ω'). Up to now, we have only used that (x, ω), (x', ω') are (δ, n_0)-close in $v_n^{Q'}$. We now require that the two points be $(\delta, 2n_0)$-close in v_n^Q. Let a be a tree automorphism realizing this close v_n-matching.

Definition 5.7: A $2n_0$-subtree of $\mathcal{T}_{n,(x,\omega)}$ is very good for the automorphism a if the fraction of branches in the subtree which are very good for a is at least $1 - C_2$ where $C_2 = \frac{1}{4}\int_{1/2}^{\infty} e^{-x^2/2}dx$.

LEMMA 5.8: *The fraction of $2n_0$-subtrees in $\mathcal{T}_{n,(x,\omega)}$ which are very good is at least $1 - 3\delta/C_1C_2$.*

Proof: By Lemma 5.4, the fraction of branches in the n-tree which are very good is at least $1 - 3\delta/C_1$. Each $2n_0$-subtree which is not very good has at least $C_2 \cdot 2^{2n_0}$ branches which are not very good. Hence, the fraction of subtrees which are not very good is at most $3\delta/C_1C_2$. ∎

LEMMA 5.9 (Local Rigidity Lemma): *Let t_1 and t_2 be two $2n_0$-subtrees of $\mathcal{T}_{n,(x,\omega)}$ which are very good. Let t_1 and t_2 be centered at r_1 and r_2 at height*

$2n_0$, and let $a(t_1)$ and $a(t_2)$ be centered at s_1 and s_2. If $|r_1 - r_2| \leq \sigma_{n_0}$, then $|(r_1 - r_2) - (s_1 - s_2)| < 2a_1 + B$.

Proof: The goal is to show that there is a very good branch β_1 in t_1 and a very good branch β_2 in t_2 such that β_1 and β_2 pass within d of each other at height n_0. Then by Theorem 5.6, $a(\beta_1)$ and $a(\beta_2)$ must pass within $2a_1$ of each other in $\mathcal{T}_{n,(x',\omega')}$ at height n_0. Because we are requiring that a act as the identity automorphism up to height $2n_0$, this implies that the distance between $a(t_1)$ and $a(t_2)$ at height $2n_0$ is within $2a_1 + B$ of the distance between t_1 and t_2.

It only remains to find the branches β_1 and β_2. This follows from the Matching Lemma (5.1). Let $\mathcal{C}$ be the set of branches which are very good. Since t_1 and t_2 are very good subtrees, the fraction of branches in each which are in $\mathcal{C}$ is at least $1 - C_2$, where C_2 was chosen to satisfy the conditions of Lemma 5.1 with $C = 1$. Since t_1 and t_2 are at most $\sigma_{n_0} < \sigma_{2n_0}$ apart, there must exist β_1 and β_2 as desired and the proof is complete. ∎

An important consequence of Lemma 5.9 is that a maps all very good $2n_0$-subtrees lying in any interval of length σ_{n_0} at height $2n_0$ to subtrees lying in an interval of length less than $\sigma_{n_0} + 3a_1$ at height $2n_0$ in $\mathcal{T}_{n,(x',\omega')}$. By taking σ_{n_0} to be at least of the order of magnitude of $e^{c^2}A$, we can now consider the map a on a scale that is visible even given the "nearsightedness" of our estimates for the distributions.

6. Relationship between v and $\bar{f}$

In the previous section, we saw that the map $b \mapsto a(b)$ at height $2n_0$ has a rigid structure on a local scale. We wish to use this information to prove that the map is essentially monotone when restricted to an appropriate set of branches. This monotone mapping will then give us the desired $\bar{f}$-match. We restrict our attention first to very good $2n_0$-subtrees which are located in intervals on the scale of σ_{n_0} with a high concentration of very good subtrees.

Definition 6.1: A $2n_0$-subtree, t, of $\mathcal{T}_{n,(x,\omega)}$ is a **map tree** for a if t is very good and at height $2n_0$ it passes through an interval $K_i = [i\sigma_{n_0}/16, (i+1)\sigma_{n_0}/16) \subset [-c\sigma_{n-2n_0}, c\sigma_{n-2n_0}]$ with $i \in \mathbb{Z}$ such that the fraction of subtrees through K_i at height $2n_0$ which are very good is greater than $1 - e^{-4c}/32$.

We say that t is a **good map tree** if it is a map tree for a at height $2n_0$ and if $a(t)$ is a map tree for a^{-1} at height $2n_0$.

LEMMA 6.2: *The fraction of $2n_0$-subtrees of $\mathcal{T}_{n,(x,\omega)}$ which are good map trees is at least* $1-\frac{4\cdot 3\cdot 32}{C_1C_2}\delta e^{4c}-\frac{8}{c}e^{-c^2/2}$.

Proof: If a $2n_0$-subtree is not a map tree, then either it is not very good or it passes through a K_i at height $2n_0$ which is not in $[-c\sigma_{n-2n_0}, c\sigma_{n-2n_0}]$ or for which the fraction of subtrees which are very good is at least $e^{-4c}/32$. By Theorem 3.1, the fraction of $2n_0$ subtrees not passing through the c-middle of the tree at height $2n_0$ is less than

$$2\int_c^\infty e^{-x^2/2}dx+\frac{2A}{\sqrt{n-2n_0}}\le\frac{2e^{-c^2/2}}{c}+\frac{2A}{\sqrt{n-2n_0}}\le\frac{4}{c}e^{-c^2/2}.$$

By Lemma 5.8, the fraction of $2n_0$-subtrees which are not very good is at most $3\delta/C_1C_2$. Thus, there can be at most a fraction

$$\frac{3\delta}{C_1C_2}+32e^{4c}\frac{3\delta}{C_1C_2}+\frac{4}{c}e^{-c^2/2}\le\frac{2\cdot 3\cdot 32}{C_1C_2}\delta e^{4c}+\frac{4}{c}e^{-c^2/2}$$

of subtrees which are not map subtrees for a. The definition of a very good subtree is symmetric in a and a^{-1}, so the fraction of subtrees which are good map trees is at least $1-\frac{4\cdot 3\cdot 32}{C_1C_2}\delta e^{4c}-\frac{8}{c}e^{-c^2/2}$. ∎

LEMMA 6.3: *Let (x,ω) and (x',ω') be in G and suppose the two points are $(\delta,2n_0)$-close in v_n. Let t_1 and t_2 be map trees at height $2n_0$, which pass through the points r_1 and r_2 respectively at height $2n_0$. Let $a(t_1)$ and $a(t_2)$ pass through s_1 and s_2 at height $2n_0$. If $\sigma_{n_0}/2\le r=|r_2-r_1|\le .1\sigma_n$ and $s=|s_2-s_1|$, then $s\le 2r$.*

Proof: The proof is by contradiction. Suppose there exist t_1 and t_2 as in the statement of the lemma with $s>2r$. Let K and K' be intervals as in the Definition 6.1, containing t_1 and t_2 respectively. By the Local Rigidity Lemma (5.9), all of the images under a of the very good subtrees through K at height $2n_0$ must pass through an interval $a(K)$ of length $\sigma_{n_0}/16+3a_1$ at height $2n_0$ and similarly we can find an interval $a(K')$ containing the images of all very good subtrees through K' at height $2n_0$.

Let $R=\lceil r^2/c\rceil$ and let J be the interval (r_1,r_2) or (r_2,r_1), whichever is non-empty. Consider t any $(2n_0+R)$-subtree of $\mathcal{T}_{n,(x,\omega)}$ which passes through J at height $2n_0+R$. Since we have assumed $s>2r$, either $|s_1-a(t)|-|r_1-t|\ge r/2$ or $|s_2-a(t)|-|r_2-t|\ge r/2$. We assume without loss of generality that the former is true for at least half of the subtrees t. For any such t, the ratio of the number of $2n_0$-subtrees of $a(t)$ in $a(K)$ to the number of $2n_0$-subtrees of t in K

is at most $2\exp(-c/2) < 1/2$. Any $2n_0$-subtree of t whose image does not lie in $a(K)$ cannot be very good since a maps the very good $2n_0$-subtrees of K into $a(K)$.

We can now estimate the fraction of $2n_0$-subtrees through K which are not very good by counting those which are subtrees of a $(2n_0+R)$-subtree t through J such that $a(t)$ passes at least $r/2$ further from s_1 than t does from r_1. K lies in the $2\sqrt{c}$-middle of each such subtree t. We assume that $0 \le |r_1| \le |r_2| \le c\sigma_n$; the other case is similar. Under this assumption, this fraction is at least

$$\frac{\frac{1}{2}\left(\frac{\exp[-\frac{1}{2}(|r_2|/\sigma_{n-2n_0-R})^2]\ell(J)-2A}{\sigma_{n-2n_0-R}}\right)\left(\frac{\exp[-4c/2]\ell(K_1)-2A}{\sigma_R}\right)}{\left(\frac{\exp[-\frac{1}{2}(|r_1|/\sigma_{n-2n_0})^2]\ell(K_1)+2A}{\sigma_{n-2n_0}}\right)}$$

$$\ge \frac{(\frac{1}{2}\exp[-\frac{1}{2}(|r_2|/\sigma_{n-2n_0-R})^2])(\frac{1}{2}\exp[-4c/2]\ell(K_1))}{8\exp[-\frac{1}{2}(r_1/\sigma_{n-2n_0})^2]\ell(K_1)} \cdot \frac{\ell(J)\cdot\sigma_{n-2n_0}}{\sigma_{n-2n_0-R}\sigma_R}$$

$$\ge \frac{1}{32}\exp\Big[-\frac{1}{2}\Big(\Big(\frac{|r_1|+.1\sigma_n}{\sigma_{n-2n_0-R}}\Big)^2-\Big(\frac{|r_1|}{\sigma_{n-2n_0}}\Big)^2+4c\Big)\Big]$$

$$\ge \frac{1}{32}\exp\Big[-\frac{1}{2}\Big(\Big(\frac{|r_1|+.1\sigma_n}{\sqrt{1-1/c}\,\sigma_{n-2n_0}}\Big)^2-\Big(\frac{|r_1|}{\sigma_{n-2n_0}}\Big)^2+4c\Big)\Big]$$

$$\ge \frac{1}{32}e^{-4c}.$$

This contradicts the fact that t_1 is a map tree, which means that the fraction of $2n_0$-subtrees through K which are very good is at least $1-e^{-4c}/32$. ∎

By applying Lemma 6.3 to $a^{-1}, a(t_1)$ and $a(t_2)$, we immediately obtain that if t_1 and t_2 are good map trees then $\frac{1}{2}r \le s \le 2r$. With one further restriction on the $2n_0$-subtrees we are interested in, we will be able to prove a monotonicity result which is valid on $2n_0$-subtrees which are on the order of σ_n apart, rather than σ_{n_0}, which was covered under Lemma 5.9.

Definition 6.4: Let $K_k = [k\frac{\sigma_{n_0}}{16}, (k+1)\frac{\sigma_{n_0}}{16}) \subset [-c\sigma_{n-2n_0}, c\sigma_{n-2n_0}]$. We say that K_k is a **nice interval** if for every interval $[i,j] \ni k$ with K_i, K_j contained in the c-middle of $\mathcal{T}_{n,(x,\omega)}$, at least a fraction $\frac{15}{16}$ of the K_l with $l \in [i,j]$ have a good map tree passing through at height $2n_0$.

A good map tree through a nice interval at height $2n_0$ is called a **nice tree**.

LEMMA 6.5: *If (x,ω) and (x',ω') are in G and are $(\delta, 2n_0)$-close in v_n with $\delta \le 1/(3\cdot 80\cdot 16F(c))$ where $F(c) = 4\cdot 3\cdot 32e^{c^2/2+4c}/C_1C_2$, then the fraction of the $K_k \in [-c\sigma_{n-2n_0}, c\sigma_{n-2n_0}]$ which are nice is at least $1-1/40c$. Further, if t_1 and t_2 are nice $2n_0$-subtrees which lie a distance at least σ_{n_0} apart at height $2n_0$,* *then $a(t_1)$ lies to the left of $a(t_2)$ at height $2n_0$ if and only if t_1 lies to the left of t_2 at height $2n_0$.*

Proof: By Lemma 6.2, the fraction of $2n_0$-subtrees in $\mathcal{T}_{n,(x,\omega)}$ which are good map trees is at least $1 - e^{-c^2/2}(F(c)\delta + 8c^{-1}) \geq 1 - 2e^{-c^2/2}F(c)\delta$. Each interval K_k in the c-middle of the n-tree has at least

$$2^{n-2n_0}\frac{e^{-c^2/2}(\sigma_{n_0}/16) - 2A}{\sigma_{n-2n_0}} \geq 2^{n-2n_0}\frac{e^{-c^2/2}(\sigma_{n_0}/16)}{2\sigma_{n-2n_0}}$$

$2n_0$-subtrees passing through at height $2n_0$. There are $2c\sigma_{n-2n_0}(\sigma_{n_0}/16)^{-1}$ such intervals. Thus, the fraction of intervals $K_k \subseteq [-c\sigma_{n-2n_0}, c\sigma_{n-2n_0}]$ which do not have a good map tree passing through at height $2n_0$ is at most $2c^{-1}F(c)\delta$.

Restricting to the c-middle at height $2n_0$, the complement of the set of nice intervals can be written as the union of all intervals $[i,j]$ for which at least $1/16$ of the intervals $K_l, l \in [i,j]$ have no good map trees. We can cover at least one third of this set by disjoint such intervals. Therefore, the fraction of the K_k which are not nice is at most

$$2 \cdot 3 \cdot c^{-1} 16 F(c)\delta \leq 1/40c.$$

We now need to prove the monotonicity result for nice branches which are sufficiently far apart. First note that the result above implies that there are at least two nice intervals K_k in any interval of length $.1\sigma_n$, so that it is sufficient to prove the result under the additional assumption that the distance between t_1 and t_2 is less than $.1\sigma_n$.

Assume without loss of generality that t_1 passes through a point r_1 at height $2n_0$ to the left of r_2, the point which t_2 passes through at height $2n_0$. Assume further that $a(t_1)$ lies to the right of $a(t_2)$, $a(t_1)$ at s_1 and $a(t_2)$ at s_2 so that $s_2 - s_1 < 0$. We may also take t_2 to be the rightmost good map tree with these properties and with $r_2 - r_1 > 3a_1$ (note that we are not assuming that t_2 is nice for this part of the argument). By the Local Rigidity Lemma (5.9), we must then have $r_2 - r_1 > \sigma_{n_0}$.

Since t_1 is a nice tree, it passes through a nice interval K_k at height $2n_0$. The tree t_2 passes through some interval $K_{k'}$ at height $2n_0$. Since $r_2 - r_1 > \sigma_{n_0}$, $k' - k \geq 16$. By the definition of a nice interval, at most $1/16$ of the intervals K_l with $l \in [k, k']$ can have no good map trees. Therefore, there is a good map tree t lying at the point r with $4(r_2 - r) < (r_2 - r_1)$. But if $a(t)$ is centered at s, then $s > s_1$ by the minimality of t_2 and, by Lemma 6.3 and the construction, we get

$$s - s_2 = (s - s_1) + (s_1 - s_2) > s_1 - s_2 \geq \frac{1}{2}(r_2 - r_1) \geq 2(r_2 - r).$$

However, $r_2 - r > \sigma_{n_0}/2$ by the Local Rigidity Lemma, so this contradicts Lemma 6.3. Therefore, we must have $s_2 - s_1 > 0$. ∎

As a consequence of Lemma 5.9 and Lemma 6.5, if t_1 and t_2 are nice $2n_0$-trees which are at least a distance $3a_1$ apart at height $2n_0$, then the mapping $t_i \mapsto a(t_i), i = 1, 2$ is monotone. We wish to use this monotonicity to give us a close $\bar{f}$-match between the symbolic names $\bar{\omega}$ and $\bar{\omega}'$ when (x, ω) and (x', ω') are close in v_n. However, the v_n matching gives us information about the Q-partition at the base of the tree and our monotone map is up at height $2n_0$. To get around this difficulty, we require that (x, ω) and (x', ω') be $(\delta, 2n_0)$-close in $v_n^{Q_{n_0}}$ where Q_{n_0} is a finite partition which is a refinement of Q giving information about the Q-name of branches at height $2n_0$. We also change part 1 in Definition 5.2 of a good branch to require that the Q_{n_0}-labels of b and $a(b)$ agree. The rest of the argument proceeds as before.

Definition 6.6: Define Q_{n_0} on Ω to be a finite partition refining Q such that $Q_{n_0}((x, \omega))$ also specifies the symbolic $\bar{Q}$-name of $\hat{T}_f^{n_0}((x, \omega))$ on the interval $[-\sigma_{n_0}/16, \sigma_{n_0}/16]$.

We are now ready to prove that a close v_n matching of two nondegenerate points gives an $\bar{f}$ matching for the scenery coordinates. We make this precise in the following theorem.

THEOREM 6.7 (Relation between v_n and $\bar{f}$): *Given c, Q, and δ_0, if we let $\delta = 1/(3 \cdot 80 \cdot 16F(c))$ and choose n_0 large enough, the following holds. If n is sufficiently large and (x, ω) and (x', ω') are in $G = G(\delta_0, \delta, n_0, n)$ and are $(\delta, 2n_0)$-close in $v_n^{Q_{n_0}}$, then*

$$\bar{f}^Q(\bar{\omega}, \bar{\omega}') < 1/c,$$

where $\bar{\omega}$ and $\bar{\omega}'$ are the symbolic Q-names for ω, ω' on $[-c\sigma_n, c\sigma_n]$.

Proof: Let a be a tree automorphism realizing the close v_n match between (x, ω) and (x', ω'). By Lemma 6.5, the fraction of the intervals K_k in the c-middle of $\mathcal{T}_{n-2n_0,(x,\omega)}$ which are nice for a is at least $1 - 1/40c$. Our $\bar{f}$-matching will be realized on the nice intervals. Each nice K_k has a nice $2n_0$-subtree t_k through a point $r_k \in K_k$. $a(t_k)$ passes through some point s_k at height $2n_0$. Since t_k is nice, it contains a good branch b_k and so the symbolic Q-name for ω on the interval of length $\sigma_{n_0}/8$ centered around r_k matches the symbolic Q-name for ω' on the interval of the same length centered around s_k. Thus, we can match the symbolic name on $K_k \subset [r_k - \sigma_{n_0}/16, r_k + \sigma_{n_0}/16]$ exactly to that on the corresponding interval for ω' contained in $[s_k - \sigma_{n_0}/16, s_k + \sigma_{n_0}/16]$. Call this map on the nice intervals ϕ.

The domain of ϕ is the set of nice intervals in $[-c\sigma_{n-2n_0}, c\sigma_{n-2n_0}]$ and the range is contained in $[-c\sigma_{n-2n_0}, c\sigma_{n-2n_0}]$ as well by the definition of good map trees. Clearly, we can take n sufficiently large that

$$|\bar{f}_{[-c\sigma_n, c\sigma_n]}(\bar{\omega}, \bar{\omega}') - \bar{f}_{[-c\sigma_{n-2n_0}, c\sigma_{n-2n_0}]}(\bar{\omega}, \bar{\omega}')| < 1/3c.$$

It only remains to show that the map ϕ is monotone. Certainly, each restriction to a nice interval K_k is monotone. Let $k < l$, where K_k and K_l are nice intervals. If t_k and t_l are at least σ_{n_0} apart at height $2n_0$, then by Lemma 6.3, $a(t_k)$ and $a(t_l)$ are at least $\sigma_{n_0}/2$ apart, so the images of $\phi|_{K_k}$ and $\phi|_{K_l}$ are disjoint. Furthermore, by Lemma 6.5, since $k < l$, $a(t_k)$ lies to the left of $a(t_l)$, so that ϕ is monotone when restricted to $K_k \cup K_l$. Now suppose t_k and t_l lie within a distance σ_{n_0} of each other at height $2n_0$. By Lemma 5.9, $|(r_k - r_l) - (s_k - s_l)| < 3a_1$. Therefore, the images of $\phi|_{K_k}$ and $\phi|_{K_l}$ can overlap only if $l = k + 1$ and, even in this case, the size of the overlap is at most $3a_1$. If we delete from each K_k any symbols whose images lie in the overlap, then ϕ is monotone and gives an $\bar{f}$-match to within $1/20c + 2/3c + 3a_1/(\sigma_{n_0}/16) < 1/c$. ∎

7. Entropy is invariant

Let $(\Omega_S, \mathcal{F}_S, \mu_S, S_t)$ and $(\Omega_T, \mathcal{F}_T, \mu_T, T_t)$ be two ergodic flows and take two functions f, g on $[0,1]$, Hölder continuous with $\int f = \int g = 0$ and with the appropriate normalizations. Suppose that $\hat{S}_g$ and $\hat{T}_f$ are σ-algebra isomorphic via a σ-algebra isomorphism Φ. We wish to conclude that $h(S) = h(T)$. Assume that $h(S) < h(T)$, so that in particular S has finite entropy.

Since Φ is a σ-algebra isomorphism, it induces a collection of invertible, measure-preserving maps

$$\Phi_n \colon [0,1] \times \Omega_S \to [0,1] \times \Omega_T$$

for $n \geq 1$ such that $\Phi(\mathcal{T}_{n,(x,\omega)}) = \mathcal{T}_{n,\Phi_n(x,\omega)}$.

Choose partitions P of Ω_S and Q of Ω_T as flows under a function as described in Section 3 and so that (S_t, P) has full entropy and $h(T, Q) > h(S)$.

Partition $[0,1] \times \Omega_S$ into sets $E_{(x,\omega),k,c',n}$ consisting of points (x', ω') such that:

1. x and x' agree on the first k coordinates and
2. $P(S_j(\omega)) = P(S_j(\omega'))$ for all integers $j \in [-c'\sigma_n, c'\sigma_n]$.

Note that σ_n is taken to be as calculated for X_f rather than X_g. We wish to prove that, except on a set of small measure and for appropriate choices of k, c', n, each set $\Phi_n(E_{(x,\omega),k,c',n})$ is contained in a small v_n^Q-neighborhood.

LEMMA 7.1: *Fix c, P, Q with $\delta = 1/(3 \cdot 80 \cdot 16F(c))$ and n_0 large as in the previous sections. When k, n are sufficiently large and $c' = \sqrt{c^2 + 8c} + 1$, there is a good set $\Gamma_{k,c',n} \subset [0,1] \times \Omega_S$ with $\mu_S(\Gamma_{k,c',n}) > 1 - 1/c$ such that if $(x,\omega), (x',\omega') \in E_{(x,\omega),k,c',n} \cap \Gamma_{k,c',n}$, then $\Phi_n(x,\omega)$ and $\Phi_n(x',\omega')$ are $(\delta, 2n_0)$-close in $v_n^{Q_{n_0}}$.*

Proof: For any point (x',ω'), $\Phi(\mathcal{T}_{n,(x',\omega')}) = \mathcal{T}_{n,\Phi_n(x',\omega')}$ and Φ restricted to the tree acts as a tree automorphism, $a_{(x',\omega')}$. For $(x',\omega') \in E_{(x,\omega),k,c',n}$, consider the tree automorphism $a = a_{(x',\omega')} \circ Id \circ a^{-1}_{(x,\omega)}$ where Id is the identity tree automorphism from $\mathcal{T}_{n,(x,\omega)}$ to $\mathcal{T}_{n,(x',\omega')}$. Then a: $\mathcal{T}_{n,\Phi_n(x,\omega)} \to \mathcal{T}_{n,\Phi_n(x',\omega')}$. If (x',ω') is in the good set, then this a will nearly give the desired $v_n^{Q_{n_0}}$-match.

To prove this, let P_m be the refinement of P achieved by cutting each full-height interval from the flow under a function defining P into m pieces of equal length. Now define a sequence of partitions on $[0,1] \times \Omega_S$ by

$$C_{m,n}((z,\eta)) = \{(z',\eta') | \, (z_0,\ldots,z_{n-1}) = (z'_0,\ldots,z'_{n-1}) \text{ and } \bar{\eta}_m = \bar{\eta}'_m\}$$

where $\bar{\eta}_m$ is the symbolic P_m-name along the orbit of ω on $[-m,m]$. Clearly, as $n, m \to \infty$, $C_{m,n}$ converges to the full σ-algebra on $[0,1] \times \Omega_S$.

For each pair m, n, we can define a finite code $\phi_{m,n}$ for Φ which is constant on the atoms of $C_{m,n}$. If $\Phi(x,\omega) = (y,\psi)$, then we wish $\phi_{m,n}$ to code for $Q_{n_0}(\psi)$ and $(y_0,\ldots,y_{2n_0})$. Choose $m > 24c/\delta$ large enough so that for all n sufficiently large, $\phi_{n,m}$ codes this information correctly except on a set of measure $\delta/12c$.

In order to get the finite codes to converge, we had to refine P to P_m, which is constant only on intervals of length a_1/m or a_2/m. We would like corresponding branches in $\mathcal{T}_{n,(x,\omega)}$ and $\mathcal{T}_{n,(x',\omega')}$ to land in the same P_m partition most of the time. To this purpose, we take k large. Because g is Hölder continuous, if x and x' agree on the first k coordinates where k is large enough, then the n-trees over x and x' will be $1/m^2$-close for all n.

We will call a point (z,η) **nice** if $\phi_{n,m}$ codes correctly for (z,η) and if η does not lie within $2/m^2$ of a P_m-cut. The set $\Gamma_{k,c',n}$ is then defined as the set of points (z,η) such that at least a fraction $1-\delta/4$ of its branches are nice and such that the P-name of η on the integers in $[-c'\sigma_n, c'\sigma_n]$ specifies the P-name on the interval up to a shift of at most $1/m^2$. The measure of all points which are not nice is less than $2\delta/12c$, so by the Chebyshev inequality, the measure of points with at least a fraction $1-\delta/4$ of nice branches is at least $1-2/3c$. By taking n sufficiently large, we can ensure that, except on a set of measure $1/3c$, the P-name of a point η on the integers in $[-c'\sigma_n, c'\sigma_n]$ determines the full P-name on this interval up to a shift of at most $1/m^2$. Therefore, $\mu_S(\Gamma_{k,c',n}) > 1-1/c$.

Now, suppose that $(x,\omega),(x',\omega') \in E_{(x,\omega),k,c',n} \cap \Gamma_{k,c',n} \subseteq [0,1]\times\Omega_S$. We wish to show that $\Phi_n(x,\omega)$ and $\Phi_n(x',\omega')$ are close in $v_n^{Q_{n_0}}$. Let b be a branch in the binary n-tree corresponding to a nice branch in both $\mathcal{T}_{n,(x,\omega)}$ and $\mathcal{T}_{n,(x',\omega')}$. Further, suppose that b lands in the interval $[-c'\sigma_n+m, c'\sigma_n-m]$ in both trees. By assumption, the P-names for ω and ω' along the integers in $[-c'\sigma_n, c'\sigma_n]$ agree and determine the full P-name on the interval up to a shift of at most $1/m^2$. By our choice of k, the realizations of b in the two trees land at points that are within $1/m^2$ of each other. Since b is nice, it does not land within $2/m^2$ of a cut for P_m in either tree. Therefore, the $C_{m,n}$-labels on b and b' agree. Since $\phi_{n,m}$ is constant on atoms of $C_{m,n}$, this implies that $Q_{n_0}(a_{(x,\omega)}(b)) = Q_{n_0}(a(a_{(x,\omega)}(b)))$ and a acts as the identity automorphism up to height $2n_0$ on $a_{(x,\omega)}(b)$.

Because (x,ω) and (x',ω') lie in the good set, the fraction of branches b which are nice in both trees is at least $1-\delta/2$. The fraction of branches in $\mathcal{T}_{n,(x,\omega)}$ which land outside the interval $[-c'\sigma_n+m, c'\sigma_n-m]$ when n is large enough is at most

$$2\left(\frac{1}{\sqrt{2\pi}}\int_{c'-(m/\sigma_n)}^{\infty} e^{-x^2/2}dx+\frac{A}{\sigma_n}\right) \leq 4\left(\frac{e^{-(c'-(m/\sigma_n))^2/2}}{c'-(m/\sigma_n)}\right)$$
$$\leq \frac{4e^{-(c^2/2+4c)}}{\sqrt{c^2+8c}} \leq \frac{1}{3\cdot 16\cdot 80F(c)} = \delta/2.$$

Therefore, a realizes a $v_n^{Q_{n_0}}$-distance between $\Phi_n(x,\omega)$ and $\Phi_n(x',\omega')$ of δ. Furthermore, a acts as the identity up to height $2n_0$ on all of the nice branches which get coded correctly, so we can perturb a to a tree automorphism a' which realizes the same $v_n^{Q_{n_0}}$-distance and which acts as the identity up to height $2n_0$. This proves that $\Phi_n(x,\omega)$ and $\Phi_n(x',\omega')$ are $(\delta, 2n_0)$-close in $v_n^{Q_{n_0}}$. ∎

Corollary 7.2: *If $(x,\omega),(x',\omega') \in E_{(x,\omega),k,c',n} \cap \Gamma_{k,c',n}$ and if $\Phi_n(x,\omega)=(y,\psi)$ and $\Phi_n(x',\omega')=(y',\psi')$ are both (n,n_0,δ,δ_0)-nondegenerate, then*

$$\bar{f}^Q_{[-c\sigma_n,c\sigma_n]}(\bar{\psi},\bar{\psi}') \leq 1/c.$$

Proof: By Lemma 7.1, $\Phi_n(x,\omega)$ and $\Phi_n(x',\omega')$ are $(\delta,2n_0)$-close in $v_n^{Q_{n_0}}$. Then by Theorem 6.7, $\bar{f}_{[-c\sigma_n,c\sigma_n]}(\bar{\psi},\bar{\psi}') \leq 1/c$. ∎

Corollary 7.2 gives us a way to map P-names on $[-c'\sigma_n, c'\sigma_n]$ in Ω_S to $\bar{f}^Q_{[-c\sigma_n,c\sigma_n]}$-neighborhoods in Ω_T. This enables us to prove the main theorem.

THEOREM 7.3: *Let $\hat{T}_f$ and $\hat{S}_g$ be σ-algebra isomorphic. Then $h(T) = h(S)$.*

Proof: Assume as above that $h(S) < h(T,Q) < \infty$. Take $\epsilon = 1/c$ and c sufficiently large. Let $F_{(x,\omega),k,c',n}$ be the projection of $\Phi_n(E_{(x,\omega),k,c',n} \cap \Gamma_{k,c',n}) \cap G$ onto the Ω_T-coordinate. Let H be the set of points $\psi \in \Omega_T$ for which the cylinder set $\psi_{-c\sigma_n}^{c\sigma_n}$ for the time-1 map of T_t has the measure predicted by the entropy, at most $2^{-(1-2\epsilon)h(T,Q)\cdot 2c\sigma_n}$. Taking n large, we get $\mu_T(H) \geq 1-\epsilon$ by the Shannon–McMillan Theorem. By Corollary 6.2, $F_{(x,\omega),k,c',n}$ is contained in a symbolic $\bar{f}$-neighborhood of radius ϵ. To estimate the size of a symbolic $\bar{f}$-neighborhood, we first choose an ϵ fraction of indices in each symbolic name on which they will disagree and the labels on these disagreements. Because we are converting from a symbolic to a time-1 name, we also need to consider that each interval of length $a_i, i = 1, 2$ may contain $\lfloor a_i \rfloor$ or $\lceil a_i \rceil$ integers. Thus,

$$\mu_T(F_{(x,\omega),k,c',n} \cap H) \leq \binom{2a_2^{-1}c\sigma_n}{2\epsilon a_2^{-1}c\sigma_n}^2 |Q|^{2\epsilon a_2^{-1}c\sigma_n} \cdot 2^{2a_2^{-1}c\sigma_n} \cdot 2^{-(1-2\epsilon)h(T,Q)\cdot 2c\sigma_n}.$$

For n large, we can cover all but ϵ of $[0,1] \times \Omega_S$ with

$$2^k 2^{(h(S)+\epsilon)2c'\sigma_n}$$

of the sets $E_{(x,\omega),k,c',n}$ and hence this number of the sets $F_{(x,\omega),k,c',n} \cap H$ cover all but 3ϵ of Ω_T. Since we can choose a_2 and then $c = 1/\epsilon$ arbitrarily large, this gives us a contradiction. There are too few sets of the given size to cover $1-3\epsilon$ of Ω_T. Therefore, $h(S) = h(T)$. ∎

Since σ-algebra isomorphism is weaker than measurable isomorphism of dynamical systems, we immediately get the following corollary.

COROLLARY 7.4: *If $\hat{T}_f$ and $\hat{S}_g$ measurably isomorphic, then $h(T) = h(S)$.*

This corollary proves the assertion from the introduction, that our family of smooth maps, $\{\widehat{T}(\alpha)\}_{\alpha>0}$, contains no two which are isomorphic, even though they all have entropy $\log 2$.

We conclude with a remark about when T, T^{-1} maps can be isomorphic to maps of our type, $\hat{S}_g$.

Remark 7.5: Suppose that the T, T^{-1} endomorphism is σ-algebra isomorphic to $\hat{S}_g$. By using a hybrid of the argument given here and the parallel argument given by Heicklen, Hoffman, and Rudolph for T, T^{-1} endomorphisms, we can use finite codes to map cylinder sets for (T, P) to symbolic $\bar{f}$-neighborhoods for (S, Q) and to map cylinder sets for (S, Q) to $\bar{f}$ neighborhoods for (T, P). In either case, the same calculations regarding the entropies of T and S are possible and we get the following corollary. ∎

COROLLARY 7.6: *The entropy of the scenery process is an invariant of σ-algebra isomorphism within the class of random walks on random sceneries which are either T, T^{-1} endomorphisms or of the form $\hat{T}_f$, as described in Section 1.*

References

[1] K. Ball, *Superpolynomial growth in the number of v_n names for random walks on random sceneries*, Ergodic Theory and Dynamical Systems, to appear.

[2] J. Feldman, *New K-automorphisms and a problem of Kakutani*, Israel Journal of Mathematics **24** (1976), 175–189.

[3] D. Heicklen and C. Hoffman, *T, T^{-1} is not standard*, Ergodic Theory and Dynamical Systems **8** (1998), 875–878.

[4] D. Heicklen, C. Hoffman and D. Rudolph, *Entropy and dyadic equivalence of random walks on a random scenery*, Advances in Mathematics **156** (2000), 157–179.

[5] S. A. Kalikow, *T, T^{-1} transformation is not loosely Bernoulli*, Annals of Mathematics **115** (1982), 393–409.

[6] D. S. Ornstein, *Ergodic Theory, Randomness, and Dynamical Systems*, Yale University Press, New Haven, 1974.

[7] W. Philipp and W. Stout, *Almost sure invariance principles for partial sums of weakly dependent random variables*, Memoirs of the American Mathematical Society **161** (1975).

[8] E. Rio, *Sur le théorème de Berry-Esseen pour les suites faiblement dépendantes*, Probability Theory and Related Fields **104** (1996), 255–282.

[9] D. Rudolph, *Asymptotically Brownian skew products give non-loosely Bernoulli K-automorphisms*, Inventiones Mathematicae **91** (1988), 105–128.