STOCHASTIC STABILITY OF BERNOULLI TORAL LINKED TWIST MAPS OF FINITE AND INFINITE ENTROPY

MATTHEW NICOL

ABSTRACT. We construct linked twist maps of the two dimensional torus which are Bernoulli and possess infinite entropy. In particular, we construct a Bernoulli toral linked twist map B^* of infinite entropy which has smooth, absolutely continuous local (un)stable manifolds and positive Lyapunov exponents defined almost everywhere. This map is continuous at each point save those on two line segments.

B^* is shown to be stochastically stable under the following random perturbation: apply the map to a point p and then jump (all points move the same distance and in the same direction) according to a B-process (not necessarily an independent process) such that the expected distance moved is equal to r. Stochastic stability means that given $\alpha > 0$ if $r > 0$ is sufficiently small then the perturbed and unperturbed systems are α-congruent. We prove a similar stability result for B^* under a perturbation in which the random jump described above is distributed according to a general stochastic process. These stability results are also shown to hold (in a slightly modified form) for a general class of finite entropy toral linked twist maps under the same perturbations.

1. INTRODUCTION

1.1. **Discussion and statement of results.** The ergodic properties of toral-linked twist maps (tltm's) have been investigated by several researchers. Burton and Easton [2] established the Bernoulli property for tltm's under fairly mild conditions. Przytycki [14] extends these results and provides a good survey of earlier work.

We consider tltm's of the following form. Let $T^2 = R^2/Z^2$ be the standard torus and let $f : [0,1] \to R^+$, $g : [0,1] \to R^+$ be piecewise, C^2 functions such that, where defined, $f' > c_1$ and $g' > c_2$ for some positive constants c_1, c_2. We allow f and g to have countably many discontinuities. Define $\tau_1(x,y) = (x + f(y), y)$ and $\tau_2(x,y) = (x, y + g(x))$. Let $F(x,y) = \tau_2 \circ \tau_1(x,y)$. F is called a tltm and preserves Lebesgue measure. If f and g have finitely many discontinuities then results of Katok et. al. [6] show that almost every point $p \in T^2$ has a local stable manifold $\lambda^s(p)$ and a local unstable manifold $\lambda^u(p)$ (in this context by a local stable (resp. unstable) manifold through a point p we mean a C^1 curve through p such that there exists a positive constant C and ρ $(0 < \rho < 1)$ such that $d(F^n(p), F^n(q)) < C\rho^n d(p,q)$ for $n \geq 0$ (resp. $n \leq 0$) for all points q contained in the curve, see Section 2 for more details). Furthermore T^2 decomposes into countably many ergodic components, on each of which the restriction of F is either a Bernoulli map or the direct product of a Bernoulli map with a rotation. We shall

1991 *Mathematics Subject Classification.* 58F11; Secondary 58F15.
Key words and phrases. Lyapunov exponent, Linked Twist map.

also require that for almost all points $p, q \in T^2$ and every pair of integers m, n large enough $F^m(\lambda^u(p)) \cap F^{-n}(\lambda^s(q)) \neq \emptyset$. This latter condition implies that the tltm F is Bernoulli [14]. Although we will concentrate on tltm's which are piecewise linear it is easy to see that the proofs we give hold also for the piecewise C^2 case and hence our stability results are valid for the more general class described above.

The main result of Section 3 is the following,

Theorem 1. *There is a Bernoulli toral linked-twist map of infinite entropy which has smooth, absolutely continuous, local (un)stable manifolds and positive Lyapunov exponents defined almost everywhere. This map is continuous everywhere save on two line segments.*

We will show that the presence of infinite entropy ensures an especially strong form of stability. We refer to this map as B^*.

The modification of Pesin theory as developed by Katok et al. [6] does not apply to the analysis of the stochastic properties of B^*. This is because the integral of the log of the absolute value of the Jacobian of B^* does not converge and this convergence is a condition for the applicability of the Katok- Strelcyn theory. For other technical reasons it is also difficult to establish directly from the literature that B^* is Bernoulli.

To overcome this we use certain results from the Bernoulli theory of Ornstein and Weiss [12], especially the notion of a finitely determined process and construct B^* as a limit in the $\bar{d}$-metric of a sequence (B_i) of finite entropy toral linked-twist maps.

Ornstein and Weiss [12] have introduced a notion of stochastic stability based upon the concept of α-congruence. We denote by (f, X, μ, P) a measure-preserving map f of the measure space (X, μ). P is a measurable function from X to a metric space (M, d) (we have left (M, d) out of the notation since it is unwieldy enough already).

Definition 1. Two abstract dynamical systems $(f_1, X_1, \mu_1, P_1), (f_2, X_2, \mu_2, P_2)$ where f_i is a measure preserving transformation of X_i and P_i is a measurable function from X_i to a metric space (M, d) are said to be α-congruent if there is an invertible measure preserving transformation $\phi : X_1 \to X_2$ such that $\phi f_1 = f_2 \phi$ and $d(P_1(p), P_2\phi(p)) < \alpha$ except for a set of points p of μ_1 measure at most α.

Remark 1. This may be phrased as an α-congruence of the *processes* (f_1, P_1) and (f_2, P_2) (see [20]).

The functions P_i are best thought of as measurements or observations on the two systems (taking values in the same metric space (M, d)). The isomorphism ϕ implements a pairing of the orbits of f_1 and f_2 such that observations on corresponding orbits lie within $\sqrt{\alpha}$ of each other except for a set of times of density at most $\sqrt{\alpha}$.

We describe a dynamical system (with observations on the system clear from context) as *stochastically stable* under a certain perturbation if given $\alpha > 0$ then if the magnitude of the perturbation is sufficiently small then the perturbed and unperturbed systems are $\alpha-$congruent.

We consider two types of random perturbations,

Type I) Apply the map to a point p and then jump (all points move the same distance and in the same direction) according to a B-process (not necessarily an independent process, see Section 2) such that the expected value of the distance of the jump is equal to r. The size of the perturbation is measured by r.

Type II) Same as above but now jump according to any stationary stochastic process (not necessarily a B-process) such that the expected value of the distance of the jump is equal to r. The size of this perturbation is measured by r. Note that no assumptions are made on the distribution of the stochastic process constituting the perturbation.

When B^* is perturbed by a type I perturbation the phase space of the perturbed system (see 4.1 for details) is $T^2 \times (T^2)^Z$ with measure $\mu \times m$ (where μ is Lebesgue measure and m is a measure on sequences in $(T^2)^Z$ induced from the B-process of the perturbation). The perturbed system is the skew-product transformation $B^{\#}(p, \sigma) = (B^*(p) + \sigma_0, S\sigma)$, where $(S\sigma)_n = \sigma_{n+1}$ is the shift on sequences. The projection function

$$\Pi : T^2 \times (T^2)^Z \to T^2, \Pi(p, \sigma) = p.$$

is the observation that is made on the torus.

Theorem 2. *B^* is stochastically stable under a type I perturbation.*

Remark 2. Hence there is an isomorphism

$$\Psi : T^2 \to T^2 \times (T^2)^Z$$

which commutes with the respective transformations ($\Psi B^* = B^{\#} \Psi$) and moves by more than a distance α less than a measure α of points. Thus there is a function (which we think of as a deterministic viewer or measuring instrument) $V : T^2 \to T^2$ defined by $V(p) = \Pi\Psi(p)$ which distorts little for most points and the perturbed system can be reproduced exactly by observing the unperturbed system through the viewer (instead of ' reading' p it reads $V(p)$).

We obtain a slightly weaker stability [1] result for B^* under a type II perturbation.

The phase space of the perturbed system is again $T^2 \times (T^2)^Z$ with measure $\mu \times m$ (where μ is Lebesgue measure and m is a measure on sequences in $(T^2)^Z$ induced from the stationary stochastic process governing the perturbation). The perturbed system is the skew-product transformation $B_N^{\#}(p, \sigma) = (B^*(p) + \sigma_0, S\sigma)$, where $(S\sigma)_n = \sigma_{n+1}$ is the corresponding shift on sequences. As before, the projection function

$$\Pi : T^2 \times (T^2)^Z \to T^2, \Pi(p, \sigma) = p.$$

'realises' the perturbed system on the torus.

Theorem 3. *Suppose B^* is perturbed according to a type II perturbation. Given $\alpha > 0$ if $r > 0$ is sufficiently small then the perturbed system is α-congruent to the direct product of the unperturbed system and the 'noise' process of the perturbation. There is an isomorphism*

$$\Psi : (B^* \times S \to (B_N^{\#}, T^2 \times (T^2)^Z),$$

such that $d(p, \Pi\Psi(p, \sigma)) < \alpha$ (for a subset of $T^2 \times (T^2)^Z$ of measure greater than $1 - \alpha$).

---

[1] This form of stability is also called by some authors stochastic stability. We reserve the term stochastic stability for strict α-congruence between the perturbed and unperturbed systems and talk of stability when we have a looser form of α-congruence result.

Remark 3. If we think of the function $V(p,\sigma) = P^{\#}\Psi(p,\sigma)(V : T^2 \times (T^2)^Z \to T^2)$ as a random viewer then the perturbed system can be reproduced exactly by observing the unperturbed system through a random viewer which distorts little most of the time. The randomness of the viewer is a consequence of the fact that it is necessary to know the state of the viewer (given by the noise process $S : (T^2)^Z \to (T^2)^Z$) to determine the exact correspondence between the orbits of the perturbed and unperturbed systems.

We also have stability results for Bernoulli tltm's of finite entropy (satisfying the conditions given above) under perturbations of type I and type II. In this situation the distinction between type I and type II is not important and the result is the same in either case. To be more precise suppose B_f is a tltm satisfing the conditions detailed in Section 1.1. Perturb B_f by a perturbation of type I or type II. The phase space of the perturbed system is $T^2 \times (T^2)^Z$ with product measure $\mu \times m$ where m is the measure induced on the space of sequences of jumps determined by the stationary stochastic process governing the distribution of the jumps. The perturbed system is the skew-product $B_f^{\#} : T^2 \times (T^2)^Z \to T^2 \times (T^2)^Z$.

Theorem 4. *Given $\alpha > 0$ if $r > 0$ is sufficiently small then $(B_f \times S, T^2 \times (T^2)^Z)$ is α-congruent to $(B_f^{\#}, T^2 \times (T^2)^Z)$.*

Remark 4. This has an identical interpretation in terms of a random viewer to that given for the stability of B^* under a random perturbation of type II.

The proofs of these results all have a similar flavor. In essence we show that the perturbed system is Bernoulli (or that the 'noise' factor has a Bernoulli complement) and then establish that if the perturbation is sufficiently small then the perturbed and unperturbed systems are close in the $\bar{d}$ metric. We give the proof of the stochastic stability of B^* under a perturbation of type I in detail and then modify this proof to establish the other stability results.

This work is organised as follows. Section 2 introduces notation and background on the Bernoulli theory. Section 3 details the construction of B^* and Section 4 establishes the stability properties of B^* and toral linked twist maps of finite entropy.

2. Background

For a detailed discussion of the Bernoulli theory and in particular a discussion of the $\bar{d}$ metric we refer the reader to [12], [3], [15] and [10]. For a discussion of the Relative Bernoulli theory we refer the reader to [19] and [12].

First we establish some notation. If $C \subset T^2$ is a set of positive Lebesgue measure then μ_C denotes the conditional measure on C, that is $\mu_C(A) = \frac{\mu(A\cap C)}{\mu(C)}$ for every set $A \subset T^2$. If P is a partition of T^2 and $E \subset T^2$ then $P|E$ denotes the induced partition of the measure space E with measure μ_E. If a property holds for all atoms of a partition P except for a union of atoms which has measure less than ϵ we say that the property holds for ϵ a.e. atom of P. $\mathcal{B}$ denotes the Borel sigma algebra. We denote by $H_S(P|Q)$ the relative entropy conditioned on the set S. This is calculated in the following manner- the set S is normalised to have measure 1 and the entropy of the partition $S \cap P$ relative to $S \cap Q$ is calculated. All finite partitions of the torus that we consider will be assumed to have atoms with piecewise smooth boundaries.

2.1. **Background On The Bernoulli Theory.** Let (f, X, μ) be a dynamical system with P a finite partition of the metric space X. We say that the points x, y have the same $P-N$ name if $f^k(x)$ and $f^k(y)$ lie in the same set of P for all $0 \leq k \leq N$. The partition of X into sets with the same $P-N$ name we denote $\vee_0^N f^{-j}P$. If (f, X, μ) is a dynamical system then we say the partition P of X is a *generator* (or *generates*) if $\vee_{-\infty}^{\infty} B^{-j}P$ is the entire sigma algebra of the system (which in the present context is $\mathcal{B}$). A dynamical system is *Bernoulli* if it has an independent generator i.e. a generating partition P such that f^nP is independent of $\vee_{-n+1}^{n-1} f^{-j}P$ for all n (for equivalent definitions see [12]).

We may model a stationary stochastic process consisting of random variables taking values in a set M as a dynamical system in a standard way as a shift $S : M^Z \to M^Z$, [12, page 48]. If the resulting dynamical system is Bernoulli then the process is called a B-process (note that this is not necessarily an independent process). There is a function $P : M^Z \to M, P(\sigma) = \sigma(0)$ which makes this system concrete (alternatively it represents measurements on the system).

Conversely if X is given the finite partition $P = \{P_1, ..., P_k\}$ then the dynamical system (f, X, μ) produces a stationary stochastic process by inducing a measure on the space $\Omega = \{1, ..., k\}^Z$. The shift map $S\omega_n = \omega_{n-1}$ preserves this measure. When X and μ are clear from context we will denote this process as (f, P). If (f, X, μ) is a Bernoulli dynamical system then for every finite partition P of X, (f, P) is a B-process.

Let f be an invertible measure preserving transformation of a measure space X and let P be a partition of X.

Definition 2. The partition P is a Very Weak Bernoulli (VWB) partition if for every $\epsilon > 0$ there exists N_0 such that for all $N' > N > N_0$ and for all $n \geq 0$ ϵ a.e. atom E of $\vee_N^{N'} f^i P$ satisfies

$$\bar{d}(\{f^{-i}P \mid E\}_0^n, \{f^{-i}P\}_0^n) < \epsilon$$

The process (f, P) is a B-process if and only if the partition P is VWB. The importance of the notion of VWB is clear from the following result,

Proposition 1. [12, page 66]

I *If (f_i, P_i) are B-processes that converge in $\bar{d}$ to (f, P) then (f, P) is a B-process.*

II *If $P_i < P_{i+1}$, $\vee P_i = \mathcal{B}$ and (f, P_i) is a B-process for all i then f is Bernoulli.*

The property of being finitely determined is another characterisation of B-processes.

Definition 3. A process (f, P) is finitely determined (FD) if given $s > 0$ there exists $\delta, N, \gamma > 0$ such that when $(\bar{f}, \bar{P})$ is a stationary process which satisfies

1) (close in finite distribution)

$$\bar{d}(\{f^{-i}P\}_0)^N, \{\bar{f}^{-i}\bar{P}\}_0^N) < \delta$$

2) (close in entropy)

$$h(\bar{f}, \bar{P}) \geq h(f, P) - \gamma$$

then

$$\bar{d}((f, P), (\bar{f}, \bar{P})) < s.$$

A finite partition $P = \{P_1, ..., P_k\}$ of a measure space X can be viewed as a map from the space X to the metric space S of k symbols $S_k = \{1, ..., k\}$ (with metric $d(i,j) = 1$ if $i \neq j$ and zero otherwise) such that $P(x) = j$ if $x \in P_j$.

The α-congruence of the following theorem refers to such induced maps $P : X \to S_k$ and $Q : Y \to S_k$.

Proposition 2. [12, page 104] *If (f, X, μ) and (g, Y, ν) are Bernoulli transformations of equal entropy and P and Q are finite partitions of the same cardinality such that*

$$\bar{d}((f,P),(g,Q)) < \frac{\alpha}{10}$$

then (f,P) is α-congruent to (g,Q).

The meaning of this theorem is that the pairing of orbits to produce the $\bar{d}$ closeness may be implemented by an isomorphism.

Local Stable and Unstable Manifolds.

Let (B, T^2, μ) be an invertible mpt of T^2 with P a finite partition of T^2. Our definition of local stable and unstable manifolds is not standard but will make many of our proofs simpler. The definition of local (un)stable manifolds that we use takes into account the partition of T^2 that we are considering. A *local stable manifold* (lsm) of (B, P) through a point p , $l^s(p)$, is a C^1 smooth curve such that there exists a positive constant C and ρ $(0 < \rho < 1)$ such that $d(B^n(p), B^n(q)) < C\rho^n d(p,q)(\ n \geq 0)$ for all points q contained in the curve and such that any two points of the curve have the same $P - N$ name for all N. The nonstandard part of this definition is the requirement that points on a local stable manifold have the same $P - N$ name for all N. A *local unstable manifold* $l^u(p)$ (lum) of (B, P) through p is a lsm of (B^{-1}, P) through p.

Hyperbolic Blocks

In this paper a *hyperbolic block* H of (B, P) is a measurable set with a distinguished point p (and distinguished lsm $l^s(p)$ and lum $l^u(p)$, both with p as midpoint) constructed in the following way. Let $\tilde{H}$ be an open, connected set in the shape of a four-sided polygon with p as center whose sides are parallel to $l^u(p)$ and $l^s(p)$ and of the same length. The hyperbolic block H is the subset of $\tilde{H}$ such that if $q_1, q_2 \in H$ then the local stable manifold $l^s(q_1)$ intersects the local unstable manifold $l^u(q_2)$ at precisely one point which lies in $\tilde{H}$. Note that $p \in H$ but that there may be points in $\tilde{H}$ whose lsm and lum do not have this intersection property. A hyperbolic block is called *ϵ-good* if $\mu_{\tilde{H}}\{H\} > 1 - \epsilon$.

For the concept of a hyperbolic block to be useful we require absolute continuity of the lsm's and lum's. See [6] and [12] for the exact definition and a discussion of this property.

3. The Construction of B^*

B^* is constructed as a $\bar{d}$ (and pointwise) limit of finite entropy linked twist maps.

We construct a sequence of finite partitions of T^2, $\{P_i\}$, such that $P_i < P_{i+1}$ and $\vee P_i = \mathcal{B}$. We denote the jth finite entropy toral linked twist map of the sequence by B_j. In Section 3.4 we show that each finite entropy process (B_i, P_j) of our sequence is a B-process. The $\bar{d}$ sequence is constructed in the following manner. Construct B_1 so that $h(B_1) \geq 10$ and let P_1 be a partition of T^2 such that $h(B_1, P_1) > 5$. In general, given (B_k, P_k) we construct B_{k+1} so that

†

a) $h(B_{k+1}) \geq 10(k+1)$
b) $\bar{d}((B_k, P_k), (B_{k+1}, P_k)) < \frac{1}{2^k}$
c) $h(B_{k+1}, P_j) > h(B_k, P_j) - \frac{1}{2^k}$ for all $j \leq k$

We then choose a partition P_{k+1}, such that $P_k < P_{k+1}$ and

d) $h(B_{k+1}, P_{k+1}) > 5k$.

The sequence of partitions may be chosen so that $\vee P_j = \mathcal{B}$.

Now fix P_1. We have constructed the sequence (B_i) so that

$$\bar{d}((B_m, P_1), (B_{m+k}, P_1)) < 2^{-m} \text{ for all } k > 0$$

Note that the sequence (B_i) converges pointwise. In Section 3.5 we show that (B^*, P_j) is the $\bar{d}$ limit of (B_k, P_j) for each j and hence by Proposition 1 (B^*, P_j) is a B-process for all j and B^* is Bernoulli. The properties c) and d) which the maps and partitions have ensure that $h(B^*) = \infty$. This is also shown in Section 3.5.

The delicate part of the construction lies in showing that it is possible to choose (B_{i+1}) so that it has measure theoretic entropy greater than $10i$ yet ensure (B_{i+1}, P_i) and (B_i, P_i) are close in the $\bar{d}$ metric (to within 2^{-i}).

3.1. **Description of the Finite Entropy Toral Linked Twist Maps.** This section describes in more detail the maps in the sequence of finite entropy toral linked twist maps used to construct B^*.

Let $\{l_i\}_{i\geq 1}$ be a sequence of points in the unit interval such that $l_{i+1} > l_i$ and $|1 - l_i| < \frac{1}{2^i}$. We define a sequence $\{f_i\}$ of positive, piecewise linear real-valued functions $f_i : [0,1] \to R^+$ satisfying the following conditions ,

a) $10 \leq f_1' \leq f_2' \leq ... \leq f_n' ...$
b) $f_i \equiv f_{i-1}$ except on the interval $(l_i, 1]$.
c) f_i is smooth except at the points $\{0\}$ and $\{l_k\}$, $k \leq i$.

Define the twist maps $\tau_{1i} : T^2 \to T^2$ and $\tau_{2i} : T^2 \to T^2$ by

$$\tau_{1i}(x, y) = (x + f_i(y), y)$$

and

$$\tau_{2i}(x, y) = (x, y + f_i(x))$$

Finally B_i is defined as the composition $\tau_{2i} \circ \tau_{1i}(x, y)$. B_i preserves Lebesgue measure.

We note here that the functions f_i may be chosen so that B_i is continuous for all i (and hence B^* will be continuous except on the union of two lines). The requirement that $f_i' > 10$ is not necessary and any positive lower bound will do.

To simplify notation we will sometimes write τ_2 and τ_1, omitting the i subscript when there is no danger of ambiguity.

At this point we avoid assigning specific values to the numbers l_i and describing the maps f_i beyond the properties already detailed. The specifications given entail that B_i is Bernoulli for each i. This is proved in Section 2.4. In Section 2.2 we show that it is possible to choose the maps (f_i), the values (l_i) and the partitions P_i so as to ensure that conditions a), b) , c) and d) of † hold.

3.2. **Construction of the $\bar{d}$ Cauchy Sequence.** In this section and the next we detail how the parameters l_{i+1}, f_{i+1} and the partitions P_{i+1} are to be chosen for the tltm B_{i+1} (assuming inductively that B_j has been defined for $j \leq i$). We need to ensure conditions a),b) c) and d) † hold.

The difficulty in the construction lies in showing that it is possible to satisfy a) and b) with a single choice of the parameters l_{i+1} and f_{i+1}. It is as a consequence of the proof of a) and b) that we show that it is possible to construct the sequence so that c) and d) hold as well.

First we discuss the metric entropy of B_i and give a lower bound for $h(B_i)$ in terms of f_i and l_i. Recall that $h(B_i)$ is equal to the integral of the positive Lyapunov exponents over T^2 with respect to Lebesgue measure.

Let $Q_{ii} = \{(x,y) : (x,y) \in (0,1] \times (l_i,1)\})$ and $\tau_{1i}(x,y) \in (l_i,1)) \times (0,1]\}$.

As $f_i' > 10$ for all i, the linear transformation $D(B_i)_{(x,y)}$ preserves the cone $\mathcal{C}$ of vectors consisting of the upper half of the first quadrant (vectors in the first quadrant with slope greater than 1). The maximum eigenvalue of $D(B_{x,y})$ is given by

$$\Lambda_i = \frac{(2+\gamma_{ii}) + \sqrt{\gamma_{ii}^2 + 4\gamma_i}}{2}$$

where $\gamma_{ii} = (f_i'|_{(l_i,1)})^2$.

The corresponding eigenvector ν_{ii} lies in the cone $\mathcal{C}$ for all iterates of DB_i. The eigenvector corresponding to the minium eigenvalue of DB_i lies in the second quadrant.

Since B_i is ergodic the orbit of almost every point will lie in Q_{ii}, $\mu(Q_{ii})$ percent of the time. Hence the positive Lyapunov exponent will be at least $\frac{\mu(Q_{ii})\log\Lambda_{ii}}{2}$ by the ergodic theorem. Thus given $\mu(Q_{ii})$ we may choose $f_i'|_{(l_i,1)}$ sufficiently large to ensure that $h(B_i) > 10i$ (in fact we choose $f_i'|_{(l_i,1)}$ so that $\frac{\mu(Q_{ii})\log\Lambda_{ii}}{2} > 10i$).

We then choose a partition P_{i+1} so that $P_i < P_{i+1}$ and $h(B_{i+1}, P_{i+1}) > 5(i+1)$. To achieve b) we use the fact that B-processes are finitely determined. Since B_i is Bernoulli, (B_i, P_k) is a finitely determined process for all k. As a consequence of the proof of a) it is easily seen that we may take conditions c) and d) to also hold.

Apply Definition 3 to (B_i, P_i) with $s = 2^{-i}$ to obtain $T > 0, \gamma > 0$ and $\delta > 0$ such that if $h(B_{i+1}, P_i) > h(B_i, P_i) - \gamma$ and $\bar{d}(\{B_{i+1}^{-j} P_i\}_0^T, \{B_i^{-j} P_i\}_0^T) < \delta$ then $\bar{d}((B_{i+1}, P_i), (B_i, P_i)) < 2^{-i}$. Note that without loss of generality we may take $\gamma < \frac{1}{2^i}$.

The main technical difficulty is to ensure that condition c) holds. To do this we use an argument involving hyperbolic blocks. For the remainder of this chapter $\gamma(\gamma < \frac{1}{2^i}), \delta$, and T will refer to the corresponding parameters obtained by taking (B_i, P_i) for (f, P) and $s = 2^{-i}$ in Definition 3. We define an *admissible perturbation* of B_i to be a map produced by a choice of the parameters l_{i+1} and f_{i+1}' subject to the conditions specified earlier, namely $l_{i+1} > l_i$ and $f_{i+1}' \geq f_i'$. To allow us to construct lsm's and lum's for (B^*, P_i) we also require that l_{i+1} be chosen so that $|1 - l_{i+1}| \leq \frac{1}{2^{i+1}}$.

We now state a lemma which will be proved in Section 3.3.

Lemma 1. *Let P_i be a finite partition of T^2. If Φ is a tltm of T^2 consisting of either B_i or an admissible perturbation of B_i then given $\delta', \eta > 0$ there exists $\beta > 0$ such that we may cover T^2 disjointly up to measure $1 - \delta'$ with hyperbolic blocks of Φ all of which have measure greater than β.*

The hyperbolic blocks H have the following property:

- $$\frac{1}{k}|H_H(\vee_0^k \Phi^{-j} P_i | \vee_1^\infty \Phi^j P_i) - H_H(\vee_0^k \Phi^{-j} P_i))| \leq \eta$$

for all $k > 0$

Remark 5. The conclusions of this lemma hold for any tltm's satisfying the conditions stipulated in the introduction. We will prove these properties only for the case where the tltm's are piecewise linear since the proofs are more transparent and we may choose whatever tltm's we wish in the construction of B^*. It is easily seen that the arguments generalise to the general class of tltm's described in the introduction.

The content of this lemma is that the properties of our cover (up to measure $1-\delta'$) hold uniformly over all possible admissible perturbations.

In this section we will slightly abuse notation and use both the notation B_{i+1} and Φ to denote an admissible perturbation of B_i.

The main idea is that we alter B_i (to produce B_{i+1}) on a set of sufficiently small measure so that the finite distributions of B_i and B_{i+1} up to a time T are close on the hyperbolic blocks of B_{i+1} and the size of the atoms in the induced partition of $\vee_0^T B_i^{-j} P_i$ on these hyperbolic blocks reflects the entropy of (B_i, P_i) (a consequence of the Shannon-McMillan-Breiman Theorem).

When we condition on the past on a hyperbolic block of B_{i+1} the relative entropy of the P_i-names of B_{i+1} till time T does not decline much- hence the entropy of (B_{i+1}, P_i) is close to that of (B_i, P_i).

In Definition 3 we take $\delta' < \delta$ such that $\mid \delta' \log \delta' \mid < \frac{\gamma}{10}$, $\delta' \log(\# P_i) < \frac{\gamma}{10}$ (this last requirement is because of the incompleteness of the cover by hyperbolic blocks and looks forward to Lemma 3) and $\eta = \frac{\gamma}{10}$ to obtain a cover with the properties described in Lemma 1. Recall that the hyperbolic blocks in the cover have measure bounded below by $\beta > 0$ (a function of δ' and η) and this bound is uniform over admissible perturbations.

Lemma 2. *Given $0 < \beta < 1$ there exists $T' > T$ such that*

1) $|H_S(\vee_0^{T"} B_i^{-j}) - T"(h(B_i, P_i))| \leq \frac{\gamma}{10}$ *for all $T" > T'$ where S is any set of measure $\geq \beta$.*

This lemma is an easy consequence of the Shannon-McMillan-Breiman Theorem and states that after a certain time T' the entropy of B_i can be estimated from the relative entropy of any set of measure greater than β.

Remark 6. Note that this lemma refers only to B_i.

We need one more lemma.

Lemma 3. *If B_{i+1} is an admissible perturbation of B_i then,*

$$\begin{aligned} &H(\vee_0^{T'} B_{i+1}^{-j} P_i | \vee_1^{\infty} B_{i+1}^{j} P_i) \\ &\geq \sum_j \mu(H_j) H_{H_j}(\vee_0^{T'} B_{i+1}^{-j} P_i | \vee_1^{\infty} B_{i+1}^{j} P_i) \end{aligned}$$

where $\cup H_j, j \neq 0$ is the cover of hyperbolic blocks of B_{i+1} and H_0 ($\mu(H_0) < \delta'$) is the complement of $\cup H_j, j \neq 0$.

The proof of this lemma is routine. It enables us to get a lower estimate for the entropy of B_{i+1} in terms of its relative entropy on hyperbolic blocks.

We now string Lemmas 1, 2 and 3 together to show how to construct B_{i+1}. Recall that we have $\gamma, (\gamma < \frac{1}{2^i}), \delta$, and T from Definition 3 applied to B_i with

$s = 2^{-i}$. In Lemma 1 we then take δ' and η so that $\delta' < \delta, |\ \delta' \log \delta'\ | < \frac{\gamma}{10}$, $\delta' \log(\# P_i) < \frac{\gamma}{10}$ (this requirement is necessary because of the incompleteness of the cover in Lemma 3) and $\eta = \frac{\gamma}{10}$. This choice of δ' and η defines a minimal measure β for the hyperbolic blocks in the cover as described in Lemma 1. For this β we then have a T' as defined in Lemma 2.

We need to choose l_{i+1} and f_{i+1} in order to define B_{i+1}. First choose l'_{i+1} so that $\mu(l_{i+1}, 1]$ is small enough to ensure

a) $|H_S(\vee_0^{T'} B_{i+1}^{-j} P_i) - H_S(\vee_0^{T'} B_i^{-j} P_i)| < \frac{\gamma}{10}$ for any set S of measure $> \beta$
b) $\bar{d}(\{B_{i+1}^{-j} P_i\}_0^{T'}, \{B_i^{-j} P_i\}_0^{T'}) < \delta$.

Note that condition b) is condition 2 of Definition 3.
Then we have,

$$\begin{aligned}
T' h(B_{i+1}, P_i) &= H(\vee_0^{T'} B_{i+1}^{-j} P_i |\ \vee_1^{\infty} B_{i+1}^{j} P_i) \\
&\geq (1-\delta') \mathrm{inf}_{j>0} H_{H_j}(\vee_0^{T'} B_{i+1}^{-j} P_i |\ \vee_1^{\infty} B_{i+1}^{-j} P_i) \text{ (by Lemma 3)} \\
&\geq \mathrm{inf}_{j>0} H_{H_j}(\vee_0^{T'} B_i^{-j} P_i) - \frac{2T'\gamma}{10} \text{ (by Lemma 1))} \\
&\geq T'(h(B_i, P_i) - \frac{2\gamma}{10}) - \frac{2T'\gamma}{10} \text{ (by Lemma 2).}
\end{aligned}$$

Therefore $h(B_{i+1}, P_i) \geq h(B_i, P_i) - \gamma$ and condition 1 of Definition 3 holds.

We note here that the argument given above applies to each partition $P_j, j < i$ so that we obtain a number l_{i+1}^j such that $h(B_{i+1}, P_j) > h(B_i, P_j) - \frac{1}{2^i}$ if $\mu(l_{i+1}^j, 1]$ is small enough. Take l_{i+1} to be $\max(\ l_{i+1}^j (j < i), l'_{i+1})$.

Thus , with this choice of l_{i+1}, $\bar{d}((B_{i+1}, P_i), (B_i, P_i)) < 2^{-i}$ and $h(B_{i+1}, P_j) > h(B_i, P_j) - \frac{1}{2^i}$, for all $j \leq i$.

We have seen that given the value of l_{i+1} chosen above we may choose f_{i+1} to ensure $h(B_{i+1}) \geq 10(i+1)$.

Now we choose a finite partition P_{i+1} so that $P_i < P_{i+1}$ and $h(B_{i+1}, P_{i+1}) > 5(i+1)$.

Therefore we have constructed a sequence of finite entropy toral linked twist maps (B_i) and finite partitions P_i such that $P_i < P_{i+1}$, $\vee P_i = \mathcal{B}$ and conditions a), b), c) and d) of † hold.

□

3.3. **Proof of Lemma 1.** First fix a choice of l_{i+1} and f_{i+1}. B_{i+1} so defined is invertible almost everywhere.

Note that the cone O in the second quadrant defined by the x-axis and the line y= $-$x is preserved under $D({B^{i+1}}^{-1})$. Furthermore a calculation shows that every vector in O expands in length by a factor greater than 10 under $D((B_{i+1})^{-1}))$. Since we have a minimum expansion rate in the invariant cone O any sector of defining angle θ in O will be sent under $D(B_{i+1})^{-1}$ to a sector of defining angle θ' where $|\frac{\theta'}{\theta}| < \zeta < 1$ for some $\zeta > 0$.

We now construct lsm's. Let p be a point on T^2. Let $O(p_k)$ denote the cone O in the tangent space of the point $p_k \equiv B_{i+1}^k(p)$. Let $O(k) = B_{i+1}^{-k}(O(p_k))$. Then $O(k)$ is a sector in the tangent space of p and $O(k+1) \subset O(k)$ strictly.

Since we have a uniform decrease in the defining angle of the sector $O(k)$ with k, $\bigcap_k O(k)$ defines a vector $w(p)$ in the tangent space of p. Note that $w(p)$ lies in the cone $O(p)$. A vector based at p in the direction $w(p)$ will contract uniformly under B_{i+1}^{k+1} for all k.

Let R be the partition formed by the join of the partition P_i with the partition into the sets $A_{j,k}$ where $A_{j,k} \equiv [l_j, l_{j+1}) \times [l_k, l_{k+1}), \quad (j, k \leq i)$.

For $p \in T^2/\partial R$ take a linear segment of length l' (a candidate lsm) centered on p in the direction of the vector $w(p)$. We consider such a lsm l_p "cut" if at some time $k \geq 0$ we have

$$B_{i+1}^k(l_p) \bigcap \partial R \neq \emptyset$$

or

$$\tau_1 B_{i+1}^{k-1}(l_p) \bigcap \partial R \neq \emptyset.$$

Until l_p is cut it contracts at a rate greater than μ (where $\mu \leq \frac{1}{10}$), so that

$$|B_{i+1}^k(l_p)| \leq \mu^k |l_p|$$

and

$$|\tau_1 B_{i+1}^k(l_p)| \leq \mu^{k-1} |l_p|$$

for all k till l_p is cut.

Now ∂R consists of finitely many line segments, say $N(i)$. The candidate lsm l_p will be cut at time k if the midpoint of $B_{i+1}^k(l_p)$ is within $\frac{\mu^k l'}{2}$ of one of these line segments at time k or if the midpoint of $\tau_1 B_{i+1}^{k-1}(l_p)$ is within $\frac{\mu^{k-1} l'}{2}$.

Thus the measure of points p such that l_p is cut at time k is bounded above by

$$N(i)\mu^{k-1} l'$$

So the measure of points whose candidate lsm's of length l' are eventually cut is bounded above by

$$\sum_{k=1}^{\infty} N(i) l' \mu^k = \frac{N(i) l'}{1-\mu}.$$

Therefore given δ' there exists l' such that up to a set of measure $\frac{\delta'}{2}$, every point p is the midpoint of a lsm (by abuse of notation we neglect the dependence on length and denote this $l^s(p)$) of length l' through it and any two points on this fiber have the same future name (with respect to the partition P_i) for all time. This is true of all admissible perturbations.

An analogous argument shows that up to a set of measure $\frac{\delta'}{2}$, every point p is the midpoint of a lum (which we denote $l^u(p)$) of length l'' through it and any two points on this fiber have the same future name (with respect to the partition P_i) for all time. Any two points on the lum fiber have the same past name (with respect to P_i) for all time. Thus if we take l to be the minimum of l' and l'' the claim is established.

The lsm's lie in the cone O and the lum's lie in the cone $\mathcal{C}$. Thus the angle between any lsm and any lum is greater than $\frac{\pi}{4}$.

This foliation enables us to construct hyperbolic blocks with the properties described in Lemma 1. To do this we use the following three sublemmas.

Sublemma 1. *Let Φ be an admissible perturbation of B_i. Given $\epsilon, \delta > 0$ there exists $l > 0$ such that except for a set of measure at most 2δ each point p satisfies the following conditions:*

a) *p is the midpoint of a lsm $l^s(p)$ and a lum $l^u(p)$ both of length greater than $2l$.*
b) *Let $N_l(p)$ be the neighborhood of p defined by an open polygon of sidelength $2l$ centred on p (with sides parallel to $l^s(p)$ and $l^u(p)$). Then $\mu_{N_l(p)}\{q : |l^c(q)| < 4l$ or $|l^e(q)| < 4l\} < \epsilon$.*

Proof of Sublemma 1:

The truth of condition a) is clear. Suppose b) is is false. Then there exist $\epsilon, \delta > 0$ such that for all $l > 0$ we have a set $B(l)$ of points and each point in $B(l))$ has a neighborhood $N_l(p)$ such that $\mu_{N_l(p)}\{q : |l^c(q)| < 4l$ or $|l^e(q)| < 4l\} > \epsilon$. Thus the set of points q such that $|l^c(q)| < 4l$ or $|l^e(q)| < 4l$ has measure greater than $\epsilon\delta$. This contradicts the fact that the measure of points which are the midpoints of lcm's and lum's of length greater than $4l$ approaches 1 as l approaches 0 uniformly over every admissible perturbation.

□

A neighborhood $N_l(p)$ of a point p satisfying the conditions of the sublemma defines an ϵ-good hyperbolic block in an obvious way.

Sublemma 2. *Given $\eta > 0$ there exist $l, \eta' > 0$ such that if B_{i+1} is an admissible perturbation of B_i, then for all $k > 0$*

$$\frac{1}{k} \mid H_H(\vee_0^k B_{i+1}^{-j} P_i \mid \vee_1^{\infty} B_{i+1}^{j} P_i) - H_H(\vee_0^k B_{i+1}^{-j} P_i) \mid < \eta$$

for any η'-good hyperbolic block H of B_{i+1} which is composed of fibers of length less than l.

Proof of Sublemma 2:

Observe that since a hyperbolic block H is composed of lsm's and lum's of length at least l which traverse from one side of the hyperbolic block to the other then as l approaches zero the distribution of an atom in the induced partition of $\vee_0^k B_{i+1}^{-j} P_i$ on H on atoms in the induced partition of $\vee_1^{\infty} B_{i+1}^{-j} P_i$ on H approaches its distribution on the whole of H. This is because atoms of the induced partition of $\vee_0^k B_{i+1}^{j} P_i$ on H are composed of lum's and atoms of the induced partition of $\vee_1^{\infty} B_{i+1}^{j} P_i$ are composed of lsm's. The lum's in H are approximately parallel straight line segments which become more nearly parallel uniformly over such hyperbolic blocks as $l \to 0$. Recall that we may cover the torus arbitrarily close to full measure with lum's of a minimal length l' and then break these lum's into shorter ones to form hyperbolic blocks. Since the original fibers lay in the cone $\mathcal{O}$ and did not intersect for a length greater than l' if we take the lum's constituting the hyperbolic blocks to be of much shorter length l then expanding fibers of our hyperbolic blocks are more nearly parallel uniformly over such hyperbolic blocks as $l \to 0$. Lsm's are by an analogous argument more nearly parallel uniformly over hyperbolic blocks as $l \to 0$. Thus uniformly over k and uniformly over the hyperbolic blocks in the cover the distribution of an atom in the induced partition of $\vee_0^k B_{i+1}^{j}$ on a hyperbolic block H on each atom of the induced partition of $\vee_1^{\infty} B_{i+1}^{-j} P_i$ on H approximates its distribution on the whole of the hyperbolic block.

We now examine the consequence of this for the difference between the relative entropy of a hyperbolic block conditioned on the past and the unconditioned relative entropy.

Let (P_α) denote atoms of $\vee_0^k B_{i+1}^{-j} P_i$ and (Q_β) denote atoms of $\vee_1^t B_{i+1}^{j} P_i$, where t is finite (for concreteness we condition on the finite past, our estimates will not depend on t and hence our results will hold when we condition on the infinite past).

We have seen that given $\epsilon > 0$ there exists a l (independent of t) such that if H is a hyperbolic block composed entirely of lsm's and lum's of length less than l which extend across H then

$$(1-\epsilon)\frac{\mu(P_\alpha \cap H)}{\mu(H)} \leq \frac{\mu(P_\alpha \cap Q_\beta \cap H)}{\mu(Q_\beta \cap H)} \leq (1+\epsilon)\frac{\mu(P_\alpha \cap H)}{\mu(H)}$$

A calculation shows

$$|H_H(\vee_0^k B_{i+1}^{-j} P_i \mid \vee_1^t B_{i+1}^{j} P_i) - H_H(\vee_0^k B_{i+1}^{-j} P_i)|$$
$$\leq \epsilon k \log(\# P_i) + \mid (1+\epsilon)\log(1-\epsilon) \mid$$

Therefore given η if we take ϵ small enough we see that there exists an l such that if a hyperbolic block H is composed solely of lsm's and lum's of length less than l then

‡i)

$$\frac{1}{k} \mid H_H(\vee_0^k B_{i+1}^{-j} P_i \mid \vee_1^t B_{i+1}^{j} P_i) - H_H(\vee_0^k B_{i+1}^{-j} P_i) \mid < \eta$$

Since l was independent of t , for this value of l we also have

‡ii)

$$\frac{1}{k} \mid H_H(\vee_0^k B_{i+1}^{-j} P_i \mid \vee_1^\infty B_{i+1}^{j} P_i) - H_H(\vee_0^k B_{i+1}^{-j} P_i) \mid < \eta$$

for all such hyperbolic blocks.

□

Sublemma 3. *If Φ is an admissible perturbation then given $\delta, \eta > 0$ there exist disjoint hyperbolic blocks $\{H_i\}$ such that $\mu(\cup_i H_i) > 1 - 2\delta$ and for each hyperbolic block H_i*

$$\frac{1}{k} \mid H_H(\vee_0^k \Phi^{-j} P_i \mid \vee_1^\infty \Phi^{j} P_i) - H_H(\vee_0^\infty \Phi^{-j} P_i) \mid \leq \eta \textit{ for all } k > 0.$$

Proof of Sublemma 3:

We have already observed (Sublemma 1) that there exist $l, \eta' > 0$ such that condition ‡ holds uniformly over η'-good hyperbolic blocks of admissible perturbations if the lsm's and lum's of the hyperbolic blocks are of length less than l'.

Note that if a hyperbolic block is $\frac{\eta'}{n^2}$- good we can disjointify it into n^2 smaller blocks of approximately equal measure, each of which is at least η'- good.

Cover the torus disjointly (up to full measure) with squares of equal sidelength w, and hence area w^2. There are $\frac{1}{w^2}$ such squares in the cover.

If S is a square of measure w^2 and it is covered by the polygon $\tilde{H}$ defined by a hyperbolic block H which is $\frac{\eta'}{n^2}$- good then if the hyperbolic block is disjointified into n^2 hyperbolic blocks of approximately equal measure which are η'- good then those blocks contained within S exhaust S (up to a set of measure at most η') as n increases. The relative measure of S covered by these disjoint hyperbolic blocks approaches $1 - \eta'$ as n increases and this depends on the alignment of the hyperbolic blocks with the square. Since we have uniform bounds (over admissible perturbations) on the alignment of lum's and lsm's we see that the increase in the relative measure of such a square covered by disjoint hyperbolic blocks of our cover is uniform in n (over admissible perturbations and the squares in our cover).

Hence given $\eta' > 0$ we may choose n large enough so that if a square is covered by a $\frac{\eta'}{n^2}$-good hyperbolic block if the block is disjointified into n^2 blocks those that lie within the square are η'-good and their union has relative measure in the square of $1-\eta'$.

Given $\delta, \eta' > 0$, $(\eta' < \delta)$ for any integer n there exists $l > 0$ (and we may assume $l < l'$) so that up to measure $1-\delta$ each point p lies in an $N_{l'}(p)$ neighborhood which defines a hyperbolic block which is $\frac{\eta'}{n^2}$-good. Cover the torus with disjoint squares of sidelength w, where say, $w < 5l$. The measure of those squares which do not contain a point p with such a $N_{l'}(p)$ neighborhood is at most δ. If a square in our cover does contain a point with such a neighborhood then there is an $N_{l'}(p)$ neighborhood which covers the square and this neighborhood defines a $\frac{\eta'}{n^2}$-good hyperbolic block H. Disjointify H into n^2 hyperbolic blocks of roughly the same measure. Each resulting hyperbolic block is η'-good. Those that lie within the square are disjoint, $\eta'-$ good and have relative measure in the square of greater than $1-\eta'$. Since this holds for each square in our cover the set of points covered by disjoint hyperbolic blocks which are $\eta'-$ good and is greater than $1-2\delta$.

Thus we have ,uniformly over admissible perturbations, a disjoint cover (up to measure $1-\delta$) of η'-good hyperbolic blocks.

Since these hyperbolic blocks are η'-good and have lsm's and lum's of length smaller than l we see that property ‡ holds for each hyperbolic block.

□

The three sublemma's given above establish Lemma 1.

3.4. **Bernoulliness of the Finite Entropy Tltm's B_i.** We now show that each tltm satisfying the template conditions of Section 3.1 is Bernoulli (note that we don't require $\mid 1-l_i \mid < \frac{1}{2^i}$). Fix a map B_j satisfying the template conditions. First we note that B_j satisfies Katok-Strelcyn conditions [6] and hence has countably many ergodic components, on each of which the map is Bernoulli or Bernoulli times a rotation. To establish the Bernoulli property we need only show that B_j is ergodic and rule out the possibility of Bernoulli times a rotation. To prove ergodicity we use a criterion given by Przytycki [14] for systems satisfying Katok-Strelcyn conditions: it is enough to show that for almost every $p, q \in T^2$ $B_j^m(l^u(p))$ intersects $B_j^{-n}(l^s(p))$ for all m, n sufficiently large.

Take a lum $l^u(p)$ through the point p. Consider the partition A into the sets $A_{s,k}$ defined by the lines where the rotation rates change. By construction $l^u(p)$ does not intersect ∂A. If $\tau_1(l^u(p))$ cuts a set $A_{s,k}$ horizontally then $B_j(l^u(p))$ will cut T^2 vertically (as $f'_k(l_{k+1}-l_k) > 1$) by construction. So suppose $\tau_1(l^u(p))$ does not cut an atom $A_{s,k}$ horizontally. Then $\tau_1(l^u(p))$ can cut at most one line of form $(l_s) \times [0,1)$ (if it cuts 2 or more it cuts horizontally). If it cuts one such line then this line divides $\tau_1(l^u(p))$ into two segments (neither of which cuts a $A_{s,k}$ horizontally by assumption , else we are done). Let the longer of the two segments be denoted $l^{e\prime}$. Hence $l^{e\prime}$ will have length greater than $\frac{10|l^u(p)|}{2\sqrt{2}}$ since its length will be at least 10 times the projection of the length of $l^u(p)$ in the vertical direction, and $l^u(p)$ lies in the cone of expanding vectors. $\tau_2 l^{e\prime}$ either cuts a $A_{s,k}$ vertically (in which case we are done as then $\tau_1\tau_2 l^{e\prime}$ will cut horizontally) or cuts at most one line of the form $[0,1) \times (l_s)$. If it cuts one such line then it is divided into two segments by this line and we take the longer to get $l^{e''}$. $l^{e''}$ lies in the cone of lum's and has length greater than $2|l^u(p)|$. If we repeat the process we either get a horizontal crossing or a segment increasing in length by a fixed factor each time (and this will give a horizontal crossing eventually).

Thus we have satisfied Przytcki's condition for ergodicity and B_j is Bernoulli or Bernoulli times a rotation. We may rule out the second possibility by noting that the Przyctcki condition is satisfied for any power of B_j. Hence any power of B_j is ergodic and so B_j cannot be Bernoulli times a rotation. Thus B_j is Bernoulli. □

3.5. **$(B^*, \mathcal{B})$ is Bernoulli and $h(B^*) = \infty$.** We show that (B^*, P_j) is the $\bar{d}$ limit of (B_k, P_j) for each j and hence, by Proposition 1 , (B^*, P_j) is a B-process for each j. Since $\vee P_j = \mathcal{B}$, B^* is Bernoulli.

Lemma 4. *$\bar{d}((B^*, P_j), (B_k, P_j)) \leq \frac{1}{2^k}$ for all $k \geq j$.*

Proof

It suffices to show

$$\bar{d}((B^*, P_k), (B_k, P_k)) \leq \frac{1}{2^k} \text{ for all } k$$

as then

$$\bar{d}((B^*, P_j), (B_k, P_j)) \leq \frac{1}{2^k} \text{ if } j \leq k \text{ (since } P_k \text{ refines } P_j)$$

Now fix P_k. Given s there exists n_s greater than k such that if $r \geq n_s$ we have

$$\bar{d}(\{B^{*-i}P_k\}_0^s, \{B_r^{-i}P_k\}_0^s) \leq \frac{1}{2^s}$$

Since $\bar{d}$ is a metric, we have

$$\bar{d}(\{B^{*-i}P_k\}_0^s, \{B_k^{-i}P_k\}_0^s) \leq \bar{d}(\{B^{*-i}P_k\}_0^s, \{B_{n_s}^{-i}P_k\}_0^s) + \bar{d}(\{B_k^{-i}P_k\}_0^s, \{B_{n_s}^{-i}P_k\}_0^s)$$

By construction of the sequence,

$$\bar{d}(\{B_k^{-i}P_k\}_0^s, \{B_{n_s}^{-i}P_k\}_0^s) \leq \frac{1}{2^k}.$$

By definition of n_s,

$$\bar{d}(\{B^{*-i}P_k\}_0^s, \{B_{n_s}^{-i}P_k\}_0^s) \leq \frac{1}{2^s}$$

Thus,

$$\bar{d}(\{B^{*-i}P_k\}_0^s, \{B_k^{-i}P_k\}_0^s) \leq \frac{1}{2^s} + \frac{1}{2^k}$$

Since $\lim_{s\to\infty} \bar{d}(\{B^{*-i}P_k\}_0^s, \{B_k^{-i}P_k\}_0^s) = \bar{d}((B^*, P_k), (B_k, P_k)$ this proves the lemma and the result follows. □

We now show that $h(B^*) = \infty$. To do this we show that given $m > 0$ there exists a P_{r_m} such that $h(B^*, P_{r_m}) \geq m$.

First we note that by construction $h(B_k, P_j) \geq h(B_{k-1}, P_j) - \frac{1}{2^{k-1}}$ for all $k \geq j$ and $h(B_j, P_j) > 5j$.

Let $m > 0$. Given $k > 0$ there exists a B_{i_k}, with $i_k > k$, such that

$$\frac{1}{k} \mid H(\vee_0^{k-1} B^{*-j} P_m) - H(\vee_0^{k-1} B_{i_k}^{-j})| \leq 1$$

But then we have

$\frac{1}{k}H(\vee_0^{k-1} B^{*-j} P_m) \geq \frac{1}{k}H(\vee_0^{k-1} B_{i_k}^{-j} P_m) - 1$

$\geq h(B_{i_k}, P_m) - 1$
$\geq h(B_{i_k-1}, P_m) - 1 - \frac{1}{2^{i_k-1}}$
$\geq h(B_m, P_m) - 2.$
Since k is arbitrary the claim is proven.

□

3.6. B^* **has Positive Lyapunov Exponents and Smooth Fibers.** The proof of the existence of positive Lyapunov exponents proceeds exactly the same way as that for B_i, a finite entropy tltm. We will not repeat it here. Similarly exactly the same arguments used in the case of the finite entropy tltm's show that under B^* almost every point p has a contracting direction $w(p)$. Vectors in this direction contract by more than the factor μ ($\mu \leq \frac{1}{10}$), under iterates of the derivative of B^*.

The construction of lsm's and lum's for B^* is, however, more difficult than the construction of lsm's and lum's for the maps B_i of the sequence. This is because there are infinitely many lines bounding the regions of different twist rates in the case of B^* and each of these lines may cut a lsm (lum).

Given a partition P_j we show that (B^*, P_j) has lsm's and lum's through almost every point.

For almost all points p we have a contracting direction $w(p)$ in the tangent space at p. Through each such point p place a line segment of length l, centered on the point p, in the contracting direction.

We say the segment $l(p)$ is "cut" at time m if either

- $$B^{*m}(l(p)) \cap \partial(A \vee P_j) \neq \emptyset \text{ or } \tau_1 B^{*m-1}(l(p)) \cap \partial(A \vee P_j) \neq \emptyset$$
- or

$B^{*m}(l(p))$ or $\tau_1 B^{*m-1}(l(p))$ is within $l\mu^m$ of the singularity lines$(0,1]\times(1),(1)\times(0,1]$

Let $N(m,l)$ be the number of line segments in ∂A (i.e. lines of form $(l_j) \times (0,1]$ or $(0,1] \times (l_i)$) which are further than $\mu^m l$ from the singularity lines. Let C be the number of line segments in ∂P_j.

The measure of points p whose line segments $l(p)$ are cut at time m is therefore bounded above by

$$2CN(m,l)\mu^m l + 2C\mu^m l.$$

Let D_m be the set of points p whose line segments $l(p)$ are cut at time m. Then the set of points whose line segment is eventually cut is contained in the set $\cup_0^\infty D_m$.

Thus the measure of points p whose line segments $l(p)$ are eventually cut is less than or equal to

$$\sum_0^\infty m(D_m) = 2C\frac{l}{1-l} + 2Cl\sum_0^\infty N(m,l)\mu^m.$$

Since we have stipulated that the sequence (l_i) be chosen so that $|1-l_i| \leq \frac{1}{2^i}$, the number of boundary lines further than δ from the singularity lines is less than $\frac{1}{\sqrt{\delta}}$. Hence $N(m,l)$ is bounded above by $\frac{1}{\sqrt{\mu^m l}}$.

Therefore

$$\sum_0^\infty m(D_m) \leq \frac{2Cl}{1-l} + \frac{2C\sqrt{l}}{1-\sqrt{\mu}}.$$

Thus almost every point p has a lsm of length $l(p)$ through it. Any two points on the same lsm have the same P_j name for all time. An analogous argument holds for lum's. We note that the arguments given in the proof of Lemma 1 apply to B^* and hence given $\epsilon, \delta > 0$ we may cover T^2 disjointly up to measure $1 - \delta$ with ϵ-good hyperbolic blocks of B^*.

□

4. Stability of Toral Linked-Twist Maps

In this section the stability properties of B^* and Bernoulli toral linked twist maps of finite entropy are established. The proofs of these results are very similar. We show that the perturbed and unperturbed systems are close in the $\bar{d}$-metric and have equal entropy (in certain cases we need to take a direct product with a stochastic process to achieve equality). Then Proposition 2 implies that the systems are α-congruent. We give the proof of the stochastic stability of B^* under a perturbation of type I in detail and then modify this proof as needed in later sections to establish the other stability results.

4.1. **Proof of the stochastic stability of B^*.** The phase space of the perturbed system is $T^2 \times \Omega$ where $\Omega \equiv (T^2)^Z$ is the space of bilaterally infinite sequences in T^2. Each sequence $\sigma = (.., \sigma_{-1}, \sigma_0, \sigma_1, ...) \in \Omega$ has a distinguished entry σ_0. A measure m is induced on the space of sequences Ω by the B- process of the perturbation in the usual way. The shift $S : \Omega \to \Omega$, defined by $(S\sigma)_i = \sigma_{i+1}$ preserves m and is Bernoulli.

The perturbed process is the skew-product transformation $B^\#$ on $T^2 \times \Omega$ given by $B^\#(p, \sigma) = (B^*(p) + \sigma_0, S\sigma)$ (where $p \in T^2$, $\sigma \in \Omega$). $B^\#$ preserves $\mu \times m$. The projection function $\Pi : T^2 \times \Omega \to T^2$ gives the location of the path at time 0 i.e. $\Pi(p, \sigma) = p$.

Steps in the proof

Given $\alpha > 0$, we show

I) $(B^\#, P_\alpha \times \Omega)$ is a B-process where P_α is a finite partition of T^2 into sets of diameter less than $\frac{\alpha}{10}$.

II) $\bar{d}((B^\#, P_\alpha \times \Omega), (B^*, P_\alpha)) \le \frac{\alpha}{10}$ if r is chosen sufficiently small.

III) $B^\#$ has infinite entropy and hence by Proposition 2 $(B^\#, P_a \times \Omega)$ is α - congruent to (B^*, P_α).

The argument we give to obtain infinite entropy essentially obtains a lower bound for entropy in terms of the asymptotic expansion rate of lum's (one direction of Pesin's Entropy Formula [6]).

4.2. **Proof of I.** Fix a realisation of the perturbed process, denoted $B_\sigma^\#$. $B_\sigma^\#$ acts as a sequence of measure-preserving transformations of T^2 such that ${B_\sigma^\#}^j = B_j \dots B_1 B_0$ if $j \ge 0$ (where $B_i(p) = B^*(p) + \sigma_i$) and $(B_\sigma^\#)^j = B_{-j} B_{-(j-1)} \dots B_{-1}$ if $j < 0$.

We construct non-stationary hyperbolic blocks in the torus T^2 above each point in the base space of random jumps. We denote the torus above the sequence with distinguished entry σ_i as $T^2(i)$.

We show,

a) Given $\eta, \delta > 0$ it is possible to cover $T^2(i)$ (uniformly over i) up to measure $1 - \delta$ with η-good hyperbolic blocks of approximately the same measure.

b) Each atom of the distant past (the tailfield) is uniformly distributed over each hyperbolic block in the cover of $T^2(i)$.

Once we have established that these two conditions hold we may repeat the argument given first by Ornstein and Weiss [13] to prove that geodesic flows on manifolds of negative curvature are Bernoulli.

Proof of a)

The construction of lsm's and lum's for B^* detailed in Section 3.5 holds with no modification in the case of a realisation of a jump perturbation of B^*, and the estimates used in the construction are uniform over the possible realisations. Similarly the construction of ϵ-good hyperbolic blocks disjointly covering the torus up to measure $1-\delta$ for each realisation of the perturbation proceeds exactly as detailed in Section 3.2 and it is easily seen that this construction makes reference only to the length of lsm's and lum's and the angle between the stable and unstable manifolds. Thus given $\epsilon, \delta > 0$, for each realisation of the perturbation it is possible to cover the torus with ϵ-good hyperbolic blocks up to measure $1-\delta$.

$\square$

Proof of b)

Fix a realisation of the perturbation with distinguished entry σ_0. We show that the tailfield $\mathcal{T}(\prime)$ of $T^2(0)$ is trivial. The realisation is a biinfinite sequence of measure preserving maps $\{B_j\}$. Denote $B^n \equiv B_{n-1}\dots B_0$ and $B^{-n} \equiv B_{-n}^{-1}\dots B_{-1}^{-1}$. Note that $(B^j)^{-1}$ is not necessarily the same as B^{-j}. Points in $T^2(0)$ whose P_α names agree till time n constitute the atoms of $\vee_0^n {(B^j)}^{-1} P_\alpha$. The tailfield $\mathcal{T}(0)$, is given by $\lim_{n\to\infty} \vee_{-\infty}^{-n} B^j P_\alpha$.

First we show that the tailfield is atomic. Note that atoms of the tailfield contain by definition lum's. Suppose, in order to obtain a contradiction, that $\mathcal{T}(0)$ has a continuous part, namely a set of measure $\delta > 0$ which is composed of atoms of arbitrarily small measure.

Cover each $T^2(i)$ with ϵ-good hyperbolic blocks ($\epsilon < \frac{\delta}{100}$) up to measure $1-\frac{\delta}{100}$. We may specify that each hyperbolic block in these covers has measure bounded below by $\nu > 0$ and lsm's and lum's of length bounded below by $l > 0$.

Denote the continuous part of $\mathcal{T}(0)$ by Γ. Disjointify Γ into sets of atoms (Γ_i) of measure β where $0 < \beta < \frac{\nu}{100}$. Let $\mathcal{H}$ denote the union of the hyperbolic blocks in the cover of $T^2(0)$. Then $\mu_\Gamma(\mathcal{H}) > .\exists$. As a consequence of the absolute continuity of the foliation into lsm's and lum's μ a.e. point $p \in \Gamma \cap \mathcal{H}$ is a point of Lebesgue density for a lsm of one of the hyperbolic blocks contained in $\mathcal{H}$ in the sense that,

$$\lim_{y\in l^s(p)\to p} \frac{\Lambda((p-y,p+y)\cap l^s(p)\cap\Gamma)}{\Lambda((p-y,p+y)\cap l^s(p))} = 1$$

where Λ is the induced Lebesgue measure on $l^s(p)$.

Take d small enough so that for all but a set B of points in $\Gamma \cap H$ of measure $\rho < \frac{\beta}{100}$ we have

$$\frac{\Lambda(p-y,p+y)\cap l^s(p)\cap\Gamma)}{\Lambda((p-y,p+y)\cap l^s(p)} > .99$$

if $p\in\Gamma$ and $\mid p-y \mid < d$.

Define $G \equiv (\Gamma\cap H)\backslash B$. Recall that we have a uniform minimal rate of contraction $\frac{1}{\mu}$. Choose m large enough to ensure that $\frac{l}{\mu^m} < d$. The intersection of $B^{-m}\Gamma$ with the hyperbolic blocks of $T^2(-m)$ has relative measure on $B^{-m}\Gamma$ greater than .9. The pushforward F_m of the lsm's of the hyperbolic blocks in the cover of $T^2(-m)$ by

$B_{-1}B_{-2}...B_{-m}$ covers $T^2(0)$ except for a set of Lebesgue measure at most $1 - \frac{2\delta}{100}$. Thus $\mu_{\Gamma\cap\mathcal{H}}(F_m) > .98$. The lsm's of F_m form a subset of the lsm's of $T^2(0)$.

Suppose (Γ_i) is such that $\mu(\Gamma_i \cap F_m \cap H_j \cap G) > 0$ for some hyperbolic block H_j in H. Suppose $p \in \Gamma_i \cap F_m \cap H_j \cap G$. Let Λ be the conditional measure on $l^s(p)$ and let l^s_{-m} be the unique stable manifold of $T^2(-m)$ such that p lies in $B_{-1} \ldots B_{-m} l^s_{-m}$.

Since

$$\frac{\Lambda(B_{-1} \ldots B_{-m} l^s_{-m} \cap \Gamma_i \cap l^s(p))}{\Lambda(B_{-1} \ldots B_{-m} l^s_{-m} \cap l^s(p))} > .99$$

we have

$$\frac{\Lambda_m(l^s_{-m} \cap B^{-1}_{-m} \ldots B^{-1}_{-1}\Gamma^i)}{\Lambda(l^s_{-m})} > .99$$

(where Λ_m is the induced Lebesgue measure on l^s_{-m}).

Hence if $B^{-1}_{-m} \ldots B^{-1}_{-1}(\Gamma^i \cap G)$ intersects a lsm of a hyperbolic block of $T^2(-m)$ then it contains most of the one dimensional Lebesgue measure of the lsm. At least one of the sets Γ_i is such that $\mu(\Gamma_i \cap F_m \cap H \cap G) > 0$ and $B^{-1}_{-m} \ldots B^{-1}_{-1}(\Gamma_i \cap G)$ intersects a hyperbolic block of T^2_{-m} in a set of positive measure.

Suppose Γ^* has this property. Thus $B^{-1}_{-m} \ldots B^{-1}_{-1}(\Gamma^* \cap G)$ contains more than the fraction .9 of the induced Lebesgue measure of a lsm if it intersects this lsm in a set of positive induced Lebesgue measure. Furthermore $B^{-1}_{-m} \ldots B^{-1}_{-1}(\Gamma^* \cap G)$ contains whole lum's. Thus if $B^{-1}_{-m} \ldots B^{-1}_{-1}(\Gamma^* \cap G)$ intersects a hyperbolic block of T^2_{-m} it contains more than half of the measure of this hyperbolic block (which has measure greater than $\frac{\nu}{2}$).

Thus $\mu(\Gamma^* \cap G) = \mu(B_{-m} \ldots B^{-1}_{-1}(\Gamma^* \cap G) \geq \frac{\nu}{2}$, a contradiction to the assumption that $\mu(\Gamma^*) = \beta$. Therefore the tailfield, $\mathcal{T}(0)$, is atomic.

Now we show that $\mathcal{T}(0)$ is actually trivial. Suppose, in order to obtain a contradiction, that we have two distinct atoms of the tailfield. Denote these atoms A and D respectively. Since the tailfield is atomic we may assume, without loss of generality, that $0 < \nu = \mu(A) \leq \mu(D)$.

Choose $0 < \epsilon < \frac{\nu}{100}$ and $0 < \delta < \frac{\nu}{100}$ and cover each fiber $T^2(i)$ (except for a set of points of measure at most δ) with ϵ-good hyperbolic blocks. Let $\mathcal{S}$ (respectively $\mathcal{U}$) denote the union of all lsm's (respectively lum's) contained in the hyperbolic block cover of $T^2(0)$. Then $\mu(A \cap \mathcal{S}) > \frac{\exists\exists}{\infty''}$ and $\mu(D \cap \mathcal{U}) > \frac{\exists\exists}{\infty''}\mu(\mathcal{D})$.

The set of points p in $A \cap \mathcal{S}$ which are points of Lebesgue density for A in the sense that

$$\lim_{y \to p} \frac{\Lambda((p-y, p+y) \cap l^s(p) \cap A)}{\Lambda((p-y, p+y) \cap l^s(p))} = 1$$

(where Λ is the induced measure on $l^u(s)$) has Lebesgue measure greater than $\frac{49}{50}$. Let $\tilde{A}$ denote this subset of A.

Recall that if $B^{-n}(\tilde{A})$ intersects a lsm of $T^2(-n)$ then $B^{-n}(\tilde{A})$ contains a fraction $\delta(n)$ of the induced Lebesgue measure of that lsm and $\delta(n) \to 1$ as $n \to \infty$. Hence given $\eta > 0$ we may choose n large enough to ensure that if $B^{-n}\tilde{A}$ intersects a lsm l^s of $T^2(-n)$ then $B^{-n}(A \cap \mathcal{C}) \cap l^s$ accounts for more than $1 - \eta$ of the one-dimensional Lebesgue measure of l^s. We now state a lemma whose proof is almost identical to the 'cut and expand' argument of Section 3.4 used to proof the Bernoulliness of the finite entropy tltm's.

Lemma 5. *Suppose $T^2(j)$ has a cover of hyperbolic blocks (except for a set of measure at most δ) and the lsm's and lum's composing the hyperbolic blocks have length greater than l. Then there exist $\beta > 0$ and $r > 0$ such that except for a set of points p of measure at most 2δ, $B^{-r}l^s(p)$ (respectively $B^r l^u(p)$) will contain segments which cross $T^2(j-r)$ horizontally (respectively $T^2(j+r)$ vertically). These segments constitute more than a fraction β of the induced Lebesgue measure (length) of $B^{-r}l^s(p)$ and $B^r l^u(p)$.*

Recall that for all m, $B^{-m}D$ contains whole lum's and for each i there is a cover of $T^2(i)$ (except for a set of measure at most δ) with ϵ-good hyperbolic blocks. We may assume that uniformly over i the lum's and lsm's in each cover have length bounded below by $l > 0$. Now apply lemma 5 with l, δ to obtain r, β satisfying the assertion of the lemma. Choose n large enough so that if $B^{-n}\tilde{A}$ intersects a lsm l^s of $T^2(-n)$ then $B^{-n}(A \cap \mathcal{C}) \cap l^s$ is more than $1-\eta$ of the one-dimensional Lebesgue measure of l^s. Define $M = 2r + N$. Hence except for a set of points of measure at most 2δ each point $p \in T^2(-N)$ has a lsm $l^s(p)$ such that $B^{-r}l^s(p)$ contains segments that traverse the fiber $T^2(\frac{M+N}{2})$ and these segments comprise more than a fraction β of the length of $B^{-r}l^s(p)$. Similarly except for a set of points of measure at most 2δ each point $q \in T^2(-M)$ has a lum $l^u(p)$ such that $B^r l^u(q)$ contains segments which traverse $T^2(-\frac{M+N}{2})$ vertically and these segments comprise more than a fraction β of the length of $B^r l^u(q)$. Note that the induced measure of

$$\cup\{B^r l^u(q) : l^u(q) \in B^{-M}(D)\}$$

on each image $B^{-r}l^s(p)$ is greater than $\nu\beta$. Note also that the induced measure of $B^{-r}(l^s(p) \cap \tilde{A})$ on $B^{-r}l^s(p)$ is greater than $1-\eta$ which is greater than $1-\alpha\beta$.

Thus except for a set of points of $T^2(-N)$ of measure at most 2δ if $l^s(p)$ intersects $B^{-N}(\tilde{A})$ then

$$I \equiv \cup_{l^u(q) \in B^{-M}D} B^r(l^u(q)) \cap B^{-r-N}\tilde{A}$$

induces positive Lebesgue measure on $B^{-r}(l^s(p))$.

Therefore the set $B^{r+N}(I)$ has positive Lebesgue measure and is contained in $A \cap D$ a contradiction to the assumption that the two atoms were disjoint.
Therefore $\mathcal{T}(0)$ is trivial.

□

Proof of II)

Note that (B^*, P_α) is a finitely determined process. Hence given $\alpha > 0$ there exists a $\delta, \gamma, T > 0$ such that if

1) $\bar{d}(\{B^{*-j}P_\alpha\}_0^T, \{B^{\#-j}(P_\alpha \times \Omega)\}_0^T) < \delta$.

2) $h(B^\#, P_\alpha \times \Omega) > h(B^\#, P_\alpha) - \gamma$

then

3) $\bar{d}((B^*, P_\alpha), (B^\#, P_\alpha \times \Omega)) < \alpha$.

In the ensuing discussion let γ, δ, T refer to those quantities arising from our choice of α. Note that given $T' > 0$ uniformly over all possible realisations of the perturbed process $B_\sigma^{\#-j}$,

$$\bar{d}(\{B^{*-j}P_\alpha\}_0^{T'}, \{B_A^{\#-j}P_\alpha\}_0^{T'}) < \delta$$

if r is sufficiently small.
Hence we need only show condition 2) holds.
For a realisation $(B_\sigma^\#, P_\alpha)$ of the perturbed process define

$$h^\sigma(P_\alpha) \equiv \lim_{n\to\infty} \frac{1}{n} H(\vee_0^{n-1} B_\sigma^{\#j} P_\alpha)$$

By the Abramov and Rohlin [1] skew-product entropy formula

$$h(B^{\#}, P_a \times \Omega) = \int h^\sigma(P_\alpha) dm + h(S, \Omega),$$

where the integration is over all possible realisations with respect to the invariant measure m.

This means that a lower bound for the entropy of the perturbed process with respect to the partition $P_\alpha \times \Omega$ is given by calculating the entropy on each realisation and then integrating over the possible realisations of the perturbation.

We may repeat the arguments of Section 3.2, in particular Lemmas 1, 2 and 3 with the obvious modifications, to obtain a lower bound for the entropy of a realisation in terms of the entropy of the unperturbed process. Once again we use the similarity in hyperbolic block structure between $B_\sigma^{\#}$ and B^* to obtain this lower bound. The only points to note are that the properties of the hyperbolic block cover described in Lemma 1 hold uniformly over realisations of the perturbed process and that in Lemma 2 we apply the Shannon-McMillan-Breiman Theorem to (B^*, P_α) and hence the T' we obtain does not depend upon the realisation we are considering.

Thus uniformly over realisations there exists an $r > 0$ and a T' such that if the perturbation is smaller than r then for each realisation Φ of the perturbed process we have,
$\frac{1}{T'} H(\vee_0^{T'} \Phi^{-j} P_\alpha \mid \vee_1^\infty \Phi^j P_\alpha)$
$\geq h(B^*, P_\alpha) - \gamma$
Let $T \equiv T' + 1$. Then we have,
$H(\vee_0^{kT-1} \Phi^{-j} P_\alpha) = H(\vee_0^{T-1} \Phi^{-j} P_\alpha) + H(\vee_T^{kT-1} \Phi^{-j} P_\alpha \mid \vee_0^{T-1} \Phi^{-j} P_\alpha)$
$= H(\vee_0^{T-1} \Phi^{-j} P_\alpha) + H(\vee_T^{2T-1} \Phi^{-j} P_\alpha \mid \vee_0^{T-1} \Phi^{-j} P_\alpha) + H(\vee_{2T}^{kT-1} \Phi^{-j} P_\alpha \mid \vee_0^{2T-1} \Phi^{-j} P_\alpha)$
$\geq \sum_{m=0}^{k-1} H(\vee_{mT}^{(m+1)T-1} \Phi^{-j} P_\alpha \mid \vee_{-\infty}^{mT-1} \Phi^{-j} P_\alpha)$
$\geq kT(h(B^*, P_\alpha) - \gamma)$ by the foregoing result.
Thus
$\lim_{k\to\infty} \frac{1}{kT} H(\vee_0^{kT-1} \Phi^{-j} P_\alpha) \geq h(B^*, P_\alpha) - \gamma$.

As this estimate holds for all realisations, by our previous remarks concerning skew-product entropy we have

$h(B^{\#}, P_\alpha \times \Omega) \geq h(B^*, P_\alpha) - \gamma$. □

Thus the perturbed process is within α in the $\bar{d}$-metric of the unperturbed process.

We remark here that if the stochastic process constituting the pertubation has infinite entropy (which would be the case if the jumps comprising the perturbation are taken to be independent and uniformly distributed in a disk) then by results of Ornstein and Weiss we would actually have α-congruence.

We have not made the assumption that the perturbation is of infinite entropy and so this must first be proved before Proposition 2 can be applied. This is done next.

□

Proof of III)

We show that given $M > 0$ and $\epsilon > 0$ there is a finite partition P_n and a k such that

$$H(\vee_0^k B_\sigma^{\#^{-j}} P_n \mid \vee_1^\infty \Phi^j P_n) > k(M-\epsilon)$$

for a set of realisations $B_\sigma^\#$ of the perturbed process of m measure greater than $1-\epsilon$.

This implies that $h^\sigma(P_n) > M-\epsilon$ for a set of realisations of m measure greater than $1-\epsilon$. This establishes the result as M was arbitrary.

Recall the definition of the sets $Q_{ss} \equiv \{p \in T^2 : p \in (0,1] \times (l_s, l_{s+1}]$ and $\tau_1(p) \in (l_s, l_{s+1}]\}$. On Q_{ss} the expansion rate of B^* in the expanding directon is equal to $f_s'|_{(l_s,l_{s+1})}$ where $\mu(Q_{ss})\log(f_s'|_{(l_s,l_{s+1})}) > 10s$.

Note that given $\epsilon > 0$ there exists $l(\epsilon) > 0$ such that we may cover each torus $T^2(i)$ of each realisation up to measure $1-\epsilon$ with lum's and lsm's of length greater than $l(\epsilon)$ (we are abusing notation here as the lum's and lsm's are not defined in reference to a partition but the meaning is clear).

Now choose $k(M,\epsilon)$ large enough so that except for a set of realisations of m measure at most ϵ each realisation $B_\sigma^\#$ has lum's l^u which cover up to measure $1-\epsilon$ the torus upon which the realisation acts and satisfy $\mid B_\sigma^{\#^{-k'}} l^u \mid < e^{-(M-\epsilon)k'} \mid l^u \mid$ for all $k' \geq k$.

We may choose such a k as almost every lum (with respect to Lebesgue measure) on almost every realisation (with respect to the measure m) is subject to a contraction of magnitude $\frac{1}{f_s'}$ under Φ^{-1}, $\mu(Q_{ss})$ percent of the time asymptotically by the ergodic theorem applied to the Bernoulli map $B^\#$.

As $\sum_s \mu(Q_{ss})\log(f'_{s|(l_s,l_{s+1})})) = \infty$ we may choose $k(\epsilon, M)$ large enough so that for all lum's except for a set of measure at most ϵ on all but at most m measure ϵ of the realisations Φ satisfy

$\mid \Phi^{-k'} l^u \mid < e^{-(M-\epsilon)k'} \mid l^u \mid$ for all $k' \geq k$.

In fact as $\sum_s \mu(Q_{ss})\log(f_s') = \infty$ we may take the asymptotic rate of contraction as large as we like. This is the property that enables us to establish infinite entropy.

We may also choose k large enough so that up to measure $1-\epsilon$ of the lum's l^u on up to m measure $1-\epsilon$ of the lum's lie wholly within sets of a partition P_n of the fibers $T^2(i)$ into squares of sidelength $\frac{1}{n} \simeq e^{-k(M-2\epsilon)} l(\epsilon)$ and have the same P_n name at all times in the past (ie, points on $\Phi^{-k-j} l^u$ have the same P_n name for all $j \geq 0$).

Each such lum l^u expands in length by at least a factor of $e^{k(M-\epsilon)}$ under application of $B_\sigma^{\#^k}$. Hence l^u will be fragmented into at least $e^{k(M-2\epsilon)}$ P_n names (since $\frac{1}{n} \simeq e^{-k(M-2\epsilon)} l(\epsilon)$).

Now $\sqrt{\epsilon}$ a.e. atom of $\vee_1^t B_\sigma^{\#^j} P_n$ consists of the image of the lum's $B_\sigma^{\#^{-k}} l^u$ up to relative measure $\sqrt{\epsilon}$. This is uniform over t. We call such atoms of $\vee_1^t B_\sigma^{\#^j} P_n$ for a fixed t 'good' atoms. Let A(t) be the subset of $\vee_1^t B_\sigma^{\#^j} P_n$ which consists of the 'good' atoms. Then $\mu(\vee_1^t B_\sigma^{\#^j} P_n \backslash A(t)) < \sqrt{\epsilon}$ for all t. Let (P_α) and (Q_β) denote the atoms of $\vee_0^k B_\sigma^{\#^{-j}} P_n$ and $\vee_1^t B_\sigma^{\#^j} P_n$ repectively. Then,

$$H(\vee_0^k B_\sigma^{\#^{-j}} P_n \mid \vee_1^t B_\sigma^{\#^j} P_n)$$

$$= -\sum_{\alpha,\beta} \mu(Q_\beta)\left(\frac{\mu(P_\alpha \cap Q_\beta)}{\mu(Q_\beta)} \log \frac{\mu(P_\alpha \cap Q_\beta)}{\mu(Q_\beta)}\right)$$

$$= -\sum_{\beta \in A(t),\alpha} \mu(Q_\beta)\left(\frac{\mu(P_\alpha \cap Q_\beta)}{\mu(Q_\beta)} \log \frac{\mu(P_\alpha \cap Q_\beta)}{\mu(Q_\beta)}\right)$$

$$(\ddagger) - \sum_{\beta \in A(t)^c,\alpha} \mu(Q_\beta)\left(\frac{\mu(P_\alpha \cap Q_\beta)}{\mu(Q_\beta)} \log \frac{\mu(P_\alpha \cap Q_\beta)}{\mu(Q_\beta)}\right)$$

$$\geq (1-\sqrt{\epsilon}) \log(e^{k(M-\epsilon)}) - (1-\sqrt{\epsilon}) \log(1-2\epsilon)$$

$$\geq (1-\sqrt{\epsilon})(k(M-\epsilon)) - (1-\sqrt{\epsilon}) \log(1-2\epsilon)$$

As this estimate is uniform over t and ϵ was arbitrary this establishes the claim.

Remark 7. As stated this argument gives the integral of the logarithm of the maximal eigenvalue $\Lambda(p)$ of a tltm satisfying the template conditions of the introduction as a lower bound for the entropy of the map but it can clearly be sharpened in line $\ddagger$ to establish equality in the case that the tltm has a bounded maximal eigenvalue.

□

Thus if r is sufficiently small then the conditions of Proposition 2 are satisfied and the perturbed and unperturbed systems are $\alpha-$congruent.

□

5. Stability of B^* under a Perturbation of Type II

See Section 1.1 for notation. We show that given $\alpha > 0$ if $r > 0$ is sufficiently small then there is an isomorphism $\Psi : (B^* \times S, T^2 \times \Omega) \to (B^\#_N, T^2 \times \Omega)$ such that $\Pi\Psi(p,w)$ lies within α of p with probability $1-\alpha$.

Since the perturbed system is not Bernoulli we cannot directly apply the arguments given in the case of a perturbation by a B-process as the form of Proposition 2 we have given only holds for Bernoulli systems. To overcome this we use the Relative Bernoulli theory in a stronger form.

As a consequence of Thouvenot's Relative Isomorphism theorem and the work of Ornstein and Weiss [12] it is known that if

1) $(B^\#_N, T^2 \times \Omega)$ is Bernoulli relative to the factor $\mathcal{A}$ defined by $B^\#_N$ restricted to Ω (in other words the 'noise' factor).

2) $(B^\#_N, I \times \Omega)$ is within $\frac{\alpha}{10}$ of (B^*, I) in the $\bar{d}$ metric relative to $\mathcal{A}$, (where I is the partition of T^2 given by the identity function. This means for almost every realisation $B^\#_{N\sigma}$ of the perturbation, $\bar{d}((B^\#_{N\sigma}, I), (B^*, I)) < \frac{\alpha}{10}$)

and

3) $h(B^\#) = h(S) + h(B^*)$,

then

$(B^*, T^2) \times (S, \Omega)$ is $\alpha-$congruent to $(B^\#_N, T^2 \times \Omega)$.

This is because the Embedding Theorem [12, page 64] states that if 1) holds and

$\ddagger$

$h(B^\#_N, T^2 \times \Omega) \geq h(S) + h(B^*)$

then given $\epsilon > 0$ there is a partition $\bar{I}$ such that

i) $d(\bar{I}, I) < \epsilon$

ii) $\vee_{-\infty}^{\infty} B_N^{\#-j} \bar{I}$ is independent of $\mathcal{A}$

iii) $(B^\#_N, \bar{I})$ is congruent (identically distributed) to (B^*, I).

If in addition equality holds in $\ddagger$ then $\bar{I}$ may be chosen so that

$\vee_{-\infty}^{\infty} B_N^{\#-j} \bar{I} \vee \mathcal{A} = \mathcal{B}$.

Thus in the case of equality in ‡ the canonical isomorphism $(B_N^{\#}, \bar{I}) \to (B^*, I)$ extends to an α-congruence between $(B \times S, T^2 \times \Omega)$ and $(B_N^{\#}, T^2 \times \Omega)$.

Hence it is only necessary to verify conditions 1),2) 3) and show that equality holds in ‡ to show that $(B_N^{\#}, T^2 \times \Omega)$ is isomorphic to $(B^*, I) \times (S, \Omega)$ by an isomorphism Ψ which ensures $\Pi\Psi^{-1}(p, \sigma)$ is within α of p for all but a set of points (p, σ) of measure less than α (this is a consequence of i) and iii) above).

Note that we have equality in ‡ (condition 3 holds) since the argument given to prove $B^{\#}$ has infinite entropy also holds for $B_N^{\#}$. Now we show that conditions 1), and 2) are also satisfied.

Condition 1) is checked exactly as in the case of the perturbation of B^* by a Bernoulli process. To verify condition 2) one uses the fact that for any partition P_a into sets of diameter less than α (B^*, P_α) is a B-process and hence finitely determined. Then the same argument given when the perturbation is a B-process holds (involving an entropy and finite distribution check on each realisation of the perturbed process) and so if $r > 0$ is sufficiently small then $(B_N^{\#}, P_{\frac{\alpha}{10}} \times \Omega)$ is within $\frac{\alpha}{10}$ of $(B^*, P_{\frac{\alpha}{10}})$ in the $\bar{d}$ metric. Since $P_{\frac{\alpha}{10}}$ was taken to be a partition into sets of diameter less than $\frac{\alpha}{10}$ this implies $\bar{d}((B_N^{\#}(I \times \Omega), (B^*, I)) < \alpha$. So the perturbed sytem is α-congruent to the direct product of the unperturbed system with a stochastic process.

$\square$

5.1. **Stability of a Bernoulli TLTM of Finite Entropy under a Perturbation of Type I or Type II.** In this section we show that a Bernoulli TLTM B_f of finite entropy (satisfing the conditions stipulated in the introduction) is stable under any form of perturbation (i.e. the sequence of jumps may be given by any stationary stochastic process (S, Ω)) in the sense that given α if the size of the perturbation is sufficiently small then the perturbed system is $\alpha-$congruent to the direct product of the "noise" (S, Ω) and the unperturbed system.

Using the ideas of the previous section the perturbed system is the skew-product transformation $B_f^{\#} : T^2 \times \Omega \to T^2 \times \Omega$ defined by $B_f^{\#}(p, \sigma) = (B_f(p) + \sigma), S\sigma)$. The possible realisations of the perturbed system are given by the stationary stochastic process (S, Ω). The shift S preserves a measure m. The function $\Pi : T^2 \times \Omega \to T^2$ gives the location of orbits at time zero i.e. $\Pi(p, \sigma) = p$.

The proof of this result follows the lines of the proof given for B^*. Exactly the same arguments show that $(B_f^{\#}, T^2 \times \Omega)$ is relatively VWB with respect to the sigma algebra defined by (S, Ω). If P_α is a partition of T^2 into sets of diameter less than $\frac{\alpha}{10}$ then by the same arguments previously employed it is shown that if the perturbation is sufficiently small then $\bar{d}((B_f^{\#}, P_\alpha \times \Omega), (B_f, P_\alpha)) < \frac{\alpha}{10}$. Now let $g : T^2 \to R$ be a continuous function. Since $(B_f^{\#}, T^2 \times \Omega)$ is VWB relative to (S, Ω) for m almost every realisation $B_{f\sigma}$ the time average

$$\lim_{n \to \infty} \frac{1}{n} \sum_{0}^{n-1} g(B_{f\sigma}^j(p))$$

converges for almost every point p to $\int_{T^2} g d\mu$. This is a consequence of the fact that as $B_f^{\#}$ preserves the product measure $\mu \times m$ for m a.e. $\sigma \in \Omega$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{0}^{n-1} g(B^j_{f\sigma}(p))$$

converges for μ a.e. $p \in T^2$. The fact that m almost every realisation has a trivial tailfield implies that this limit must be a constant (and hence must equal $\int_{T^2} g d\mu$). Hence we have a form of nonstationary ergodicity which enables us to estimate the asymptotic expansion of lum's in terms of the integral of the logarithm of the maximum eigenvalue $\Lambda(p)$ of B_f over T^2.

The argument given to calculate the skew-product entropy of B^* in terms of the asymptotic expansion rate of lum's may be duplicated (see remark 7) to show

$$\begin{aligned} \int_\Omega h^\sigma(B_{f\sigma}) &= \int_{T^2} \log|\Lambda(p)| \\ &= h(B_f) \end{aligned}$$

Thus $h(B^{\#}_f) = h(S) + h(B_f)$ so that , as in the previous section, there is an α-congruence between $(B_f \times S, T^2 \times \Omega)$ and $(B^{\#}_f, T^2 \times \Omega)$.

□

Acknowledgement:

I would like to thank Don Ornstein for many valuable conversations. This research was supported in part by NSF Grant DMS-9101836, ONR Grant N00014-94-1-0317 and the Texas Advanced Research Program (003652037).

References

[1] L. Abramov and V. A. Rohlin. The entropy of a skew product of measure-preserving transformations. *Vestnik Leningrad Univ.* 17, 5-13, 1962.

[2] R. Burton and R. Easton. Ergodicity of linked twist maps. *Global Theory of Dynamical Systems* 819, 35-49, 1980.

[3] N. Chernov and C. Haskell. Nonuniformly hyperbolic K-systems are Bernoulli. *preprint* September 1994.

[4] G. Gallavotti and D.S Ornstein. Billiards and Bernoulli schemes. *Comm. Math. Phys.* 38, 83-101,1974.

[5] E. Hopf. Statistik der geodatischen linien in mannigfaltigkeiten krummung. *Ber. Verh. Sachs* 91, 261-304, 1939.

[6] A. Katok and J.M. Strelcyn and F. Ledrappier and F. Przytycki. *Invariant manifolds, entropy and billiards, smooth maps with singularities* 1222, Springer Verlag LNS, 1986.

[7] A. N. Kolmogorov. A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces. *Dokl. Akad. Nauk. SSSR* 119, 861-864, 1958.

[8] A. G. Kouchnirenko. An estimate from above for the entropy of a classical system. *Dokl. Akad. Nauk.* 161, 37-38, Berlin, 1965.

[9] R. Mane. *Ergodic theory and differentiable dynamics*, Springer Verlag, Berlin, 1982.

[10] M. Nicol. A Bernoulli toral linked twist map without positive Lyapunov exponents. *Proc. AMS, to appear*, 1994.

[11] D. S. Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Adv. in Math* 4, 337-352, 1970.

[12] D. S. Ornstein and B. Weiss. Statistical properties of chaotic systems. *Bull. Amer. Math. Soc* 24, 11-116, 1991.

[13] D. S. Ornstein and B. Weiss. Geodesic flows are Bernoullian. *Israel J. Math* 14, 184-198, 1973.

[14] F. Pryztycki. Ergodicity of total linked twist mappings. *Ann. Scient. Ec. Norm. Sup. Series 1* 16, 345-355, 1983.

[15] D. Rudolph. *Fundamentals of measurable dynamics*, Clarendon Press, Oxford, 1990.

[16] Y. Sinai. Dynamical systems with elastic reflections. *Uspekh. Mat. Nauk.* 27, 137-189, 1972.

[17] L. S. Young. Stochastic stability of hyperbolic attractors. *Ergod. Th. and Dynam. Sys.* 6, 311-319, 1986.

[18] S. Troubetzkoy. A comparison of elastic and inelastic billiards. *Nonlinearity* 3, 3, (1990), 947-961.

[19] J. P. Thouvenot. Quelques propriétés des systèmes dynamiques qui se décomposent en un product de deux systèmes dont l'un est un schéma de Bernoulli. *Israel J. Math.* 21, 177-207, 1975.

[20] S. Troubetzkoy. A comparison of elastic and inelastic billiards. *Nonlinearity* 3, 3, (1990), 947-961.

[21] M. P. Wojtkowski, *Invariant families of cones and Lyapunov exponents*, Erg. Th. Dyn. Syst. 5 (1985), 145-161.

Mathematics Institute, University of Warwick,Coventry, CV4 7AL
E-mail address: `matt@@maths.warwick.ac.uk`