Purdue University
Purdue e-Pubs

Open Access Dissertations   Theses and Dissertations

Spring 2015

Ergodic properties of countable extensions

Samuel Joshua Roth
Purdue University

Follow this and additional works at: https://docs.lib.purdue.edu/open_access_dissertations

Part of the Mathematics Commons

Recommended Citation

Roth, Samuel Joshua, "Ergodic properties of countable extensions" (2015). *Open Access Dissertations*. 553.
https://docs.lib.purdue.edu/open_access_dissertations/553

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Samuel Joshua Roth

Entitled
Ergodic Properties of Countable Extensions

For the degree of   Doctor of Philosophy

Is approved by the final examining committee:

Michal Misiurewicz
Chair

Alexandre Eremenko

Roland Roeder

Rodrigo Perez

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Michal Misiurewicz

Approved by: Evgeny Mukhin   3/27/2015

Head of the Departmental Graduate Program   Date

ERGODIC PROPERTIES OF COUNTABLE EXTENSIONS

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Samuel Joshua Roth

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 2015

Purdue University

Indianapolis, Indiana

Serdeczne podziękowania Michałowi Misiurewiczowi za nieocenioną pomoc i opiekę.

Michał, ta praca jest dla Ciebie.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Roth, Samuel Joshua PhD, Purdue University, May 2015. Ergodic Properties of Countable Extensions. Major Professor: Michał Misiurewicz.

First, we study countably piecewise continuous, piecewise monotone interval maps. We establish a necessary and sufficient criterion for the existence of a non-decreasing semiconjugacy to an interval map of constant slope in terms of the existence of an eigenvector of an operator acting on a space of measures. Then we give examples, both Markov and non-Markov, for which the criterion is violated.

Next, we establish a criterion for the existence of a constant slope map on the extended real line conjugate to a given countably piecewise monotone interval map. We require the given interval map to be continuous, Markov, and topologically mixing, and show by example that the mixing hypothesis is essential.

Next, we study a class of countable state subshifts of finite type which admit finite-state factors. Our systems carry a displacement function, analogous to that used in the rotation theory of circle maps. Among those invariant measures on the factor system for which the average displacement is zero, we identify a unique measure of maximal entropy. As a corollary we obtain an efficient computational tool for the Gurevich entropy of the countable state system. We also prove that the countable state systems in our class do not admit any measure of maximal entropy.

Finally, we apply our findings to the study of degree one circle maps with Markov partitions and with transitive liftings to the real line. After compactifying by adjoining fixed points at plus and minus infinity, we show how to compute the topological entropy of the lifting and how to find all conjugate maps of constant slope. We prove that there are conjugate maps of constant slope for every slope greater than or equal to the exponential of the entropy.

1. NO SEMICONJUGACY TO AN INTERVAL MAP OF CONSTANT SLOPE

The idea that some interval maps should be conjugate or semiconjugate to maps of constant slope (we use the term "constant slope" instead of the more accurate but clumsy "constant absolute value of the slope") appeared first about 50 years ago. Parry [18] proved that continuous, transitive, piecewise monotone (with finite number of pieces) interval maps are conjugate to maps of constant slope. Later, Milnor and Thurston [14] proved an analogous result, removing the assumption of transitivity, but replacing conjugacy by semiconjugacy. Another proof of the Milnor-Thurston theorem appeared in [1]. That proof, after small modifications, can be used for maps of graphs or for piecewise continuous maps (with finitely many pieces) [2].

In all cases we require that the (semi)conjugacy is via monotone maps preserving orientation. This is a natural requirement; if we drop it then we get a completely different, and less interesting, problem. Also in all cases the logarithm of the slope is equal to the topological entropy of the initial map. This is because the same is true for the constant slope maps (see [16], [1]).

A natural question is what can be said if the map is piecewise monotone, but with *countably* many pieces. When trying to make such generalizations, one immediately encounters some basic problems.

The first problem is a definition of a countably piecewise monotone map. For such continuous maps, what should we assume about the set of turning points (local extrema)? For instance, if we allow the closure of this set to be a Cantor set, there may be substantial dynamics on it, not captured by our considerations. Thus, it is reasonable to assume that the closure of the set of turning points is countable.

The second problem is, what should the slope be? For countably piecewise monotone maps of constant slope it is no longer true that the entropy is the logarithm of the slope. There are obvious counterexamples, where all points of the interval, except the endpoints, are moved to the right, so the entropy is zero, but the slope is larger than 1. Thus, there is no natural choice of the slope of the map to which our map should be (semi)conjugate.

Recently, Bobok [6] considered the case of continuous, Markov, countably piecewise monotone interval maps. He found a necessary and sufficient condition for the existence of a nondecreasing semiconjugacy to a map of constant slope in terms of the existence of an eigenvector for a certain operator. The operator is given by a countably infinite 0-1 matrix representing the transitions in the Markov system, and the criterion asks for a nonnegative eigenvector in the sequence space ℓ^1. Bobok described a rich class of examples satisfying this criterion and proved that for many of these examples the constant slope so obtained is the exponential of the topological entropy of the original interval map. However, he did not give any examples that violate the criterion.

In this chapter, we study the general case of countably piecewise continuous, piecewise monotone interval maps without any Markov assumption. We also establish a necessary and sufficient criterion – analogous to Bobok's – for the existence of a nondecreasing semiconjugacy to a map of constant slope. It is given, like Bobok's criterion, in terms of existence of an eigenvector of some operator, but the operator acts on measures rather than on sequences. Then we construct a class of examples which violate that criterion. Our examples are continuous and transitive and thus have positive topological entropy.

The chapter is organized as follows. In Section 1.1 we present notation and define objects used in the next sections. In Section 1.2 we establish the criterion for the semiconjugacy to a map of constant slope. In Section 1.3 we produce sufficient conditions for this criterion being not satisfied. Theorem 1.11 can be considered the main technical result of the chapter. Section 1.4 contains the main theorem of the chapter,

Theorem 1.12, which gives us a large class of continuous transitive interval maps for which Theorem 1.11 applies. Finally, in Section 1.5 we give a concrete example of a one-parameter family of continuous transitive countably piecewise monotone interval maps that are not semiconjugate to a map of constant slope via a nondecreasing map. We show that in this family there are uncountably many Markov and uncountably many non-Markov maps.

This entire chapter has also been published as a standalone paper [15].

1.1 Notation and definitions

Let us introduce the various notations and definitions required to address the problem more clearly.

If P is a closed, countable subset of $[0, 1]$, then a component of the complement of P will be called a P-*basic interval*, and the set of all P-basic intervals will be denoted $\mathcal{B}(P)$.

We want to consider countably piecewise monotone interval maps, but not only continuous, but also piecewise continuous. That would mean that there exists a closed countable set $P \subset [0, 1]$ such that our map is continuous and monotone on each P-basic interval. However, this creates a question: what should be the values of our map at the points of P? If the map is continuous, this is not a problem. However, in general there is no good answer. Even if we allow two values at those points (one-sided limits from both sides), P may have accumulation points, and there is no natural way of extending our map to those points. Therefore we choose the simplest solution – we do not define the map at all at the points of P. This is not a new idea; a similar solution is normally used for instance in the holomorphic dynamics on the complex projective spaces of dimension larger than 1.

Thus, we define a class $\mathcal{C}$ of maps f for which there exists a closed, countable set $P \subset [0, 1]$, $f : [0, 1] \smallsetminus P \to [0, 1]$, and f is continuous and strictly monotone on each P-basic interval. Note that we assume strict monotonicity; while it is possible

to do everything that we do assuming only monotonicity, the technical details would be much more involved and they would obscure the ideas.

Similarly as in measure theory where two functions are considered equal if they differ only on a set of measure zero, we will consider two elements of $\mathcal{C}$ equal if they are equal on the complement of a closed countable set. This gives us a possibility of using different sets P for a given map $f \in \mathcal{C}$. Each such set for which f is continuous and strictly monotone on P-basic intervals will be called f-admissible.

Lemma 1.1 *A composition of two maps from $\mathcal{C}$ belongs to $\mathcal{C}$.*

Proof We will show that if $f, g \in \mathcal{C}$, the set P is f-admissible, and Q is g-admissible, then the set $P \cup f^{-1}(Q)$ is $g \circ f$-admissible. First observe that the set $(f|_I)^{-1}(Q)$ is countable for every P-basic interval I. Moreover, there are only countably many P-basic intervals. Thus, the set $P \cup f^{-1}(Q)$ is countable.

We may assume that $0, 1 \in P$. Let $[a, b]$ be the closure of a P-basic interval. Then $(f|_{(a,b)})^{-1}(Q)$ is closed in (a, b). Since $a, b \in P$, the set $(P \cup f^{-1}(Q)) \cap [a, b]$ is closed in $[a, b]$. Since P is closed in $[0, 1]$, this proves that $P \cup f^{-1}(Q)$ is closed in $[0, 1]$.

Let I be a component of the complement of $P \cup f^{-1}(Q)$. Then I is a subset of a P-basic interval and $f(I)$ is a subset of a Q-basic interval, so $g \circ f$ is continuous and strictly monotone on I. ∎

Using this lemma, by induction we get that if $f \in \mathcal{C}$ then $f^n \in \mathcal{C}$ for every natural n. That is, we can iterate a map from $\mathcal{C}$ without leaving this class of maps.

We do not want to abandon continuous maps. Therefore we consider the class $\mathcal{CC}$ of continuous maps $f : [0, 1] \to [0, 1]$ for which there exists a countable closed set $P \subset [0, 1]$ such that $f|_{[0,1] \smallsetminus P} \in \mathcal{C}$. Results for maps from this class will follow easily from the results for maps from $\mathcal{C}$. In view of Lemma 1.1, composition of maps from $\mathcal{CC}$ belongs to $\mathcal{CC}$, so in particular, iterates of a map from $\mathcal{CC}$ belong to $\mathcal{CC}$.

We will say that a map $f \in \mathcal{C}$ (or $f \in \mathcal{CC}$) has *constant slope* λ, if for some f-admissible set P, f restricted to each P-basic interval is affine with slope of absolute

value λ. Clearly, this property depends only on the map f, and not on the choice of an f-admissible set P.

We will say that a nonatomic measure defined on the Borel σ-algebra on the interval $[0,1]$ is *strongly σ-finite* if there is a closed countable set $P \subset [0,1]$ such that each P-basic interval has finite measure. We denote by $\mathcal{M}$ the set of all such measures. Observe that $\mathcal{M}$ is closed under addition and under multiplication by positive real scalars.

If $g : X \to Y$ is a map and ν is a measure on X, then we can always push this measure forward and define a measure $g_*\nu$ on Y by the formula $g_*\nu(A) = \nu(g^{-1}(A))$. If g is invertible and μ is a measure on Y, we can push this measure to X by g^{-1}. This defines a pull-back of μ, that is, $g^*\mu = g_*^{-1}\mu$. In terms of measures of sets, we have $g^*\mu(A) = \mu(g(A))$.

For a map $f \in \mathcal{C}$, we put all those pull-backs together to get an operator $T_f : \mathcal{M} \to \mathcal{M}$. It acts on a measure $\mu \in \mathcal{M}$ as follows. Choose an f-admissible set P. For each P-basic interval I, consider the homeomorphism $f|_I : I \to f(I)$. Pull back the measure $\mu|_{f(I)}$ by $f|_I$ to a measure on I. This defines $T_f\mu$ on the interval I:

$$(T_f\mu)|_I = (f|_I)^*(\mu|_{f(I)}), \quad I \in \mathcal{B}(P).$$

More explicitly,

$$(T_f\mu)(A) = \sum_{I \in \mathcal{B}(P)} \mu(f(I \cap A))$$

for all Borel sets A. By a common refinement argument, the definition of T_f depends only on the map f, and not on the choice of an f-admissible set P. Moreover, as in the proof of Lemma 1.1, if for a closed countable set Q each Q-basic interval has measure μ finite, then each R-basic interval, where $R = P \cup f^{-1}(Q)$, has measure $T_f\mu$ finite. This shows that indeed T_f maps $\mathcal{M}$ to $\mathcal{M}$.

Note the linearity properties $T_f(\mu + \nu) = T_f\mu + T_f\nu$ and $T_f(\alpha\mu) = \alpha T_f\mu$ for $\mu, \nu \in \mathcal{M}$ and $\alpha \geq 0$.

Instead of maps of the interval $[0, 1]$ into itself, we can consider maps of the circle or of the real line into itself. We will use for them the same notation as for the interval maps.

1.2   Semiconjugacy

Although $\mathcal{M}$ is not a true linear space (multiplication by negative scalars is not permitted), it is nevertheless quite fruitful to consider eigenvectors for positive eigenvalues. Let us consider the meaning of an eigenvector in this setting. Fix a map $f \in \mathcal{C}$ and an f-admissible set P. The condition $T_f \mu = \lambda \mu$, $\lambda > 0$, is equivalent to the condition that for all P-basic intervals I, $(f|_I)^*(\mu|_{f(I)}) = \lambda \mu|_I$. We will suppress subscripts and write $f^* \mu = \lambda \mu$ when the context is clear. However, if $f^* \mu = \lambda \mu$ on a P-basic interval I, then the Radon-Nikodym derivative $\frac{df^* \mu}{d\mu}$ is identically λ on I. This Radon-Nikodym derivative is the measure-theoretic version of the Jacobian of f on I as defined by Parry [19]. Thus, a measure $\mu \in \mathcal{M}$ satisfying $T_f \mu = \lambda \mu$ is a measure of constant Jacobian λ for f. If μ is an eigenvector for T_f, then for any subinterval J contained in a single P-basic interval I, $\mu(f(J)) = \lambda \mu(J)$. In words, if μ is an eigenvector for T_f, then within each P-basic interval, f uniformly stretches the μ measures of intervals by a factor λ. It is suggestive to note that constant slope maps have the same property, but with lengths in place of measures. We will show that this has a deeper meaning. Let us denote the Lebesgue measure by m.

Lemma 1.2 *The map $f \in \mathcal{C}$ has constant slope λ if and only if $T_f m = \lambda m$.*

Proof   Assume that f has constant slope λ and let P be f-admissible. Then f is affine with slope $\pm \lambda$ on each P-basic interval. Consider an arbitrary P-basic interval I and the restricted map $f|_I : I \to f(I)$. It suffices to prove that $f^* m = \lambda m$ on this interval. For every subinterval $(a, b) \subset I$ we have

$$(f^* m)(a, b) = m(f((a, b))) = |f(b) - f(a)| = \lambda(b - a) = (\lambda m)(a, b).$$

Since the two Borel measures f^*m and λm on I agree on all open intervals, they are in fact equal.

Conversely, assume that $T_f m = \lambda m$. Let I be a P-basic interval. Then for every subinterval $(a, b) \subset I$ we have

$$|f(b) - f(a)| = m(f((a,b))) = (f^*m)(a,b) = \lambda m(a,b) = \lambda(b - a).$$

Therefore $f|_I$ is affine with slope $\pm\lambda$. But I was arbitrary. Therefore f has constant slope λ. ∎

Remark 1.3 *With only slight modifications in wording, the proof of Lemma 1.2 goes through for maps of the real line into itself.*

Theorem 1.4 *Let $f \in \mathcal{C}$ and let $\lambda > 0$. Then f is semiconjugate via a nondecreasing map φ to some map $g \in \mathcal{C}$ of constant slope λ if and only if there exists a probability measure $\mu \in \mathcal{M}$ such that $T_f \mu = \lambda\mu$.*

Proof Assume that f is semiconjugate to a map g of constant slope λ by a nondecreasing map φ. Let P be an f-admissible set; it is clear that then $\varphi(P)$ is a g-admissible set. By Lemma 1.2, $T_g m = \lambda m$. Since m is nonatomic, it can be pulled back by the nondecreasing map φ to define a measure $\mu = \varphi^*m$; explicitly, $\mu(A) = m(\varphi(A))$ for all Borel sets A. Then μ is a nonatomic Borel probability measure. Now let I be any P-basic interval. Take restrictions of f, g, φ, m, and μ to the appropriate domains in the following commutative diagram:

$$\begin{array}{ccc}
I & \xrightarrow{\ f\ } & f(I) \\
\varphi \downarrow & & \downarrow \varphi \\
\varphi(I) & \xrightarrow[\ g\ ]{} & g(\varphi(I))
\end{array} \qquad (1.2.1)$$

Using these restricted maps and measures, $f^*\mu$ may be computed on I as

$$f^*\mu = f^*(\varphi^*m) = \varphi^*(g^*m) = \varphi^*(\lambda m) = \lambda\varphi^*m = \lambda\mu.$$

But I was an arbitrary P-basic interval. Therefore $T_f \mu = \lambda\mu$.

Conversely, assume that there exists a probability measure $\mu \in \mathcal{M}$ such that $T_f\mu = \lambda\mu$. Define a map φ by $\varphi(x) = \mu([0, x])$. This map is continuous, nondecreasing, and maps $[0, 1]$ onto $[0, 1]$. To see that φ induces a well-defined factor map g, suppose that $x_1 < x_2$ and $\varphi(x_1) = \varphi(x_2)$. If the interval $[x_1, x_2]$ contains a point of P, then $\varphi(x_i)$ belongs to $\varphi(P)$, which will be a g-admissible set, so there is no need to define g at $\varphi(x_i)$. Otherwise, the interval $[x_1, x_2]$ is contained in a P-basic set I, $\mu([x_1, x_2]) = 0$, and thus

$$\mu(f([x_1, x_2])) = (T_f\mu)([x_1, x_2]) = \lambda\mu([x_1, x_2]) = 0.$$

Therefore $\varphi(f(x_1)) = \varphi(f(x_2))$. This shows that we can define our map g by the equation $g(\varphi(x)) = \varphi(f(x))$, and such g will be monotone and continuous on every $g(P)$-basic interval. By construction, $\varphi_*\mu$ is the Lebesgue measure m. It remains to consider $T_g m$. Using the same restricted maps and measures as in diagram (1.2.1), we get

$$g^*m = g^*(\varphi_*\mu) = \varphi_*(f^*\mu) = \varphi_*(\lambda\mu) = \lambda\varphi_*\mu = \lambda m.$$

Since I was arbitrary, this shows that $T_g m = \lambda m$. By Lemma 1.2, g has constant slope λ. ∎

Remark 1.5 *If $f \in \mathcal{CC}$, then the only way we could get a discontinuity of g in the above construction was when $x_1 < x_2$, $\varphi(x_1) = \varphi(x_2)$, and there is a point of P between x_1 and x_2. However, then*

$$\mu(f([x_1, x_2])) = \mu(\bigcup_{I \in \mathcal{B}(P)} f(I \cap [x_1, x_2])) \leq \sum_{I \in \mathcal{B}(P)} \mu(f(I \cap [x_1, x_2])) =$$
$$= (T_f\mu)([x_1, x_2]) = \lambda\mu([x_1, x_2]) = 0$$

By the continuity of f, the set $f([x_1, x_2])$ includes the interval with endpoints $f(x_1)$, $f(x_2)$. Therefore $\varphi(f(x_1)) = \varphi(f(x_2))$, so no discontinuity is created. This shows that Theorem 1.4 holds with $\mathcal{C}$ replaced by $\mathcal{CC}$, for both f and g.

Remark 1.6 *With only slight modifications in wording, the proof of Theorem 1.4 and the considerations in Remark 1.5 go through for circle maps.*

1.3   No semiconjugacy

Now we want to find conditions that prevent semiconjugacy to a map of constant slope. We start with a technical lemma. One of our assumptions is that $\lambda > 2$. For piecewise monotone maps with finite number of pieces this type of an assumption is usually circumvented by taking a sufficiently high iterate of the map. If $\lambda > 1$ then for some large n we get $\lambda^n > 2$. However, here we have another assumption, that the measures of P-basic intervals are bounded away from 0, and taking an iterate of a map could lead to this condition being violated.

Lemma 1.7 *Let $f \in \mathcal{C}$. Suppose that there exist $\lambda > 2$, $\delta > 0$, $\mu \in \mathcal{M}$ and an f-admissible set P such that $T_f \mu = \lambda \mu$ and the measure of every P-basic interval I satisfies $\delta \leq \mu(I) < \infty$. Then for μ almost every x in $[0,1]$ there exist infinitely many times $n_1 < n_2 < \ldots$ such that x belongs to an interval which is mapped monotonically by f^{n_k} to an interval of μ-measure at least δ.*

Proof   Fix an arbitrarily large natural number N and choose an arbitrary P^N-basic interval J. A P^N-basic interval means a component of the complement of $\bigcup_{i=0}^{N-1} f^{-i}(P)$; thus, f^N is monotone and continuous on each P^N-basic interval. It suffices to prove that for μ almost every x in J there exists a time $n \geq N$ and a P^n-basic interval $L \subset J$ such that $x \in L$ and $\mu(f^n(L)) \geq \delta$.

If $\mu(f^N(J)) \geq \delta$, then we are done. Otherwise, J is a "bad" interval, and we subdivide it at all the points of intersection $J \cap f^{-N}(P)$ into P^{N+1}-basic intervals, which we classify as either "good" or "bad" according as $\mu(f^{N+1}(L))$ is either at least δ or smaller than δ, respectively. For points x in the good P^{N+1}-basic intervals, the claim holds. But if any of these intervals L is bad, we subdivide it further at all the points of intersection $L \cap f^{-(N+1)}(P)$ into P^{N+2}-basic intervals, which we then classify as good or bad, and so on.

To be more precise, we define $\mathcal{B}_0 = \{J\}$ and we recursively define

$$\mathcal{B}_{i+1} = \{M \in \mathcal{B}(P^{N+i+1}) : \mu(f^{N+i+1}(M)) < \delta \text{ and } \exists L \in \mathcal{B}_i, M \subset L\}.$$

Now observe that each bad interval L at stage i subdivides into at most two bad intervals at stage $i + 1$, because $\#(L \cap f^{-(N+i)}(P)) = \#(f^{N+i}(L) \cap P)$, and by hypothesis, an interval of measure less than δ never contains more than one point of P. It follows that $\#\mathcal{B}_i \leq 2^i$. The hypothesis $T_f\mu = \lambda\mu$ means that wherever f is monotone and continuous, it stretches μ measures uniformly by the factor λ. This provides an upper bound on the measures of bad intervals. If $L \in \mathcal{B}_i$, then

$$\mu(L) = \lambda^{-(N+i)}\mu(f^{N+i}(L)) \leq \lambda^{-(N+i)}\delta.$$

It follows that

$$\sum_{L \in \mathcal{B}_i} \mu(L) \leq \left(\frac{2}{\lambda}\right)^i \lambda^{-N}\delta,$$

and this quantity tends to zero as $i \to \infty$. Therefore almost every point of J falls at some stage of the process into a good basic interval, and this proves our claim. ∎

The following lemma is an analog of Lebesgue's density theorem. While it is known, it is difficult to find in the literature the statement we want. Usually statements with one-sided neighborhoods are only about the Lebesgue measure, while statements about more general measures use balls around the density point. Therefore we show how to deduce what we need from the statement about the Lebesgue measure.

Lemma 1.8 *Let $\mu \in \mathcal{M}$ and let A be a Borel set. Then for μ almost every $x \in A$ the measures of all one-sided neighborhoods of x are positive, and*

$$\lim_{\delta \searrow 0} \frac{\mu(A \cap [x, x + \delta))}{\mu([x, x + \delta))} = \lim_{\delta \searrow 0} \frac{\mu(A \cap (x - \delta, x])}{\mu((x - \delta, x])} = 1. \tag{1.3.1}$$

Proof Let P be a countable closed subset of $[0, 1]$ such that the measure of every P-basic interval is finite. Let L be a P-basic interval, and let a denote the left endpoint of L. It suffices to prove the claim for μ almost every x in $A \cap L$; therefore we may restrict everything to L.

There may exist closed subintervals of L of measure zero. There are countably many of such maximal intervals, so their union has measure zero. Hence, we are free

to remove this union, as well as the endpoints of L, from A. Then, for each $x \in A$ and each positive δ, the measures $\mu([x, x+\delta))$ and $\mu((x-\delta, x])$ are nonzero. Since we restrict everything to L, and $\mu(L)$ is finite, we see that the ratios under consideration have finite numerators and denominators, and so are well-defined.

Introduce a map $\varphi : L \to Y$, where $Y = [0, \mu(L)]$, by $\varphi(x) = \mu((a, x))$ By construction, φ is nondecreasing. Moreover, φ is continuous because μ is nonatomic. By the definition of φ, every interval $I \subset Y$ enjoys the property $m(I) = \mu(\varphi^{-1}(I))$. It follows that this property holds for all measurable sets $I \subset Y$. Therefore

$$\lim_{\delta \searrow 0} \frac{\mu(A \cap [x, x+\delta))}{\mu([x, x+\delta))} = \lim_{\eta \searrow 0} \frac{m(\varphi(A) \cap [\varphi(x), \varphi(x) + \eta))}{m([\varphi(x), \varphi(x) + \eta))} \qquad (1.3.2)$$

and

$$\lim_{\delta \searrow 0} \frac{\mu(A \cap (x-\delta, x])}{\mu((x-\delta, x])} = \lim_{\eta \searrow 0} \frac{m(\varphi(A) \cap (\varphi(x) - \eta, \varphi(x)])}{m((\varphi(x) - \eta, \varphi(x)])}. \qquad (1.3.3)$$

The preimage under φ of a set of full Lebesgue measure in Y has full μ measure in L. By the Lebesgue density theorem (see, e.g., [10]), the limits of the right-hand sides of (1.3.2) and (1.3.3) are 1 for Lebesgue almost all $x \in \varphi(A)$. Therefore (1.3.1) holds for μ almost all $x \in A$. ∎

In the next theorem we need an assumption stronger than transitivity. For a continuous map f on a topological space X, the usual definition of (topological) transitivity is that for every pair of nonempty open sets U and V in X, there is a positive integer k such that $f^k(U) \cap V \neq \emptyset$. The term "strong transitivity" is sometimes used for the stronger property that for every nonempty open set U, the union $\bigcup_{n=0}^{\infty} f^n(U)$ is the whole space X. Let us make an appropriate modification of this notion for the class $\mathcal{C}$ in which countable closed sets are negligible. We will say that a map $f \in \mathcal{C}$ is *substantially transitive* if for every nonempty open set $U \in [0, 1]$ the set $[0, 1] \smallsetminus \bigcup_{n=0}^{\infty} f^n(U)$ has countable closure.

If $f \in \mathcal{CC}$ is transitive, we get substantial transitivity automatically. We will need this later also for continuous circle maps, so we will state a lemma for graph maps. Again, this lemma is known, but it is easier to prove it than to look for it in the literature.

Lemma 1.9 *Let X be a graph and let f be a topologically transitive continuous map of X to X. Let U be a nonempty open subset of X. Then the set $X \smallsetminus \bigcup_{n=0}^{\infty} f^n(U)$ is finite.*

Proof By replacing U by one of its connected components, we may assume that U is connected. Let V be the connected component of the set $W = \bigcup_{n=0}^{\infty} f^n(U)$ that contains U. By transitivity, there is N such that $f^N(V) \cap V \neq \emptyset$. However, V is a component of a forward invariant set, so $f^N(V) \subset V$. It follows that $W = \bigcup_{n=0}^{N-1} f^n(V)$, and thus W has only finitely many connected components. On the other hand, by transitivity W is dense in X. Therefore W excludes only finitely many points of X. ∎

Remark 1.10 *If $X = \mathbb{R}$ and f is a lifting of a degree one circle map, the same proof shows that $\bigcup_{n=0}^{\infty} f^n(U) = \mathbb{R}$. For such f there is a constant $M > 0$ such that $|f(x) - x| < M$ for every $x \in \mathbb{R}$, so W is contained in the set of points whose distance from V is smaller than NM. Since V is connected and W is dense, we must have $V = \mathbb{R}$.*

Now we can prove the main technical result of the chapter.

Theorem 1.11 *Let $f \in \mathcal{C}$ be a substantially transitive map and let $\lambda > 2$. Assume that there exist $\delta > 0$, an infinite measure $\mu \in \mathcal{M}$, and an f-admissible set P, such that $T_f \mu = \lambda \mu$ and the measure of every P-basic interval I satisfies $\delta \leq \mu(I) < \infty$. Then there is no probability measure $\nu \in \mathcal{M}$ such that $T_f \nu = \lambda \nu$.*

Proof Suppose that such measure ν exists. The measure $\mu + \nu$ is an infinite measure in $\mathcal{M}$ such that $T_f(\mu + \nu) = \lambda(\mu + \nu)$ and for every P-basic interval I we have $\delta \leq (\mu + \nu)(I) < \infty$. Replace μ by $\mu + \nu$ if necessary to obtain absolute continuity of ν with respect to μ. Now take the Radon-Nikodym derivative, $\xi = \frac{d\nu}{d\mu}$ and write

$d\nu = \xi d\mu$. Integrate the function $\xi \circ f$ over any Borel set A contained in any P-basic interval I:

$$\int_A \xi \circ f \; d\mu|_I = \int_{f(A)} \xi \; df_*(\mu|_I) = \frac{1}{\lambda} \int_{f(A)} \xi \; d\mu|_{f(I)} = \frac{1}{\lambda} \int_{f(A)} d\nu|_{f(I)}$$

$$= \int_{f(A)} df_*(\nu|_I) = \int_A 1 \circ f \; d\nu|_I = \int_A d\nu|_I = \int_A \xi \; d\mu|_I.$$

This shows that the equality $\xi \circ f = \xi$ holds μ almost everywhere; that is, that up to a set of μ measure zero, the function ξ is constant along the orbits of f.

By the definition of the Radon-Nikodym derivative, $\int_0^1 \xi \; d\mu = \nu([0,1]) = 1$. Therefore there exists a positive real number ε such that the measurable set $E = \xi^{-1}([\varepsilon, \infty))$ has positive μ measure. This measure cannot be infinite, because then $\int_0^1 \xi \; d\mu$ would be infinite. Thus, $0 < \mu(E) < \infty$. Because ξ is constant along orbits, this set E is fully invariant; that is, $f^{-1}(E) = E$. While this is μ almost everywhere, we can modify E by adding/subtracting a set of μ measure zero so that it holds everywhere.

The plan of the rest of the proof is as follows. We use Lemma 1.8 to get high density of E in a small interval, then Lemma 1.7 to transport it to a long interval, and then substantial transitivity to transport it to the whole space. By the invariance of E this construction gives us infinite measure of E, which is impossible.

There is a point $x \in E$ which satisfies conclusions of both Lemmas 1.7 and 1.8 (with $A = E$). In particular, there exist a sequence of times n_k and a sequence of intervals $L_k = [a_k, b_k]$ containing x such that f^{n_k} is monotone on L_k and such that $\mu(f^{n_k}(L_k)) \geq \delta$. Trimming the intervals L_k (but keeping $x \in L_k$), we can achieve equality $\mu(f^{n_k}(L_k)) = \delta$. Therefore $\mu(L_k) = \lambda^{-n_k}\delta$, and this decreases to zero. But every neighborhood (both two-sided and one-sided) of x has positive μ measure. Therefore, $a_k \to x$ and $b_k \to x$ as $k \to \infty$. Now we may use the fact that x is a density point of E to conclude that $\mu(E \cap L_k)/\mu(L_k) \to 1$.

Next, we show the density of E in an interval at the large scale. By compactness, after passing to subsequences we may assume that $f^{n_k}(a_k)$ converges to some point $a \in [0,1]$ and $f^{n_k}(b_k)$ converges to some point $b \in [0,1]$. Let L denote the interval

$[a, b]$. We claim that the points a and b are distinct. Indeed, if $a = b$, then for each j take the interval $\left(a - \frac{1}{j}, a + \frac{1}{j}\right)$. It contains infinitely many intervals $f^{n_k}(L_k)$ and $\mu(E \cap f^{n_k}(L_k))$ is approaching δ, so the measure of $E \cap \left(a - \frac{1}{j}, a + \frac{1}{j}\right)$ is at least δ. Since the measure of E is finite, we may send j to infinity and find by the continuity of measure that there is an atom at a, which is a contradiction. Thus $a \neq b$. As k grows, the endpoints $f^{n_k}(a_k), f^{n_k}(b_k)$ of $f^{n_k}(L_k)$ eventually draw nearer to the respective endpoints a, b of L than half the distance between a and b. Therefore, for sufficiently large k, the symmetric difference $f^{n_k}(L_k) \triangle L$ consist of two intervals; one with endpoints $a, f^{n_k}(a_k)$, and the other with endpoints $b, f^{n_k}(b_k)$. Again by the continuity of measure, each of these intervals has μ measure converging to zero. Therefore $\mu(L \triangle f^{n_k}(L_k)) \to 0$ as $k \to \infty$. Together with the invariance of E and the monotonicity of f^{n_k} on L_k, this shows that

$$\frac{\mu(E \cap L)}{\mu(L)} \geq \lim_{k \to \infty} \frac{\mu(E \cap f^{n_k}(L_k)) - \mu(f^{n_k}(L_k) \triangle L)}{\mu(f^{n_k}(L_k)) + \mu(f^{n_k}(L_k) \triangle L)} = \lim_{k \to \infty} \frac{\mu(E \cap f^{n_k}(L_k))}{\mu(f^{n_k}(L_k))}$$

$$= \lim_{k \to \infty} \frac{\mu(f^{n_k}(E \cap L_k))}{\mu(f^{n_k}(L_k))} = \lim_{k \to \infty} \frac{\lambda^{n_k} \cdot \mu(E \cap L_k)}{\lambda^{n_k} \cdot \mu(L_k)} = 1.$$

Therefore $E \cap L$ has full measure in the interval L; that is, $\mu(L \smallsetminus E) = 0$.

If the invariant set E fills L, then it must also fill all the images $f^n(L)$, $n \in \mathbb{N}$. Indeed, $\mu(f(L) \smallsetminus E) = \mu(f(L \smallsetminus E))$ by the f-invariance of E. However,

$$\mu\big(f(L \smallsetminus E)\big) \leq \sum_{I \in \mathcal{B}(P)} \mu\big(f(I \cap (L \smallsetminus E))\big) = (T_f \mu)(L \smallsetminus E) = \lambda \mu(L \smallsetminus E) = 0.$$

This shows that $\mu(f(L) \smallsetminus E) = 0$, and it follows inductively that $\mu(f^n(L) \smallsetminus E) = 0$ for all $n \in \mathbb{N}$.

The interval L has nonempty interior, so by substantial transitivity of f the set $\bigcup_{n=0}^{\infty} f^n(L)$ excludes at most countably many points of $[0, 1]$, and hence has full μ measure in $[0, 1]$. But E has full measure in $\bigcup_{n=0}^{\infty} f^n(L)$, and therefore E has full measure in $[0, 1]$. This is a contradiction because E has finite μ measure, while by the assumption, $\mu([0, 1]) = \infty$. Therefore it is impossible for a probability measure $\nu \in \mathcal{M}$ to satisfy $T_f \nu = \lambda \nu$. ∎

1.4   Liftings of circle maps

While Theorem 1.11, together with Theorem 1.4, provides sufficient conditions for a map $f \in \mathcal{C}$ to have no semiconjugacy to a map of constant slope, it is not immediately obvious how to construct concrete examples. In particular, even if we use those theorems to exclude semiconjugacy to a map with a given slope λ, how can we exclude other λ's? In this section we provide tools to do it for a large class of maps. Those maps additionally will be continuous (formally, they will be restrictions of continuous maps to $[0,1] \smallsetminus P$). We denote the circle $\mathbb{R}/\mathbb{Z}$ by $\mathbb{R}/\mathbb{Z}$.

Theorem 1.12 *Assume that $f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ is a continuous degree one map that is piecewise monotone with finitely many pieces and has constant slope $\lambda > 1$. Assume also that f has a lifting $F : \mathbb{R} \to \mathbb{R}$ that is topologically transitive. Take any continuous interval map $g : [0,1] \to [0,1]$ such that $g|_{(0,1)}$ is topologically conjugate to F. Then there does not exist any nondecreasing semiconjugacy of g to an interval map of constant slope.*

Proof   Fix n sufficiently large so that $\lambda^n > 2$ and consider the iterates g^n, f^n, and F^n. The map $g^n|_{(0,1)}$ is conjugate to F^n and F^n is a lifting of the degree one circle map f^n of constant slope λ^n. A priori, a transitive map need not have transitive iterates. But F is a transitive map on the real line, and therefore either all iterates of F are transitive, or else there exists a point $y \in \mathbb{R}$ such that $F((-\infty, y]) = [y, \infty)$ and $F([y, \infty)) = (-\infty, y]$ (see [3, pgs 156-7] – the statements are for interval maps but the proofs also hold in $\mathbb{R}$). This latter alternative is impossible for a lifting of a degree one circle map. Therefore F^n is transitive. If there exists any nondecreasing map that semiconjugates g with a constant slope interval map, then the same map also conjugates g^n with a constant slope interval map. This shows that after replacing g, f, and F by some suitably high iterates, we may assume that $\lambda > 2$.

Let $h : (0,1) \to \mathbb{R}$ denote the homeomorphism that conjugates $g|_{(0,1)}$ with F. Let $\pi : \mathbb{R} \to \mathbb{R}/\mathbb{Z}$, denote the natural projection (that semiconjugates F with f).

Moreover, let $s : \mathbb{R} \to \mathbb{R}$, given by $s(x) = x + 1$, denote the deck transformation; then $s \circ F = F \circ s$. Thus, the following diagrams commute.

$$
\begin{array}{ccc}
(0,1) \xrightarrow{\ g\ } (0,1) & \mathbb{R} \xrightarrow{\ F\ } \mathbb{R} & \mathbb{R} \xrightarrow{\ F\ } \mathbb{R} \\
\ \downarrow{h} \qquad\quad \downarrow{h} & \ \downarrow{\pi} \qquad\quad \downarrow{\pi} & \ \downarrow{s} \qquad\quad \downarrow{s} \\
\mathbb{R} \xrightarrow[\ F\ ]{} \mathbb{R} & \mathbb{R}/\mathbb{Z} \xrightarrow[\ f\ ]{} \mathbb{R}/\mathbb{Z} & \mathbb{R} \xrightarrow[\ F\ ]{} \mathbb{R}
\end{array}
$$

The hypothesis of piecewise monotonicity means that there exists a finite set $P \subset \mathbb{R}/\mathbb{Z}$ such that f is monotone on each P-basic arc, and because of the constant slope, that monotonicity is strict. In the circle, strict monotonicity does not guarantee injectivity, but after adjoining the finite set $f^{-1}(x)$ for some $x \in \mathbb{R}/\mathbb{Z}$ to P we have also injectivity of f on each P-basic arc, so that the restriction of f to any P-basic arc is then a homeomorphism onto its image. Just as for interval maps we define the operator T_f acting on the space of nonatomic, strongly σ-finite, Borel measures on the circle. Let $P_F = \pi^{-1}(P)$. Then P_F is a closed, countable set, invariant under the integer translation map s, and F is strictly monotone on each P_F-basic interval. Let $P_g = h^{-1}(P_F) \cup \{0, 1\}$. Then $g \in \mathcal{CC}$ and P_g is a g-admissible set.

Suppose that there is a nondecreasing semiconjugacy of g to a constant slope interval map, say, with slope λ'. Then by Theorem 1.4 there exists a probability measure $\nu_g \in \mathcal{M}$ such that $T_g \nu_g = \lambda' \nu_g$. Push this measure down to a measure $\nu_F = h_*(\nu_g)$ on the real line. Then h gives not only a topological conjugacy, but also a measure-theoretic isomorphism of $((0,1), g, \nu_g)$ with $(\mathbb{R}, F, \nu_F)$. It follows that $T_F \nu_F = \lambda' \nu_F$.

Now push this measure down to the circle, defining $\nu_f = \pi_* \nu_F$. If A is a Borel subset of a P-basic arc in $\mathbb{R}/\mathbb{Z}$, then its preimage in the covering space $\mathbb{R}$ can be expressed as a disjoint union $\pi^{-1}(A) = \bigcup_{n=-\infty}^{\infty} s^n(B)$ in such a way that B is a subset of a P_F-basic interval. Then for each $n \in \mathbb{N}$, $s^n(B)$ is also a subset of a P_F-basic interval, because P_F is s-invariant. By the injectivity of f on each P-

basic arc, it follows that the sets $F(s^n(B)), n \in \mathbb{N}$, are pairwise disjoint. Therefore $\pi^{-1}(f(A)) = \bigcup_{n=-\infty}^{\infty} F(s^n(B))$ is also a disjoint union. Now we can calculate

$$(T_f \nu_f)(A) = \nu_f(f(A)) = \nu_F(\pi^{-1}(f(A))) = \nu_F\left(\bigcup_{n=-\infty}^{\infty} F(s^n(B))\right) =$$

$$= \sum_{n=-\infty}^{\infty} \nu_F(F(s^n(B))) = \sum_{n=-\infty}^{\infty} \lambda' \nu_F(s^n(B)) = \lambda' \nu_F\left(\bigcup_{n=-\infty}^{\infty} s^n(B)\right) = \lambda' \nu_f(A).$$

It follows that $T_f \nu_f = \lambda' \nu_f$.

By Theorem 1.4 and Remarks 1.5 and 1.6, there is a nondecreasing semiconjugacy of f to a circle map of constant slope λ'. But f, being a factor of a transitive map, is itself transitive. Therefore a nondecreasing semiconjugacy is automatically a conjugacy (see [1]). Thus, f is conjugate to a circle map of constant slope λ'. The topological entropy of a constant slope circle map is the logarithm of the slope (see [16], [1]), and topological entropy is a conjugacy invariant. In such a way we have shown that if there exists a nondecreasing semiconjugacy of g to an interval map of constant slope λ', then in fact $\lambda' = \lambda$, the constant slope of f.

Let us push the Lebesgue measure m on $\mathbb{R}$ via the homeomorphism h^{-1}, and denote $\mu = (h^{-1})_*(m)$. The μ measures of P_g-basic intervals are the same as m measures of corresponding P_F-basic intervals. Since F is a lifting of a piecewise monotone map with finite number of pieces, those measures take only finitely many values, all of them finite. Moreover F, as the lifting of a map of constant slope, has also constant slope. By Lemma 1.2 and Remark 1.3, the Lebesgue measure m satisfies $T_F m = \lambda m$. Since h is a measure-theoretic ismorphism of $((0,1), g, \mu)$ with $(\mathbb{R}, F, m)$ it follows that $T_g \mu = \lambda \mu$. Substantial transitivity of g follows from transitivity of F and Lemma 1.9. All this shows that μ belongs to $\mathcal{M}$ and that g and μ satisfy the assumptions of Theorem 1.11. Therefore there is no probability measure $\nu \in \mathcal{M}$ with $T_g \nu = \lambda \nu$, and consequently, there is no nondecreasing semiconjugacy of g to a constant slope interval map. ∎

Remark 1.13 *In Theorem 1.12 there is no difficulty in finding a continuous interval map* $g : [0,1] \to [0,1]$ *such that* $g|_{(0,1)}$ *is topologically conjugate to* F. *Let* h *denote*

any homeomorphism of $(0, 1)$ with $\mathbb{R}$ and define $g = h^{-1} \circ F \circ h$ with additional fixed points at 0 and 1. We obtain continuity of g at the points $0, 1$, because F was assumed to be the lifting of a degree one circle map.

Remark 1.14 *There are trivial examples of maps with zero topological entropy for which there is no semiconjugacy to a map of constant slope, for instance, the map $f : [0, 1] \to [0, 1]$ given by $f(x) = x^2$. Theorem 1.12 is nontrivial in that it applies to continuous and transitive interval maps, and by [4] [5], such maps always have topological entropy at least $\log \sqrt{2}$.*

If we wish to construct explicit examples that satisfy the hypotheses of Theorem 1.12, the only possible difficulty is in verifying the transitivity of the lifting F. Fortunately, there is a simple condition, broadly applicable and easy to verify, that guarantees transitivity.

Theorem 1.15 *Assume that $f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ is a continuous degree one map that is piecewise monotone with finitely many pieces and has constant slope. Assume also that $F : \mathbb{R} \to \mathbb{R}$ is a lifting of f. Let P denote the set of turning points of F. If for each P-basic interval I there are points x_L, x_R in the closure of I such that $F(x_L) = x_L - 1$ and $F(x_R) = x_R + 1$, then F is topologically transitive.*

Proof The sets $R := \{x \in \mathbb{R} : F(x) - x = 1\}$ and $L := \{x \in \mathbb{R} : F(x) - x = -1\}$ are both invariant under integer translations and are both nonempty by hypothesis. Choose a point $x_L \in L$ and let x_R be the smallest element of R that is larger than x_L. Then $x_R - x_L < 1$ and $F(x_R) - F(x_L) > 2$. Since F has constant slope λ, this shows that $\lambda > 2$.

Let $U \subset \mathbb{R}$ be any open interval. As n grows, the successive images $F^n(U)$ grow in length by a factor at least $\lambda/2 > 1$ until some image $F^N(U)$ contains an entire P-basic interval. Within the closure of this P-basic interval there are points $x_L \in L$, $x_R \in R$. Then, in the next steps, $F^{N+1}(U)$ contains $(x_L - 1, x_R + 1)$, $F^{N+2}(U)$ contains $(x_L - 2, x_r + 2)$, and so on. Therefore the union of all images of U is all of $\mathbb{R}$, and this proves transitivity of F. ∎

Remark 1.16 *We can immediately verify the hypothesis of Theorem 1.15 by super-imposing the diagonal lines $y = x + 1$, $y = x - 1$ on the graph $y = F(x)$. Each piece of monotonicity of the graph of F should intersect both diagonal lines.*

1.5 Examples

In this section we provide a concrete example of a one-parameter family of circle maps of degree one with transitive liftings and constant slope. In such a way we will have examples where our theorems apply and there is no nondecreasing semiconjugacy to a map of a constant slope.

Let us describe a lifting F_λ in our family. Choose a real parameter $\lambda \geq 2 + \sqrt{5}$. Let F_λ be the "connect the dots" map (the graph of F_λ consists of straight line segments connecting the dots) with the dots $(k, k-1)$ and $(k+b, k+c)$, where $k \in \mathbb{Z}$, $b = (\lambda + 1)/2\lambda$, and $c = (\lambda - 1)/2$ (see Figure 1.1). On the interval $[k, k+b]$ the slope is $(c+1)/b = \lambda$, and on $[k+b, k+1]$ it is $-c/(1-b) = -\lambda$, so the map has constant slope λ. We have

$$F_\lambda(k + b) - (k + b) - 1 = \frac{\lambda^2 - 4\lambda - 1}{2\lambda} = \frac{\left(\lambda - \left(2 + \sqrt{5}\right)\right)\left(\lambda - \left(2 - \sqrt{5}\right)\right)}{2\lambda} \geq 0,$$

and therefore $F_\lambda(k + b) - (k + b) \geq 1$. Moreover, $F_\lambda(k) - k = -1$, so by the Intermediate Value Theorem the assumptions of Theorem 1.15 are satisfied. Thus, F_λ is topologically transitive. Now if we choose any homeomorphism $h : (0, 1) \to \mathbb{R}$, we will get a map $g_\lambda = h^{-1} \circ F_\lambda \circ h$ (with additional fixed points at 0 and 1), which belongs to $\mathcal{CC}$, but is not semiconjugate by a nondecreasing map to a map of constant slope. If we want really concrete examples, we can even specify h, for instance $h(x) = \ln(x/(1 - x))$ (then $h^{-1}(x) = e^x/(e^x + 1)$).

We would like to have in our family both maps that are and are not Markov. Remember that "Markov" means countably Markov, so F_λ being Markov means that for the corresponding circle map the trajectory of the local maximum has countable closure (the local minimum is always a fixed point). Of course F_λ is Markov if and only if the map g_λ is Markov.
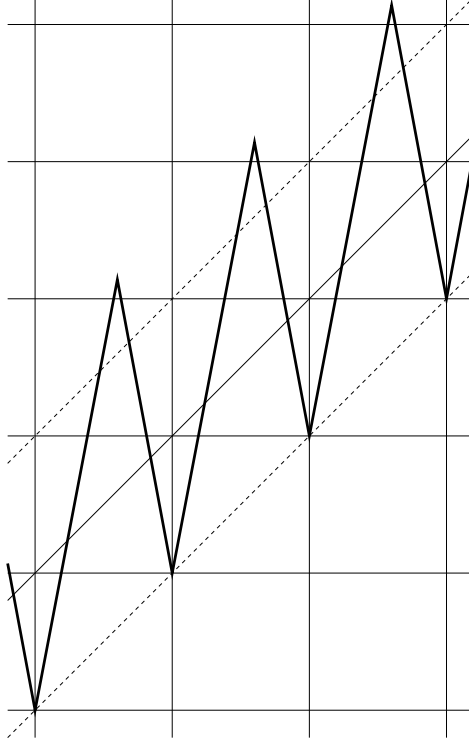
Figure 1.1. The Map F_λ with $\lambda = 5.28$

We start with a lemma on the one-sided full 2-shift $\sigma : \Sigma \to \Sigma$, where $\Sigma = \{0, 1\}^{\mathbb{N}}$.

Lemma 1.17 *Let D be the set of those points $s \in \Sigma$ for which the closure of the trajectory $\{\sigma^n(s)\}_{n=0}^{\infty}$ is countable. Then both sets D and $\Sigma \smallsetminus D$ are uncountable.*

Proof Each element of Σ is a 0-1 sequence. Let E be the set of those sequences that are built of alternating blocks of 0's and 1's, and the length of the n-th block is n or $n + 1$. Since we have to choose between the lengths n and $n + 1$ for each n, the set E is uncountable. We claim that $E \subset D$. Fix an element $s \in E$. The trajectory of s is of course countable. It remains to count the accumulation points of this trajectory (that is, of the ω-set of s). If we fix the size of a window and slide it sufficiently far to the right along the sequence s, we see in this window only one or two blocks. This means that every element of the ω-limit set of s will consist of one or two blocks.

However, there are only countably many such sequences. This proves our claim, and hence D is uncountable.

The 2-shift is transitive, and therefore the set of points with dense trajectories contains a dense G_δ set G. If G is countable, then for every $s \in G$ the set $\Sigma \setminus \{s\}$ is open and dense, so by the Baire Category Theorem the intersection of all those sets, $\Sigma \setminus G$, is a dense G_δ set. Therefore $(\Sigma \setminus G) \cap G$ is also a dense G_δ set, but it is empty. This contradiction shows that G is uncountable. Since for every $s \in G$ the closure of the trajectory of S is Σ, we have $G \subset \Sigma \setminus D$, and thus $\Sigma \setminus D$ is uncountable. ∎

Theorem 1.18 *Fix an integer $n \geq 2$. Then there are uncountable sets $\Lambda_M \subset [2n+1, 2n+3]$ and $\Lambda_{nM} \subset [2n+1, 2n+3]$ such that for every $\lambda \in \Lambda_M$ the lifting F_λ is Markov and for every $\lambda \in \Lambda_{nM}$ the lifting F_λ is not Markov.*

Proof Let A_λ be the set of those points x such that $F_\lambda^i(x) \in [0,1]$ for $i = 0, 1, 2, \ldots$. Perform the standard coding procedure, using the left and right subintervals of $[0,1] \cap F_\lambda^{-1}([0,1])$. It shows that A_λ is a Cantor set, and F_λ restricted to this set is conjugate to the one-sided full 2-shift. By the standard argument, for any given itinerary the corresponding point of A_λ depends continuously on λ. By Lemma 1.17, uncountably many itineraries correspond to points whose trajectories have countable closures, and uncountably many itineraries correspond to points whose trajectories have uncountable closures. As λ varies from $2n+1$ to $2n+3$ then the image $-n + (\lambda - 1)/2$ under F_λ of the local maximum $-n + (\lambda + 1)/(2\lambda)$ sweeps the interval $[0,1]$. When it meets a point with a countable closure of the trajectory, the corresponding map F_λ is Markov; when it meets a point with an uncountable closure of the trajectory, it is not Markov. This completes the proof. ∎

2. CONJUGACY TO A CONSTANT SLOPE MAP ON THE EXTENDED REAL LINE

In Theorem 1.12 we identify a class of transitive interval maps which are not conjugate to any interval map of constant slope. These maps are constructed, however, from constant slope maps on the real line. This raises a natural follow-up question. What happens if we dispense with the requirement that our constant slope maps act on a finite-length interval? Can we make a more systematic study of constant slope maps when the underlying space is allowed to have infinite length?

2.1 Definitions and Background

The extended real line $[-\infty, \infty]$ is the ordered set $\mathbb{R} \cup \{\infty, -\infty\}$ equipped with the order topology; this topological space is a two-point compactification of the real line and is homeomorphic to the closed unit interval $[0, 1]$.

Suppose f is a continuous self-map of some interval $[a, b]$, $-\infty \le a < b \le \infty$, and suppose there exists a closed, countable set $P \subseteq [a, b]$, $a, b \in P$, such that $f(P) \subseteq P$ and f is monotone on each component of $[a, b] \setminus P$. Such a map is said to be *countably piecewise monotone and Markov* with respect to P, the components of $[a, b] \setminus P$ are called P-*basic intervals*, and the set of all P-basic intervals is denoted $\mathcal{B}(P)$. If additionally the restriction of f to each P-basic interval is affine with slope of absolute value λ, then we say that f has *constant slope* λ. This is a geometric, rather than a topological property, and it is the reason we must distinguish finite from infinite length intervals. The class of all countably piecewise monotone and Markov maps is denoted $\mathcal{CPMM}$. The subclass of those maps which act on the closed unit interval $[0, 1]$ is denoted $\mathcal{CPMM}_{[0,1]}$.

This class $\mathcal{CPMM}$ is different from the class of interval maps $\mathcal{C}$ of Chapter 1 in three ways. First, the underlying interval $[a, b]$ depends on the map f and is permitted to be infinite in length. Second, the map f is required to be continuous; this is essential for our use of the intermediate value theorem. And third, the set P is required to be forward-invariant. This is the Markov condition; it means that if I, J are P-basic intervals and $f(I) \cap J \neq \emptyset$, then $f(I) \supseteq J$.

If f is countably piecewise monotone and Markov with respect to P, then we define the binary *transition matrix* $T = T(f, P)$ with rows and columns indexed by $\mathcal{B}(P)$ and entries

$$T(I, J) = \begin{cases} 1, & \text{if } f(I) \supseteq J \\ 0, & \text{otherwise.} \end{cases}$$

If we only allow for constant slope maps on a finite length interval, say, $[0, 1]$, then there is an established necessary and sufficient condition to determine when a map is semiconjugate to a map of constant slope.

Theorem 2.1 *(Bobok, [6]) Let $f \in \mathcal{CPMM}_{[0,1]}$ with transition matrix T, and fix $\lambda > 1$. Then f is semiconjugate via a continuous nondecreasing map ψ to some map $g \in \mathcal{CPMM}_{[0,1]}$ of constant slope λ if and only if T has a nonnegative eigenvector $v = (v_I) \in \ell^1(\mathbb{R}^{\mathcal{B}(P)})$ with eigenvalue λ.*

To be clear, the notation $v \in \ell^1(\mathbb{R}^{\mathcal{B}(P)})$ means that we require the eigenvector to be summable. If we read the proof in [6], the reason for this is clear. If we are given the semiconjugacy ψ to the constant slope map, then we construct the eigenvector v by setting $v_I = |\psi(I)|$ for each P-basic interval I, where $|\cdot|$ denotes the length of an interval, and therefore the sum of the entries v_I is just the length of the unit interval $[0, 1]$. Conversely, if we are given an eigenvector v, then we rescale it so that the sum of entries is 1 and then construct the semiconjugacy in such a way that $|\psi(I)| = v_I$ for all I, obtaining a map g of an interval of length 1.

2.2 Eigenvector Criterion

We return now to the question, when does a map $f \in \mathcal{CPMM}$ admit a nondecreasing semiconjugacy ψ to a map $g \in \mathcal{CPMM}$ of constant slope on any compact subinterval of $[-\infty, \infty]$, whether finite or infinite in length? It is clear that g must belong to the class $\mathcal{CPMM}$, because g will necessarily be piecewise monotone and Markov with respect to $\psi(P)$ – see [1, Lemma 4.6.1]. To avoid pathological examples, we demand that f be topologically mixing, that is, that for every pair of nonempty open sets U, V there is $N \in \mathbb{N}$ such that for all $n \geq N$, $U \cap f^{-n}(V) \neq \emptyset$. Without the mixing hypothesis, many things can go wrong; for instance, we might obtain a map g defined on a space that is even "longer" than the real line. We return to this idea in Section 2.7. Here is the statement of our main result.

Theorem 2.2 *Let $f \in \mathcal{CPMM}$ with transition matrix T, and fix $\lambda > 1$. Assume f is topologically mixing. Then f is conjugate via a homeomorphism ψ to some map $g \in \mathcal{CPMM}$ of constant slope λ if and only if*

$$T \text{ has a nonnegative eigenvector } v = (v_I) \in \mathbb{R}^{\mathcal{B}(P)} \text{ with eigenvalue } \lambda. \qquad (2.2.1)$$

Since the map f defines a topological dynamical system without regard to geometry, there is no loss of generality if we assume that $f \in \mathcal{CPMM}_{[0,1]}$. We will make this assumption from now on.

The proof proceeds in several pieces. First we show the easy implication, that Condition 2.2.1 is necessary. Showing the sufficiency of Condition 2.2.1 requires much more work. We give an explicit construction of the conjugating map ψ in several stages. Our construction closely follows the lines of the proofs of [6, Theorem 2.5] and [1, Theorem 4.6.8], but the unsummability of v introduces some additional difficulties not present in these previous works. The topological mixing hypothesis was introduced to overcome these difficulties.

Lemma 2.3 (Bobok) *Condition 2.2.1 is necessary.*

Proof This proof is due to private communication with Jozef Bobok [7]. As mentioned before, we may suppose that $f \in \mathcal{CPMM}_{[0,1]}$. Let ψ be the conjugating map, $\psi \circ f = g \circ \psi$. Define v by $v_I = |\psi(I)|$, $I \in \mathcal{B}(P)$, where $|\cdot|$ denotes the length of an interval. A priori, we may have $|\psi(I)| = \infty$; this happens if and only if I contains one of the endpoints $0, 1$ and ψ maps this endpoint to one of $\pm\infty$. (Recall that if $0, 1$ are accumulation points of P, then they are not endpoints of any P-basic interval). We want to show that all the entries of v are finite. Since g is monotone with slope of absolute value λ on each $\psi(P)$-basic interval, we have

$$|g(\psi(I))| = \lambda|\psi(I)|, \qquad I \in \mathcal{B}(P), \tag{2.2.2}$$

where if one side of the equality is infinite then so is the other. Let $\mathcal{F}$ denote the collection of all P-basic intervals I such that $|\psi(I)| = \infty$. If $I \in \mathcal{F}$ and if $f(J) \supseteq I$, then by the conjugacy of f, g and by Equation 2.2.2, it follows that $J \in \mathcal{F}$. Now invoke the topological mixing property, and it follows that either $\mathcal{F} = \emptyset$ or $\mathcal{F} = \mathcal{B}(P)$.

Suppose toward contradiction that $\mathcal{F} = \mathcal{B}(P)$. Then there are at most two P-basic intervals. There cannot be only one P-basic interval, because a monotone map is not mixing. Therefore there are exactly two P-basic intervals and $\psi([0,1]) = [-\infty, \infty]$. Since g has constant slope, $g(x)$ is finite whenever x is finite. By the mixing hypothesis, g is surjective, and therefore g either fixes or interchanges ∞, $-\infty$. In either case it follows that g is monotone, contradicting the mixing hypothesis. We may conclude that $\mathcal{F} = \emptyset$ and all entries of v are finite.

We still need to show that v is an eigenvector for T. Applying Equation 2.2.2 we have

$$\lambda v_I = \lambda|\psi(I)| = |g(\psi(I))| = |\psi(f(I))| = \sum_{J \subseteq f(I)} |\psi(J)| = \sum_{J \in \mathcal{B}(P)} T_{IJ} v_J.$$

■

Remark 2.4 *The proof of Lemma 2.3 does not use the full strength of the mixing hypothesis. The lemma continues to hold if we relax the hypotheses of Theorem 2.2 and assume only that f is topologically transitive.*

Now we begin the long work of proving the sufficiency of Condition 2.2.1. Let f, T, be as in the statement of the theorem, fix $\lambda > 0$, and suppose $Tv = \lambda v$ for some nonzero vector $v = (v_I) \in \mathbb{R}^{\mathcal{B}(P)}$ with nonnegative entries. We still assume $f \in \mathcal{CPMM}_{[0,1]}$. We will construct a map $\psi : [0,1] \to [-\infty, \infty]$ which is a homeomorphism onto its image in such a way that $g := \psi \circ f \circ \psi^{-1}$ has constant slope λ. Define the sets

$$P_n = \cup_{i=0}^{n} f^{-i}(P),\ n \in \mathbb{N}, \quad Q = \cup_{i=0}^{\infty} f^{-i}(P)$$

The set Q is backward invariant by construction and forward invariant because P is forward invariant. Q is a dense subset of $[0,1]$ because f is mixing. Choose a basepoint $p_0 \in P$ and define ψ on Q by the formula

$$\psi(x) = \begin{cases} 0, & \text{if } x = p_0 \\[2mm] \lambda^{-n} \displaystyle\sum_{\substack{J \in B(P^n) \\ p_0 < J < x}} v_{f^n(J)}, & \text{if } x \in P_n, x > p_0 \\[4mm] -\lambda^{-n} \displaystyle\sum_{\substack{J \in B(P^n) \\ x < J < p_0}} v_{f^n(J)}, & \text{if } x \in P_n, x < p_0 \end{cases} \tag{2.2.3}$$

The choice of p_0 is somewhat arbitrary, but to simplify the proof of Lemma 2.5 (v), we insist that $0 < p_0 < 1$ and that p_0 is an endpoint of some P-basic interval (i.e., p_0 is not a 2-sided accumulation point of P). This is possible because P is a closed, countable subset of $[0,1]$ and hence cannot be perfect.

Lemma 2.5 *The function $\psi : Q \to [-\infty, \infty]$ has the following properties:*

(i) ψ is well-defined; i.e. when $x \in P_{n_1}$ and $x \in P_{n_2}$, the sums agree.

(ii) ψ is strictly monotone increasing.

(iii) If $x, x' \in Q$ belong to an interval of monotonicity of f, then

$$|\psi(f(x)) - \psi(f(x'))| = \lambda|\psi(x) - \psi(x')|,$$

where if one side of the equality is infinite, then so is the other.

(iv) For arbitrary $x, x' \in Q$:

$$|\psi(f(x)) - \psi(f(x'))| \leq \lambda |\psi(x) - \psi(x')|,$$

and we allow for the possibility that one or both sides of this inequality are infinite.

(v) For $0 < x < 1$, $\psi(x)$ is finite.

Proof Our proof of (i) is borrowed from [6].

(i) Suppose $K \in \mathcal{B}(P_n)$. Then $f^n|K$ is monotone and $f^n(K) \in \mathcal{B}(P)$. Therefore

$$\lambda^{-n-1} \sum_{\substack{J \in \mathcal{B}(P_{n+1}) \\ J \subseteq K}} v_{f^{n+1}(J)} = \lambda^{-n-1} \sum_{\substack{J \in \mathcal{B}(P_1) \\ J \subseteq f^n(k)}} v_{f(J)} =$$

$$= \lambda^{-n-1} \sum_{J \in \mathcal{B}(P_0)} T_{f^n(K)J} v_J = \lambda^{-n-1} \lambda v_{f^n(K)} = \lambda^{-n} v_{f^n(K)}$$

This shows that ψ is well-defined.

(ii) We will use the nonnegativity of the eigenvector v together with the mixing hypothesis to show that the entries of v must be strictly positive. Strict monotonicity of ψ then follows from the definition. Since v is not the zero vector, there must be some P-basic interval I_0 with $v_{I_0} \neq 0$. Let $I \in \mathcal{B}(P)$. By the mixing hypothesis, there is $n \in \mathbb{N}$ such that $(T^n)_{II_0} \neq 0$. Then $v_I = \lambda^{-n} \sum_J (T^n)_{IJ} v_J \geq \lambda^{-n} v_{I_0} > 0$.

(iii) Consider the case when $(x, x') = K \in \mathcal{B}(P_n)$. Then $|\psi(x) - \psi(x')| = \lambda^{-n} v_{f^n(K)}$ by the definition of ψ. Moreover, $f(K) \in \mathcal{B}(P_{n-1})$, so that $|\psi(f(x)) - \psi(f(x'))| = \lambda^{-(n-1)} v_{f^{n-1}(f(K))}$, so the claim holds in this case. By taking sums and limits, the claim holds for every $x, x' \in Q$ contained in a single interval of monotonicity of f.

(iv) This is the inequality that survives from (iii) when we allow for folding between x and x'.

(v) Let x be given, $0 < x < 1$. Assume $x < p_0$; the proof when $x > p_0$ is similar. Fix a P-basic interval J_0 with p_0 at one endpoint. Because f is mixing, there exists n such that $J_0 \cap f^{-n}((p_0, 1)) \neq \emptyset$ and $J_0 \cap f^{-n}((0, x)) \neq \emptyset$. By the intermediate value theorem there exist $x_1, x_2 \in J_0$ with $f^n(x_1) = x$ and $f^n(x_2) = p_0$. By (iv) applied n times, $|\psi(x)| \leq \lambda^n |\psi(x_2) - \psi(x_1)|$. But by (ii), $|\psi(x_2) - \psi(x_1)| \leq |\psi(\sup J_0) - \psi(\inf J_0)|$. At the two endpoints of J_0, ψ takes the finite values 0 and v_{J_0}.

∎

The main problem to tackle before we can extend ψ to the desired homeomorphism is to show that the map we have defined so far has no jump discontinuities.

Problem 2.6 *Show that for each $x \in [0, 1]$,*

$$\inf \psi(Q \cap (x, 1]) = \sup \psi(Q \cap [0, x)),$$

except that for $x = 0$ we write $\psi(0)$ in place of the supremum and for $x = 1$ we write $\psi(1)$ in place of the infimum.

The resolution of this problem makes essential use of the topological mixing hypothesis as well as the order structure of the interval $[0, 1]$. Moreover, special treatment is required for the points $x \in Q$ – we must show the continuity of ψ from each side separately. We do this by introducing a notion of "half-points."

2.3   Half-Points

Construct the sets

$$\tilde{Q} = (Q \times \{+, -\}) \smallsetminus \{(0, -), (1, +)\}, \quad S = ([0, 1] \smallsetminus Q) \cup \tilde{Q}$$

The way to think of this definition is that we are splitting each point $x \in Q$ into the two *half-points* $(x, +)$ and $(x, -)$. S is the interval $[0, 1]$ with each point of Q

replaced by half-points. We use boldface notation to represent points in S, whether half or whole. Thus, $\boldsymbol{x}$ may mean x or $(x, +)$ or $(x, -)$, depending on the context.

Let us extend the dynamics of f from $[0, 1]$ to S. Recall that Q is both forward and backward invariant. On $S \setminus \tilde{Q} = [0, 1] \setminus Q$ we keep the map f without change. To extend f from Q to $\tilde{Q}$ we define a notion of the orientation of the map at half-points. We say that f is *orientation-preserving* (resp. *orientation-reversing*) at the half-point $(x, +)$ if some half-neighborhood $[x, x + \varepsilon)$ is contained in some $J \in \mathcal{B}(P)$ with $f|_J$ increasing (resp. decreasing). For a half-point $(x, -)$, the definition is the same, except that we look at a half-neighborhood of the form $(x - \varepsilon, x]$. It is not clear how to decide if f is orientation-preserving or orientation-reversing at the accumulation points of P. It may happen that every half-neighborhood of x contains $f(x)$ in the interior of its image, so that neither definition is appropriate. Nevertheless, we define the extended map f on $\tilde{Q}$ by the following formula:

$$
\begin{aligned}
f(x, +) &= \begin{cases}
(f(x), +), & \text{if } f \text{ is orientation-preserving at } (x, +) \\
(f(x), -), & \text{if } f \text{ is orientation-reversing at } (x, +) \\
(f(x), +), & \text{if } \forall_{\varepsilon > 0} \ \exists_{x' \in P \cap [x, x+\varepsilon)} \ f(x') > f(x) \\
(f(x), -), & \text{otherwise}
\end{cases} \\[2em]
f(x, -) &= \begin{cases}
(f(x), +), & \text{if } f \text{ is orientation-reversing at } (x, -) \\
(f(x), -), & \text{if } f \text{ is orientation-preserving at } (x, -) \\
(f(x), +), & \text{if } \forall_{\varepsilon > 0} \ \exists_{x' \in P \cap (x-\varepsilon, x]} \ f(x') > f(x) \\
(f(x), -), & \text{otherwise}
\end{cases}
\end{aligned}
\tag{2.3.1}
$$

Let us say a few words about the "otherwise" cases. Consider a half-point $(x, +)$ which does not fit into any of the first three cases. We claim that for such a point, $\forall_{\varepsilon > 0} \ \exists_{x' \in P \cap [x, x+\varepsilon)} \ f(x') < f(x)$. If not, we would have to conclude that $\exists_{\varepsilon > 0} \ \forall_{x' \in P \cap [x, x+\varepsilon)} \ f(x') = f(x)$. But this is impossible, because the half-neighborhood $[x, x + \varepsilon)$ must contain some P-basic interval J, and by the strict monotonicity of $f|_J$ the two endpoints of this interval have distinct images. Similarly, if a half-point

$(x, -)$ falls into the "otherwise" case, then $\forall_{\varepsilon>0} \exists_{x' \in P \cap (x-\varepsilon, x]} \ f(x') < f(x)$. This is relevant in the proofs of Lemmas 2.7 and 2.8.

Now we define a real-valued function Δ_ψ on S by the formula

$$\Delta_\psi(\boldsymbol{x}) = \begin{cases} \inf \psi(Q \cap (x, 1]) - \psi(x) & \text{if } \boldsymbol{x} = (x, +) \in \tilde{Q} \\ \psi(x) - \sup \psi(Q \cap [0, x)) & \text{if } \boldsymbol{x} = (x, -) \in \tilde{Q} \\ \inf \psi(Q \cap (x, 1]) - \sup \psi(Q \cap [0, x)) & \text{if } \boldsymbol{x} = x \in S \setminus \tilde{Q} \end{cases}$$

If $\Delta_\psi(\boldsymbol{x}) > 0$, then we say that $\boldsymbol{x}$ is an *atom* for ψ and $\Delta_\psi(\boldsymbol{x})$ is its *mass*. In this language, Problem 2.6 asks us to show that ψ has no atoms.

2.4   Lemmas About Half-Points

The next lemma is an analog of Lemma 2.5 (iii) for a single point (or half-point) $\boldsymbol{x}$. We introduced half-points for the purpose of proving this lemma even at the folding points of f.

Lemma 2.7 *Let $\boldsymbol{x} \in S$. Then $\Delta_\psi(f(\boldsymbol{x})) = \lambda \Delta_\psi(\boldsymbol{x})$.*

Proof   Consider first the case when $\boldsymbol{x} = x$ is a whole-point, i.e. $\boldsymbol{x} \in S \setminus \tilde{Q}$. Then x belongs to the interior of some P-basic interval J. We may choose a sequence y_i in $Q \cap J$ converging to x from the left-hand side, and a sequence z_i in $Q \cap J$ converging to x from the right-hand side. Then $f(y_i)$ and $f(z_i)$ are sequences in Q converging to $f(x)$ from opposite sides. By the monotonicity of ψ and the definition of Δ_ψ we have $|\psi(z_i) - \psi(y_i)| \to \Delta_\psi(\boldsymbol{x})$ and $|\psi(f(z_i)) - \psi(f(y_i))| \to \Delta_\psi(f(\boldsymbol{x}))$. Since J is an interval of monotonicity of f, the result follows from Lemma 2.5 (iii).

Now consider the case when $\boldsymbol{x} = (x, +)$ or $\boldsymbol{x} = (x, -)$, and suppose an appropriate half-neighborhood of $\boldsymbol{x}$ is contained in a single P-basic interval J so that f is either orientation-preserving or orientation-reversing at $\boldsymbol{x}$. We may repeat the proof from the previous case, with one modification. If $\boldsymbol{x} = (x, +)$, then we take y_i to be instead the constant sequence with each member equal to x. If $\boldsymbol{x} = (x, -)$, then we take z_i

to be instead the constant sequence with each member equal to x. Then the rest of the proof holds as written.

Now consider the case when $\boldsymbol{x} = (x, +)$ and $f(\boldsymbol{x}) = (f(x), +)$, but every half-neighborhood $[x, x+\varepsilon)$ meets P. We will show in this case that $\Delta_\psi(\boldsymbol{x})$ and $\Delta_\psi(f(\boldsymbol{x}))$ are both zero. Choose points $z_i \in P$ which converge monotonically to x from the right and such that each $f(z_i) > f(x)$. By continuity, $f(z_i) \to f(x)$, and after passing to a subsequence, we may assume that this convergence is also monotone. Now we calculate $\Delta_\psi(\boldsymbol{x})$ using the sequence z_i and appealing back to the definition of ψ.

$$\Delta_\psi(\boldsymbol{x}) = \lim_{i\to\infty}\left(\psi(z_i) - \psi(x)\right) = \lim_{i\to\infty}\sum_{\substack{J\in B(P)\\ x<J<z_i}} v_J = \lim_{i\to\infty}\sum_{j=i}^{\infty}\sum_{\substack{J\in B(P)\\ z_{j+1}<J<z_j}} v_J = 0$$

The rearrangement of the sum is justified because for each P-basic interval J between x and z_i there is exactly one $j \geq i$ such that J lies between z_{j+1} and z_j. But by Lemma 2.5 (v), when $i = 1$ we have already a convergent series. Thus, when we sum smaller and smaller tails of the series, we obtain 0 in the limit. We may apply exactly the same argument to compute $\Delta_\psi(f(\boldsymbol{x}))$ along the sequence $f(z_i)$, because these points also belong to the invariant set P and decrease monotonically to $f(x)$.

There are three other cases in which every appropriate half-neighborhood of $\boldsymbol{x}$ meets P; again in each of these cases $\Delta_\psi(\boldsymbol{x}) = 0$ and $\Delta_\psi(f(\boldsymbol{x})) = 0$ by similar arguments. ∎

The next lemma extends the intermediate value theorem to S.

Lemma 2.8 *Let $x_1 < x_2$ be any two points in $[0, 1]$, not necessarily in Q, and let $k \in \mathbb{N}$. Suppose that there exists a point $\boldsymbol{y} \in S$ with y strictly between $f^k(x_1)$ and $f^k(x_2)$. Then there exists $\boldsymbol{x} \in S$ with x between x_1 and x_2 such that $f^k(\boldsymbol{x}) = (\boldsymbol{y})$.*

Proof If $\boldsymbol{y} = y \in S \setminus \tilde{Q}$, we just apply the invariance of Q and the usual intermediate value theorem. If $\boldsymbol{y} \in \tilde{Q}$, then we consider the set $A = [x_1, x_2] \cap f^{-k}(y)$. It is nonempty by the usual intermediate value theorem, compact by the continuity of f^k, and contained in Q by the invariance of Q. Suppose first that $f^k(x_1) <$

$f^k(x_2)$. If x' satisfies $x_1 < x' < \min A$, then $f^k(x') < y$ by the usual interme-diate value theorem and the minimality of $\min A$. It follows that $f^k(\min A, -) = (y, -)$. Similarly, $f^k(\max A, +) = (y, +)$. Thus $\boldsymbol{x}$ may be taken as one of the points $(\min A, -), (\max A, +)$. The proof when $f^k(x_1) > f^k(x_2)$ is similar, except that $f^k(\min A, -) = (y, +)$ and $f^k(\max A, +) = (y, -)$. ∎

2.5   No Atoms

Now we are ready to answer Problem 2.6.

Lemma 2.9 ψ *has no atoms; that is, Δ_ψ is identically zero.*

Proof   Assume toward contradiction that there is a point $\boldsymbol{b} \in S$ such that $\Delta_\psi(\boldsymbol{b}) > 0$. For $n = 0, 1, 2, \ldots$, let $\boldsymbol{b}_n := f^n(\boldsymbol{b}) \in S$ and denote the corresponding point in $[0, 1]$ by b_n. We denote the orbit of b by $Orb(b) = \{b_0, b_1, b_2, \ldots\}$. By Lemma 2.7,

$$\Delta_\psi(\boldsymbol{b}_n) = \lambda^n \Delta_\psi(\boldsymbol{b}), \quad n \in \mathbb{N} \tag{2.5.1}$$

and this grows to ∞ because $\lambda > 1$. If $Orb(b)$ has an accumulation point in the open interval $(0, 1)$, then the increment of ψ across a small neighborhood of this accumulation point is ∞, contradicting Lemma 2.5 (v) and we are done. Henceforth, we may assume that the orbit of b only accumulates at (one or both) endpoints of $[0, 1]$. Consider first the case when $Orb(b)$ accumulates at only one endpoint of $[0, 1]$, and assume without loss of generality that $\lim_{n \to \infty} b_n = 1$.

Since f is mixing, it must have a fixed point w with $0 < w < 1$. Since $b_n \to 1$, it follows that $b_n > w$ for all sufficiently large n. Thus, after replacing $\boldsymbol{b}$ and b with their appropriate images, we may assume that $b_n > w$ for all $n \in \mathbb{N}$. Equation 2.5.1 continues to hold, and it follows that b is not a fixed point for f, so $b \neq 1$.

Now consider the following claim:

For all $N \in \mathbb{N}$ there exist $n > N$ and $\boldsymbol{a} \in S$

$$\text{such that } a \notin Orb(b) \text{ and } f(\boldsymbol{a}) = \boldsymbol{b}_n \text{ and } w < a < b_{n+1}. \quad (\star)$$

The proof of Claim $(\star)$ proceeds in two cases. First, assume that $b_N < b_{N+1} < b_{N+2} < \ldots$; i.e., starting from time N, the orbit of b moves monotonically to the right. Since f is mixing, the interval $[b_{N+1}, 1]$ cannot be invariant, so there must exist $c > b_{N+1}$ with $f(c) < b_{N+1}$. Take $n = \max\{i : b_i < c\}$. Clearly $n > N$. The relevant ordering of points is $b_{n-1} < b_n < c < b_{n+1}$. Since $f(b_n) > b_n$ and $f(c) < b_n$, it follows by Lemma (2.8) that there exists $\boldsymbol{a}$ with a between b_n and c such that $f(\boldsymbol{a}) = \boldsymbol{b}_n$. Clearly, $a \neq b_{n-1}$. It follows that $a \notin Orb(b)$. Moreover, $w < a < b_{n+1}$.

The remaining case is that there exists $i \geq N$ such that $b_{i+1} < b_i$; i.e., at some time later than N, the orbit moves to the left. But our orbit is converging to the right-hand endpoint of $[0, 1]$, so it cannot go on moving to the left forever. Let $n = \min\{j > i : b_{j+1} > b_j\}$. We have $n > N$, and the relevant ordering of points is $b_{n-1} > b_n$ and $b_{n+1} > b_n$. Since $f(w) < b_n$ and $f(b_n) > b_n$, it follows by Lemma (2.8) that there exists $\boldsymbol{a}$ with a between w and b_n such that $f(\boldsymbol{a}) = \boldsymbol{b}_n$. Again, we see that $a \neq b_{n-1}$, so $a \notin Orb(b)$. Finally, $a < b_{n+1}$. This concludes the proof of Claim $(\star)$.

Now we apply Claim $(\star)$ recursively to find infinitely many distinct atoms between w and b, each with the same positive mass. At stage 1, find n_1 and $\boldsymbol{a}_1$ with $a_1 \notin Orb(b)$ such that $f(\boldsymbol{a}_1) = \boldsymbol{b}_{n_1}$ and $w < a_1 < b_{n_1+1}$. Now we apply Lemma (2.8) to f^{n_1+1} to find $\boldsymbol{x}_1$ with x_1 between w and b such that $f^{n_1+1}(\boldsymbol{x}_1) = \boldsymbol{a}_1$. Then $f^{n_1+2}(\boldsymbol{x}_1) = \boldsymbol{b}_{n_1}$, so by applying Lemma (2.7) and Equation (2.5.1) we have $\Delta_\psi(\boldsymbol{x}_1) = \lambda^{-(n_1+2)}\Delta_\psi(\boldsymbol{b}_{n_1}) = \lambda^{-2}\Delta_\psi(\boldsymbol{b})$. The point $\boldsymbol{x}_1$ will serve as the first of infinitely many points between w and b at which ψ has this particular increment. At stage i, set $N = n_{i-1}$ and apply Claim $(\star)$ to find n_i and $\boldsymbol{a}_i$ with $n_i > n_{i-1}$. Again, we can find $\boldsymbol{x}_i$ with x_i between w and b and $f^{n_i+1}(\boldsymbol{x}_i) = \boldsymbol{a}_i$, whence $\Delta_\psi(\boldsymbol{x}_i) = \lambda^{-2}\Delta_\psi(\boldsymbol{b})$ as before. It remains to check that the points $\{x_i\}$ are distinct. Observe that $f^{n_i+1}(x_i) = a_i$ does not belong to the invariant set $Orb(b)$, whereas $f^{n_i+2}(x_i) = b_{n_i} \in Orb(b)$. By construction, the numbers $\{n_i\}$ are all distinct. Thus, the points $\{x_i\}$ are distinguished from one another by the time required to make first entrance into $Orb(b)$.

Now we use our atoms to produce a contradiction. By Lemma 2.5 (v), the increment $\psi(b) - \psi(w)$ is finite. Choose an integer n large enough that $n\lambda^{-2}\Delta_\psi(\boldsymbol{b}) >$

$\psi(b) - \psi(w)$. Consider the points $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$, and let δ be the minimum distance between two adjacent points of the set $\{w, b\} \cup \{x_1, x_2, \ldots, x_n\}$. For each $i = 1, \ldots, n$ there exist $y_i, z_i \in Q$ with $y_i < x_i < z_i$ and $\max\{z_i - x_i, x_i - y_i\} < \delta/2$. Then $\psi(z_i) - \psi(y_i) \geq \lambda^{-2} \Delta_\psi(\boldsymbol{b})$. By the monotonicity of ψ,

$$\psi(b) - \psi(w) \geq \sum_{i=1}^{n} \psi(z_i) - \psi(y_i) > n\lambda^{-2}\Delta_\psi(\boldsymbol{b}) > \psi(b) - \psi(w).$$

This is a contradiction; in words, we cannot have infinitely many atoms between w and b all having the same positive mass when the total increment of ψ between w and b is finite. This completes the proof in the case that $Orb(b)$ accumulates at only one endpoint of $[0, 1]$.

Finally, let us say a few words about the case when $Orb(b)$ accumulates at both endpoints of $[0, 1]$. In this case, $f(0) = 1$ and $f(1) = 0$ by continuity. Again by continuity, for sufficiently large n the points b_n belong alternately to a small neighborhood of 0 and a small neighborhood of 1. Thus, the subsequence b_{2n} accumulates only on a single endpoint of $[0, 1]$. The map f^2 is again topologically mixing. It is straightforward, then, to modify the above proof to deal with this case, by working along the subsequence b_{2n} and writing f^2 and λ^2 in place of f and λ.

∎

2.6 Sufficiency of the Eigenvector Criterion

Having resolved Problem 2.6, we are ready to finish the proof of Theorem 2.2.

Proof It remains to show that Condition 2.2.1 is sufficient. We have defined on the dense subset $Q \subset [0, 1]$ a strictly monotone map $\psi : Q \to [-\infty, \infty]$. In light of Lemma 2.9, the formula $\psi(x) = \sup \psi(Q \cap [0, x)) = \inf \psi(Q \cap (x, 1])$ gives a well-defined extension $\psi : [0, 1] \to [-\infty, \infty]$. Strict monotonicity of the extension follows from the strict monotonicity of $\psi|_Q$ and the density of Q. We claim that the extended function ψ is continuous. It suffices to verify for each x that $\psi(x) = \lim_{y \to x^-} \psi(y) = \lim_{z \to x^+} \psi(z)$. By monotonicity of ψ and the density of Q we may evaluate these

one-sided limits using points $y, z \in Q$, and by our definition of the extended map ψ the claim follows. Finally, from strict monotonicity and continuity, it follows that $\psi : [0,1] \to [-\infty, \infty]$ is a homeomorphism onto its image.

Define a map $g : \psi([0,1]) \to \psi([0,1])$ by the composition $g := \psi \circ f \circ \psi^{-1}$. It is countably piecewise monotone and Markov with respect to $\psi(P)$. If $y = \psi(x)$ and $y' = \psi(x')$ belong to a single $\psi(P)$-basic interval, then x and x' belong to an interval of monotonicity of f. By Lemma 2.5 (iii) and the density of Q we may conclude that $|g(y) - g(y')| = \lambda |y - y'|$. This shows that g has constant slope λ.

∎

2.7   The Mixing Hypothesis

Now we show that the mixing hypothesis in Theorem 2.2 is essential. We give an example of a map f in $\mathcal{CPMM}$ which is topologically transitive but not mixing. We give a nonnegative eigenvector v for the transition matrix T, but prove that f is not conjugate to any map on any subinterval $[a, b] \subseteq [-\infty, \infty]$ with constant slope the eigenvalue of v.

Fix $\lambda = 2 + \sqrt{5}$ and take the corresponding maps F_λ, h, and g_λ defined in Section 1.5. By way of reminder, F_λ is the piecewise affine "connect-the-dots" map with "dots" at $(k, k-1)$, $(k+b, k+b+1)$, $k \in \mathbb{Z}$, where $b = (\sqrt{5} - 1)/2$; it is piecewise monotone and Markov with respect to the set $\{k, k+b : k \in \mathbb{Z}\}$. For concreteness, we take $h(x) = \ln(x/(1-x))$. The map $g_\lambda = h^{-1} \circ F_\lambda \circ h$ with additional fixed points at 0, 1, is piecewise monotone and Markov with respect to the set $\{0, 1\} \cup \{h^{-1}(k), h^{-1}(k+b) : k \in \mathbb{Z}\}$; it is also transitive, as explained in Section 1.5. Figure 2.1 shows the graph of F_λ together with its Markov partition.

Now define a map $f : [-1, 1] \to [-1, 1]$ by the formula

$$
f(x) = \begin{cases} -g_\lambda(x), & \text{if } x \in [0,1] \\ -x & \text{if } x \in [-1,0] \end{cases}
$$

Figure 2.1. The Constant Slope Map $F_\lambda : \mathbb{R} \to \mathbb{R}$

This map f is piecewise monotone and Markov with respect to the set $P = \{0, \pm 1\} \cup \{\pm h^{-1}(k), \pm h^{-1}(k+b) : k \in \mathbb{Z}\}$. We enumerate the P-basic intervals as follows:

$$I_{2k} = [h^{-1}(k), h^{-1}(k+b)], \quad I_{2k+1} = [h^{-1}(k+b), h^{-1}(k+1)], \quad J_k = -I_k, \quad k \in \mathbb{Z}.$$

The Markov transitions are given by

$$f(I_{2k}) = \bigcup_{i=2k-2}^{2k+2} J_i, \quad f(I_{2k+1}) = \bigcup_{i=2k}^{2k+2} J_i, \quad f(J_k) = I_k, \quad k \in \mathbb{Z}. \tag{2.7.1}$$

Figure 2.2 shows the graph of f (in bold) as well as the corresponding Markov partition. Superimposed is the graph of the second iterate f^2. By construction, $f^2|_{[0,1]}$ and $f^2|_{[-1,0]}$ are both isomorphic copies of the map g_λ. In this sense, the map f is the dynamical square root of g_λ.

We claim that f is topologically transitive, but not topologically mixing. To see the transitivity, let U, V be arbitrary nonempty open subsets of $[-1, 1]$. After shrinking these sets, we may assume that $0 \notin U, V$. Consider first the case when $U, V \subset [0, 1]$. By the transitivity of g there exists n such that $U \cap g^{-n}(V) \neq \emptyset$, but then $U \cap f^{-2n}(V) \neq \emptyset$. The case when $U, V \subset [-1, 0]$ is similar. Now consider the case when $U \subset [0, 1]$ and $V \subset [-1, 0]$. Using the reflected set $-V$ and the transitivity of g, find n such that $U \cap g^{-n}(-V) \neq \emptyset$. Then $U \cap f^{2n-1}(V) \neq \emptyset$. The case when $U \subset [-1, 0]$ and $V \subset [0, 1]$ is similar. This shows topological transitivity of f. To see
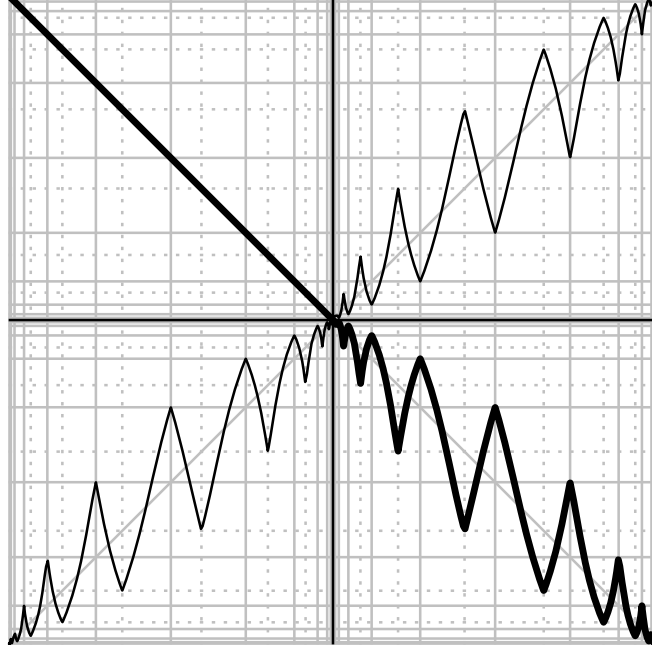
Figure 2.2. A Map f Demonstrating the Sharpness of Theorem 2.2

that f is not topologically mixing, notice that the set $\{n \in \mathbb{N} : (0,1) \cap f^{-n}(0,1) \neq \emptyset\}$ consists of only the even natural numbers.

Let T be the binary transition matrix for the map f. Let us find all nonnegative solutions $v \in \mathbb{R}^{\mathcal{B}(P)}$ to the equation $Tv = \sqrt{\lambda}v$. Comparing Equation 2.7.1 with the definition of T, we are looking for all nonnegative solutions to the infinite system of equations

$$\begin{cases} \sqrt{\lambda}\, v_{I_{2k}} &= \sum_{i=2k-2}^{2k+2} v_{J_i} \\[2mm] \sqrt{\lambda}\, v_{I_{2k+1}} &= \sum_{i=2k}^{2k+2} v_{J_i} \qquad\qquad k \in \mathbb{Z} \qquad\qquad (2.7.2) \\[2mm] \sqrt{\lambda}\, v_{J_k} &= v_{I_k} \end{cases}$$

By direct verification (remembering that we fixed $\lambda = 2 + \sqrt{5}$), Equation 2.7.2 is satisfied by

$$v_{I_{2k}} = 2, \quad v_{I_{2k+1}} = \sqrt{5} - 1, \quad v_{J_{2k}} = \frac{2}{\sqrt{\lambda}}, \quad v_{J_{2k+1}} = \frac{\sqrt{5} - 1}{\sqrt{\lambda}}, \qquad k \in \mathbb{Z}. \qquad (2.7.3)$$

We claim that, up to scalar multiples, Equation 2.7.3 defines the unique non-negative solution $v \in \mathbb{R}^{\mathcal{B}(P)}$ to Equation 2.7.2. This may be seen as follows. First, substitute the last line in Equation 2.7.2 into the first two lines to obtain

$$\begin{cases} \lambda\, v_{I_{2k}} & = \ \sum_{i=2k-2}^{2k+2} v_{I_i} \\[2mm] \lambda\, v_{I_{2k+1}} & = \ \sum_{i=2k}^{2k+2} v_{I_i} \end{cases} \qquad k \in \mathbb{Z}.$$

Adding and subtracting equations, we obtain

$$\begin{cases} \lambda\big(v_{I_{2k+1}} + v_{I_{2k-1}} - v_{I_{2k}}\big) = v_{I_{2k}} \\[2mm] \lambda v_{I_{2k+1}} = v_{I_{2k}} + v_{I_{2k+1}} + v_{I_{2k+2}} \end{cases} \qquad k \in \mathbb{Z}.$$

Solving for later variables in terms of earlier ones, we obtain

$$\begin{bmatrix} v_{I_{2k+1}} \\[2mm] v_{I_{2k+2}} \end{bmatrix} = \begin{bmatrix} -1 & 1 + \frac{1}{\lambda} \\[2mm] -\lambda + 1 & \lambda - 1 - \frac{1}{\lambda} \end{bmatrix} \begin{bmatrix} v_{I_{2k-1}} \\[2mm] v_{I_{2k}} \end{bmatrix} \qquad k \in \mathbb{Z}. \qquad (2.7.4)$$

Equation 2.7.4 should be regarded as a linear recurrence relation on v. Substituting our fixed value $\lambda = 2 + \sqrt{5}$ and observing that the matrix is invertible, we may conclude inductively that

$$\begin{bmatrix} v_{I_{2k+1}} \\[2mm] v_{I_{2k+2}} \end{bmatrix} = \begin{bmatrix} -1 & -1 + \sqrt{5} \\[2mm] -1 - \sqrt{5} & 3 \end{bmatrix}^k \begin{bmatrix} v_{I_1} \\[2mm] v_{I_2} \end{bmatrix}, \qquad k \in \mathbb{Z}.$$

The action of this matrix and its iterates on $\mathbb{R}^2$ may be regarded as a dynamical system, and the entries of v are the orbit of the initial point (v_{I_1}, v_{I_2}). To obtain nonnegative entries for v, we must choose the initial point so that the whole orbit remains in the first quadrant. The point $(v_{I_1}, v_{I_2}) = (\sqrt{5} - 1, 2)$ is a fixed point of this dynamical system (an eigenvector with eigenvalue 1), and so is every scalar multiple thereof. There are no other eigenvectors, and it follows that our matrix acts as a shear on $\mathbb{R}^2$ parallel to this line of fixed points. Thus, the only way to obtain a

whole orbit in the first quadrant is to choose the initial point from the line of fixed points, recovering (up to a scalar multiple) the vector given in Equation 2.7.3. This completes the proof that Equation 2.7.3 gives (up to a scalar multiple) the unique nonnegative vector $v \in \mathbb{R}^{\mathcal{B}(P)}$ satisfying $Tv = \sqrt{\lambda}v$.

Now we show that despite the existence of this eigenvector v, there does not exist any conjugacy ψ of the map f to a map g of constant slope $\sqrt{\lambda}$. Assume the contrary. Then by the uniqueness of v and by Remark 2.4, we have

$$|\psi(I_{2k})| = 2c, \quad |\psi(I_{2k+1})| = (\sqrt{5}-1)c, \quad |\psi(J_{2k})| = \frac{2c}{\sqrt{\lambda}}, \quad |\psi(J_{2k+1})| = \frac{(\sqrt{5}-1)c}{\sqrt{\lambda}}, \qquad k \in \mathbb{Z},$$

for some positive real scalar c. But the P-basic intervals accumulate at the center of $[-1, 1]$ so that a small open interval $(-\varepsilon, \varepsilon)$ contains infinitely many P-basic intervals. Thus, $\psi(-\varepsilon, \varepsilon)$ has infinite length. On the other hand, a nondecreasing homeomorphism $\psi : [-1, 1] \to [-\infty, \infty]$ must take finite values at every interior point of the interval $[-1, 1]$. This is a contradiction.

3. COUNTABLE EXTENSIONS IN DIMENSION ZERO

In one-dimensional dynamics there is a well-developed theory of degree one liftings, that is, continuous transformations $F : \mathbb{R} \to \mathbb{R}$ on the real line which factor through the topological covering map $\pi : \mathbb{R} \to \mathbb{R}/\mathbb{Z}$ to a transformation $f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ on the circle. The degree one property means that $F(X + 1) = F(X) + 1$, i.e., that the map F commutes with the deck transformations associated with the covering map. Consequently, there is a well-defined displacement function $\mathbb{R}/\mathbb{Z} \to \mathbb{R}$ given by $x \mapsto F(X) - X$ for $X \in \pi^{-1}(\{x\})$, which measures in some sense how far around the circle the transformation f carries each point. One-dimensional rotation theory consists in large part of studying ergodic averages of this displacement function and the implications for the dynamics of the maps f and F.

In this chapter, we study a zero-dimensional analog of degree one liftings which we call countable extensions. They are countable state subshifts of finite type (topological Markov chains). They factor through a countable-to-one topological covering map onto a finite-state chain, and the group of shift-commuting deck transformations is isomorphic to $\mathbb{Z}$. Countable extensions also come with a displacement function, analogous to the one-dimensional case. Treating this displacement function as a potential and applying the theory of thermodynamic formalism we obtain results regarding entropy and maximal measures.

3.1 Definitions and Basic Properties.

Start with a pair (Σ, φ) where $\Sigma \subseteq \mathcal{A}^{\mathbb{Z}}$ is a two-sided subshift of finite type in a finite alphabet $\mathcal{A}$ equipped with the shift transformation $(x_i)_{i\in\mathbb{Z}} \mapsto (x_{i+1})_{i\in\mathbb{Z}}$, and φ is an integer-valued observable which depends only on the zeroth coordinate, $\varphi((x_i)_{i\in\mathbb{Z}}) = \varphi(x_0) \in \mathbb{Z}$. The function φ will be called the *displacement*. The

countable extension of (Σ, φ) is defined as the two-sided subshift of finite type $\hat{\Sigma}$ in the countable alphabet $\mathcal{A} \times \mathbb{Z}$ with the transition rules

$$(a, m) \to (b, n) \text{ if and only if } a \to b \text{ and } n - m = \varphi(a). \qquad (3.1.1)$$

The notation $a \to b$ means that the symbol a may be followed by the symbol b in sequences belonging to Σ; another way to say this is that the word ab belongs to the language of Σ.
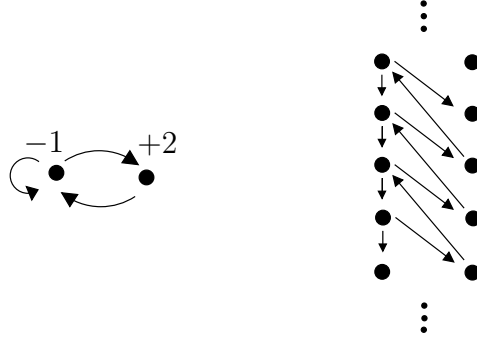

Figure 3.1. Transition Graphs of a Countable Extension

Example 3.1 *Let Σ be the golden mean subshift with the alphabet $\mathcal{A} = \{a, b\}$ in which consecutive b's are forbidden. Assign values $\varphi(a) = -1, \varphi(b) = 2$. The transition graphs of Σ and the induced countable extension $\hat{\Sigma}$ are shown in Figure 3.1. In the graph for $\hat{\Sigma}$ the vertices are arranged by levels and the function φ tells us how many levels up or down each arrow should point.*

It is easy to tell from the transition graph when a subshift of finite type is topologically transitive or topologically mixing (see [13]). Topological transitivity is equivalent to irreducibility of the transition graph, which is the condition that for any pair of vertices a, b, there is a path from a to b and there is a path from b to a. Topological mixing is equivalent to irreducibility and aperiodicity of the transition graph, which requires additionally the existence of two loops in the transition graph whose lengths are relatively prime. In Example 3.1, we see that $\hat{\Sigma}$ is topologically transitive but

not topologically mixing; every loop in the transition graph has length a multiple of three. Σ, on the other hand, is both topologically transitive and topologically mixing.

Let us try to develop the analogy between countable extensions and degree-one liftings. Our findings are summarized in Table 3.1. We notice that in both settings we have a dynamical system on a noncompact space which factors through a countable-to-one topological covering map to a system on a compact space. In both cases, the group of deck transformations that commute with the dynamics is isomorphic to $\mathbb{Z}$. And in both cases there is a displacement function which assigns to a point x the number of fundamental domains to the right or to the left that a point in the fiber above x is carried under the dynamics. In the following paragraphs we develop these ideas in more detail.

<div align="center">

Table 3.1.
Countable Extensions and Degree One Liftings

</div>

Zero-Dimensional Dynamics	One-Dimensional Dynamics
$\begin{array}{ccc} \hat{\Sigma} & \xrightarrow{\hat{\sigma}} & \hat{\Sigma} \\ \pi\downarrow & & \downarrow\pi \\ \Sigma & \xrightarrow{\sigma} & \Sigma \end{array}$	$\begin{array}{ccc} \mathbb{R} & \xrightarrow{F} & \mathbb{R} \\ \pi\downarrow & & \downarrow\pi \\ \mathbb{R}/\mathbb{Z} & \xrightarrow{f} & \mathbb{R}/\mathbb{Z} \end{array}$
$(x_i, n_i)_i \mapsto (x_i, n_i + 1)_i$	$X \mapsto X + 1$
$\varphi(x) = n_1 - n_0,$ for $(x_i, n_i)_i \in \pi^{-1}(x)$	$\varphi_{floor}(x) = \lfloor F(X) \rfloor - \lfloor X \rfloor,$ for $X \in \pi^{-1}(x)$

Countable extensions come with a natural factor structure. Suppose $\hat{\Sigma}$ is the countable extension of (Σ, φ), and denote the left shift transformations by $\hat{\sigma}$ and σ respectively. Then the map

$$\pi\left((x_i, n_i)_{i\in\mathbb{Z}}\right) = (x_i)_{i\in\mathbb{Z}} \tag{3.1.2}$$

gives the semiconjugacy $\sigma \circ \pi = \pi \circ \hat{\sigma}$. This projection is countable-to-one. If a point $(x_i)_{i \in \mathbb{Z}}$ is given, then for each $l \in \mathbb{Z}$ there exists a preimage $(x_i, n_i)_{i \in \mathbb{Z}}$ satisfying $n_0 = l$ and $n_{i+1} - n_i = \varphi(x_i)$ for all i, and by the transition rules in $\hat{\Sigma}$ there are no other preimages. Moreover, the projection $\pi : \hat{\Sigma} \to \Sigma$ is a topological covering map. Indeed, $\pi^{-1}(\Sigma)$ is the countable disjoint union $\bigsqcup_{l \in \mathbb{Z}} \left\{ (x_i, n_i)_{i \in \mathbb{Z}} : n_0 = l \right\}$ and restricting π to any one of these summands yields a homeomorphism onto Σ.

Because our spaces are totally disconnected, the group of deck transformations may be quite large. But from the dynamical point of view, we should only consider those deck transformations that commute with the shift

$$\Gamma := \left\{ \gamma \in Aut\left(\hat{\Sigma}\right) : \pi \circ \gamma = \pi, \ \gamma \circ \hat{\sigma} = \hat{\sigma} \circ \gamma \right\}.$$

Proposition 3.2 *Let $\hat{\Sigma}$ be the countable extension of (Σ, φ). If $\hat{\Sigma}$ is topologically transitive, then the group Γ of shift-commuting deck transformations is isomorphic to $\mathbb{Z}$ with generator $\gamma\left((x_i, n_i)_{i \in \mathbb{Z}}\right) = (x_i, n_i + 1)_{i \in \mathbb{Z}}$.*

Proof It is clear from the definitions that γ and its iterates are shift-commuting deck transformations and form an infinite cyclic group. It remains to show that there are no other shift-commuting deck transformations. Suppose $\eta \in \Gamma$ is arbitrary. Since η preserves the fibers of π, it follows that η must be of the form

$$x = (x_i, n_i)_{i \in \mathbb{Z}} \mapsto (x_i, n_i + k(x))_{i \in \mathbb{Z}}$$

for some $k : \hat{\Sigma} \to \mathbb{Z}$. We must show that k is constant.

Assume temporarily that k is discontinuous. Then there is a point $x \in \hat{\Sigma}$ and there are points y arbitrarily near to x with $k(y) \neq k(x)$. Remember that in shift-spaces, nearness is measured by the number of symbols around the zeroth position that x and y have in common. But if x and y agree in the zeroth position and $k(x) \neq k(y)$, then $\eta(x)$ and $\eta(y)$ will already differ in the zeroth position. This contradicts the continuity of η. Therefore k must be continuous.

Since η commutes with the shift, it follows that $k(\hat{\sigma}(x)) = k(x)$ for all x. That means that k is constant along orbits. By hypothesis, $\hat{\Sigma}$ is topologically mixing, and

therefore transitive. In a complete metric space, transitivity implies the existence of a point with a dense orbit [22]. Since k is integer-valued, continuous, and constant along a dense orbit, it must be constant everywhere. ∎

Finally, we remark that the more conventional displacement function in the one-dimensional theory is $\varphi_{con}(x) = F(X) - X$. But there is no harm in introducing the floor function $\lfloor \cdot \rfloor$ into the definition because φ_{con} and φ_{floor} are cohomologous. Explicitly, $\varphi_{floor} = \varphi_{con} + g - g \circ f$ where $g(x) = X - \lfloor X \rfloor$. Therefore the limiting ergodic averages of φ_{con} and of φ_{floor} behave identically, so both functions yield the same rotation-theoretic results.

Thus far, our discussion of countable extensions has been purely topological. If we want a fuller understanding, we must consider the measures supported by these systems. The projection $\pi : \hat{\Sigma} \to \Sigma$ associated with a countable extension induces a projection map π_* on measures, which sends a Borel measure ν on $\hat{\Sigma}$ to the Borel measure on Σ given by the formula $(\pi_*\nu)(A) = \nu(\pi^{-1}A)$. If ν is shift-invariant (resp. ergodic, finite, a probability measure), then so is $\pi_*\nu$. However, unlike in compact dynamics, as a map on the spaces of shift-invariant Borel probability measures, π_* need not be surjective. We will use this fact to great advantage, arguing that certain measures on Σ do not lift, i.e., are not the projection of any invariant measure from $\hat{\Sigma}$. We record now a crude but necessary condition for a measure to lift. If μ is a shift-invariant probability measure on Σ, then we call the average displacement $\int \varphi \, d\mu$ the *drift* of μ. A measure is called *drift-free* if its drift is zero.

Theorem 3.3 *Let $\hat{\Sigma}$ be the countable extension of (Σ, φ). If ν is an invariant, Borel probability measure for $\hat{\Sigma}$, then its projection $\pi_*\nu$ is necessarily drift-free, that is, $\int \varphi \, d\pi_*\nu = 0$.*

Proof By considering ergodic decompositions, we may assume without loss of generality that ν is ergodic. Then $\pi_*\nu$ is also ergodic. Now suppose toward contradiction

that $\int \varphi \, d\pi_* \nu = c > 0$; the proof for $c < 0$ is analogous. By Birkhoff's pointwise ergodic theorem,

$$\lim_{t\to\infty} \frac{1}{t} \sum_{i=0}^{t-1} \varphi(x_i) = c, \quad \text{for } \pi_* \nu \text{ a.e. } (x_i)_{i\in\mathbb{Z}} \text{ in } \Sigma.$$

Passing to preimages and remembering the transition rules in $\hat{\Sigma}$, we have

$$\lim_{t\to\infty} \frac{n_t - n_0}{t} = c, \quad \text{for } \nu \text{ a.e. } (x_i, n_i)_{i\in\mathbb{Z}} \text{ in } \hat{\Sigma}. \tag{3.1.3}$$

We claim that Equation 3.1.3 is incompatible with the shift-invariance of the measure ν. If the measure ν has any atoms, then by ergodicity it is concentrated on a periodic orbit, which already contradicts Equation 3.1.3. Now assume that ν is nonatomic. Let $A_{[-m,m]} = \{(x_i, n_i)_{i\in\mathbb{Z}} : -m \le n_0 \le m\}$, and fix m sufficiently large that $\nu(A_{[-m,m]}) > \frac{1}{2}$. The convergence in Equation 3.1.3 is pointwise, but by Egoroff's theorem we can find a slightly smaller subset $B \subseteq A_{[-m,m]}$, but still with $\frac{1}{2} < \nu(B)$, on which the convergence is uniform. Fix $T > 4m/c$ sufficiently large so that

$$\frac{n_T - n_0}{T} > \frac{c}{2}, \quad \text{for all } (x_i, n_i)_{i\in\mathbb{Z}} \in B.$$

So if $(x_i, n_i) \in B$, then $n_0 \ge m$ and $n_T - n_0 > \frac{4m}{c}\frac{c}{2} = 2m$, whence $n_T > m$. This shows that the Tth preimage of A under the shift is disjoint from B. But B has measure greater than $\frac{1}{2}$, and by the invariance of ν, so does the Tth preimage of A. This is a contradiction. ∎

Theorem 3.3 is not surprising if we think in terms of rotation theory. Invariance of the measure ν on $\hat{\Sigma}$ should mean that in some sense there is just as much displacement of mass in the positive direction as there is in the negative direction. So we should expect the projected measure $\pi_* \nu$ to have average rotation zero. There is also the heuristic reasoning $\int_\Sigma \varphi \, d\pi_* \nu = \int_{\hat{\Sigma}} n_1 - n_0 \, d\nu = \int_{\hat{\Sigma}} n_1 \, d\nu - \int_{\hat{\Sigma}} n_0 \, d\nu = 0$ by the shift-invariance of ν. Unfortunately, this argument is not rigorous; when ν has heavy enough tails the last two integrals diverge.

The drift-free condition is necessary for a measure to lift, but not sufficient. In the proof of Theorem 3.7 we give more delicate arguments showing that the most

important drift-free measure associated to a countable extension nevertheless does not lift.

3.2   Entropy and the Maximal Drift-Free Measure

We wish to study our countable extensions from the point of view of entropy. For a countable state topological Markov chain, there are several possible definitions of the entropy. We are most interested in the *Gurevich entropy*, denoted $h_{Gur}(\cdot)$. It may be defined as the supremum of metric entropies over all shift-invariant Borel probability measures supported on the countable state chain [12]. Thus, when we prove in Theorem 3.7 that a countable extension has no measure of maximal entropy, we mean quite naturally the Gurevich entropy.

Two other characterizations of the Gurevich entropy will be relevant for us. For transitive chains, the Gurevich entropy is given by the limit

$$\lim_{n\to\infty} \frac{1}{n} \log \#\{\text{length } n \text{ words which start and end with some fixed symbol } a\},$$

$$(3.2.1)$$

where by transitivity, this quantity does not depend on the choice of a. In the transition graph representation of the chain, this limit measures the growth rate of the number of loops which start and end at some fixed vertex. The third characterization of Gurevich entropy may also be given in terms of the transition graph model. Each finite subgraph corresponds to a finite-state subchain which is a compact dynamical system with its own well-defined topological entropy. The Gurevich entropy is equal to the supremum of topological entropies over all such subchains. The equivalence of these three characterizations was proved by Gurevich [11], [12].

Our study of the entropy of countable extensions stems from the work of Misiurewicz and Tolosa [17]. Although the terminology is slightly different, their work contains the following restricted variational principle for countable extensions.

Theorem 3.4 *(Misiurewicz, Tolosa, [17]) Suppose $\hat{\Sigma}$ is the countable extension of (Σ, φ). Then the Gurevich entropy of $\hat{\Sigma}$ is the supremum of metric entropies over all drift-free invariant ergodic probability measures on Σ.*

We want to show that this supremum is uniquely attained and we want to describe explicitly the measure which attains it. To do that, we will need to apply a few results from thermodynamic formalism. In particular, we will study the *pressure* of the scaled displacement function $\beta\varphi$ for $\beta \in \mathbb{R}$. For our purposes, we may define this pressure in terms of the following variational principle:

$$\mathcal{P}(\beta\varphi) := \sup \left\{ h_\Sigma(\mu) + \beta \int \varphi \, d\mu \; \middle| \; \mu \text{ invariant probability measure on } \Sigma \right\},$$

where $h_\Sigma(\mu)$ denotes the metric entropy of μ. The quantity $h_\Sigma(\mu) + \beta \int \varphi \, d\mu$ is called the *free energy* of μ (with respect to the potential function $\beta\varphi$). A shift-invariant probability measure whose free energy attains this supremum is called an *equilibrium state* for the observable $\beta\varphi$.

Theorem 3.5 *Let $\hat{\Sigma}$ be the countable extension of (Σ, φ). If $\hat{\Sigma}$ is topologically transitive and Σ is topologically mixing, then there exists a measure μ_0 on Σ, called the* maximal drift-free measure, *with the following properties:*

- *μ_0 is an ergodic drift-free shift-invariant probability measure and has strictly larger entropy than any other drift-free shift-invariant probability measure. Thus, it uniquely achieves the supremum in Theorem 3.4*

- *μ_0 is the unique equilibrium state for $\beta_0\varphi$, where β_0 minimizes the pressure $\mathcal{P}(\beta_0\varphi) = \min_{\beta \in \mathbb{R}} \mathcal{P}(\beta\varphi)$ (and this uniquely determines β_0).*

Moreover, the Gurevich entropy of $\hat{\Sigma}$ is given by

$$h_{Gur}(\hat{\Sigma}) = h_\Sigma(\mu_0) = \mathcal{P}(\beta_0\varphi)$$

Proof Since we have a locally constant observable function φ on a topologically mixing subshift of finite type Σ, we may apply some strong results from thermodynamic

formalism. The first of these results tells us that the pressure function $\beta \mapsto \mathcal{P}(\beta\varphi)$ is convex and real analytic (see [21]). Now we will show that the pressure function takes a minimum value. Since $\hat{\Sigma}$ is topologically mixing, there must be a path in its transition graph from some vertex $(a, 0)$ to the vertex $(a, 1)$; let n denote the length of this path. Projecting to the transition graph for Σ, we find a loop of length n from a to a with the sum of φ along the loop equal to 1. This loop corresponds to a periodic point of Σ. Consider the purely atomic probability measure distributed uniformly along this periodic orbit. It is an invariant measure with entropy zero, and the integral of φ with respect to this measure is $1/n$. Thus, the free energy of this measure is β/n, so we obtain the inequality $\mathcal{P}(\beta\varphi) \geq \beta/n$, $\beta \in \mathbb{R}$. It follows that $\mathcal{P}(\beta\varphi) \to +\infty$ as $\beta \to +\infty$. Similarly, if we use a path in the transition graph of $\hat{\Sigma}$ from some vertex $(a, 0)$ to the vertex $(a, -1)$, we may conclude that $\mathcal{P}(\beta\varphi) \to +\infty$ as $\beta \to -\infty$. Now from convexity and real analyticity, it follows that $\mathcal{P}(\beta\varphi)$ is minimized at a unique point β_0.

The second major result we need from thermodynamic formalism is that for each β, the observable $\beta\varphi$ has a unique equilibrium state (see [20]). Define μ_0 to be the unique equilibrium state corresponding to β_0. By its uniqueness it is ergodic. Consider the graph of the line $y = h_\Sigma(\mu_0) + \beta \int \varphi \, d\mu_0$ and the pressure curve $y = \mathcal{P}(\beta\varphi)$. This line intersects the pressure curve at the point β_0 because μ_0 is an equilibrium measure. This line lies below the pressure curve by the definition of pressure. By real analyticity, this line must be a tangent line, and since it is tangent at the minimum point, it must have slope zero. We conclude that μ_0 has zero drift.

For measures with zero drift, free energy equals entropy. But μ_0 is the unique equilibrium state for the parameter β_0. This implies that the entropy of μ_0 is equal to $\mathcal{P}(\beta_0\varphi)$ and is strictly larger than the entropy of any other drift-free shift-invariant probability measure.

∎

In the course of the proof we also demonstrated the following fact, which will be useful in Chapter 4.

Corollary 3.6 *Let $\hat{\Sigma}$ be the countable extension of (Σ, φ). If $\hat{\Sigma}$ is topologically transitive and Σ is topologically mixing, then the pressure function $\beta \mapsto \mathcal{P}(\beta\varphi)$ maps $\mathbb{R}$ surjectively onto $[h_{Gur}(\hat{\Sigma}), \infty)$.*

Let us illustrate our theorem by making explicit calculations for the countable extension in Example 3.1. The measure μ_0 is known to be Markov, i.e., it is given by a probability vector and a stochastic matrix. By the drift-free condition, the probability vector must be $\begin{bmatrix} \frac{2}{3} & \frac{1}{3} \end{bmatrix}$. By the invariance of the measure, the stochastic matrix must be $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix}$. The Gurevich entropy of $\hat{\Sigma}$ is therefore $\frac{2}{3}\log(2)$, the metric entropy of this Markov measure. This number is smaller than $\log(\frac{1+\sqrt{5}}{2})$, the topological entropy of Σ, which reflects the fact that the Parry measure (measure of maximal entropy for Σ) has nonzero drift.

3.3 No Measure of Maximal Entropy

Theorem 3.7 *Let $\hat{\Sigma}$ be the countable extension of (Σ, φ). If $\hat{\Sigma}$ is topologically transitive and Σ is topologically mixing, then $\hat{\Sigma}$ has no measure of maximal entropy.*

Proof Consider the maximal drift-free measure μ_0 on Σ identified in Theorem 3.5. In the first step of the proof, we regard ergodic sums of the displacement function φ as random variables on the measure space (Σ, μ_0). A central limit theorem applies. We verify that the asymptotic variance term is positive. In the second step of the proof, we use the central limit theorem to show that no invariant Borel probability measure on $\hat{\Sigma}$ projects to the measure μ_0. Thus, even though μ_0 is drift-free, it does not lift. Finally, in the third step of the proof we use the fact that μ_0 does not lift to show that there is no measure of maximal entropy for $\hat{\Sigma}$.

Step One: Equip the topological Markov chain Σ with the measure μ_0 from Theorem 3.5. The measure μ_0 is an equilibrium state, so we may apply a central limit theorem (see [8], Theorem 1.27) to the distribution of ergodic sums of any Hölder continuous function on Σ. The displacement function φ is Hölder continuous because it depends on only the zeroth coordinate. Its expected value is zero because the

measure μ_0 is drift-free. Introduce random variables S_φ^t, $t = 1, 2, \ldots$ on Σ by the formula

$$S_\varphi^t \left((x_i)_{i \in \mathbb{Z}} \right) = \varphi(x_0) + \varphi(x_1) + \ldots + \varphi(x_{t-1}).$$

S_φ^t records the sum of the displacement function φ along the first t symbols (ignoring negative coordinates; at this moment we do not care that our shift-space is two-sided and we forget about the past). Now consider the *asymptotic variance* $\sigma_{asy}^2 = \lim_{t \to \infty} \frac{1}{t} Var(S_\varphi^t)$. The central limit theorem states that this limit exists and is finite, and moreover, if $\sigma_{asy}^2 > 0$, then the random variables S_φ^t, properly scaled, converge in distribution to the standard normal distribution:

$$\text{If } \sigma_{asy}^2 > 0, \text{ then } \frac{S_\varphi^t}{\sigma_{asy}\sqrt{t}} \xrightarrow{dist} \text{Standard Normal.} \tag{3.3.1}$$

Bowen gives a condition for determining when the asymptotic variance σ_{asy}^2 is positive [8]. We have $\sigma_{asy}^2 = 0$ if and only if φ is homologous to zero by a Hölder continuous function, that is, there is some Hölder continuous u such that

$$\varphi = u \circ \sigma - u, \quad \mu_0\text{-almost everywhere.}$$

Composing with the shift and taking sums, we must have

$$S_\varphi^t = u \circ \sigma^t - u, \quad \mu_0\text{-almost everywhere, } t \in \mathbb{N}. \tag{3.3.2}$$

Since u is continuous function on a compact metric space, it is bounded by some constant M, and so the right-hand side of equation 3.3.2 is bounded by $2M$. This bound is independent of t. By hypothesis, $\hat{\Sigma}$ is topologically transitive, so we can find a path in its transition graph from some vertex $(a, 0)$ to the vertex $(a, 1)$ with some path length n. Projecting this path into the transition graph for Σ we have a length n loop from a to a with net displacement 1. Fix $k > 2M$ and consider the cylinder set corresponding to k repetitions of this loop. Then S_φ^{kn} is identically equal to k on this cylinder set. Moreover, Theorem 3.5 tells us that μ_0 has full support, so that this cylinder set has positive measure. This contradicts Equation 3.3.2. We may conclude that φ is not homologous to zero and therefore $\sigma_{asy}^2 > 0$ and Equation 3.3.1 is valid.

Step Two: Consider the projection $\pi : \hat{\Sigma} \to \Sigma$ defined in Equation 3.1.2 and the corresponding push-forward operator on measures. Suppose that ν is a Borel probability measure on $\hat{\Sigma}$ with $\pi_*(\nu) = \mu_0$. Since μ_0 is nonatomic, so is ν. We will prove that ν is not shift-invariant.

Let $A_l \subset \hat{\Sigma}$ denote those sequences which begin at level l, that is, $A_l = \{(x_i, n_i)_{i \in \mathbb{Z}} : n_0 = l\}$. By an appropriate choice of l we may assume that $\nu(A_l) > 0$, and after relabeling the levels, we may assume that $l = 0$. Let $A_{[-m,m]} := \{(x_i, n_i)_{i \in \mathbb{Z}} : -m \le n_0 \le m\}$. By the continuity of measure, we may make an appropriate choice of m so that

$$\nu(\hat{\Sigma} \setminus A_{[-m,m]}) < \frac{1}{2}\nu(A_0). \tag{3.3.3}$$

Choose $\delta > 0$ small enough so that the measure of the interval $[-\delta, \delta]$ under the Gaussian distribution is strictly less than $\frac{1}{2}\nu(A_0)$. Applying the central limit theorem (Equation 3.3.1), we may fix a value of t sufficiently large so that $m < \delta \sigma_{asy}\sqrt{t}$ and

$$\mu_0\left(\left\{x \in \Sigma : \left|\frac{S_\varphi^t(x)}{\sigma_{asy}\sqrt{t}}\right| \le \delta\right\}\right) < \frac{1}{2}\nu(A_0).$$

Passing to a subset, we have

$$\mu_0(\{x \in \Sigma : |S_\varphi^t(x)| \le m\}) < \frac{1}{2}\nu(A_0). \tag{3.3.4}$$

Consider the set $B = \{(x_i, n_i)_{i \in \mathbb{Z}} : n_t = 0\}$ of sequences which reach level 0 at time t. It is the tth preimage of A_0 under the shift transformation in $\hat{\Sigma}$. Partition this set into $B \cap A_{[-m,m]}$ and $B \setminus A_{[-m,m]}$. We have

$$B \cap A_{[-m,m]} \subseteq \{(x_i, n_i)_{i \in \mathbb{Z}} : |n_t - n_0| \le m\} = \pi^{-1}(\{x \in \Sigma : |S_\varphi^t(x)| \le m\}).$$

Combining Inequalities 3.3.3 and 3.3.4 and remembering that ν projects to μ, we have

$$\nu(B) = \nu(B \cap A_{[-m,m]}) + \nu(B \setminus A_{[-m,m]})$$
$$\le \mu_0(\{x \in \Sigma : |S_\varphi^t(x)| \le m\}) + \nu(\hat{\Sigma} \setminus A_{[-m,m]})$$
$$< \frac{1}{2}\nu(A_0) + \frac{1}{2}\nu(A_0) = \nu(A_0).$$

It follows that ν is not a shift-invariant measure.

Step Three: The Gurevich entropy $h_{Gur}(\hat{\Sigma})$ is the supremum of metric entropies among all invariant probability measures on $\hat{\Sigma}$. We wish to show that this supremum is not attained.

Suppose ν is an arbitrary shift-invariant probability measure on $\hat{\Sigma}$. We know from step two that $\pi_*\nu \neq \mu_0$. We know from Theorem 3.3 that $\pi_*\nu$ is a drift-free measure. Therefore, by Theorem 3.5,

$$h_\Sigma(\pi_*\nu) < h_\Sigma(\mu_0) = h_{Gur}(\hat{\Sigma}). \tag{3.3.5}$$

Since π is is a countable-to-one factor map, π_* preserves metric entropy (see [9] Theorem 4.1.15),

$$h_{\hat{\Sigma}}(\nu) = h_\Sigma(\pi_*\nu). \tag{3.3.6}$$

Combining Equations 3.3.5 and 3.3.6, we see that $\hat{\Sigma}$ has no measure of maximal entropy.

∎

3.4   Explicit Calculations

Let us show how to apply Theorem 3.5 to make explicit calculations for concrete examples of countable extensions. We begin by recording some formulas from the general theory of equilibrium states. Suppose we have a countable extension $\hat{\Sigma}$ of a pair (Σ, φ), with $\hat{\Sigma}$ topologically transitive and Σ topologically mixing. We continue to use $\mathcal{A}$ to denote both the (finite) alphabet of Σ and the vertex set of the corresponding transition graph. The *transition matrix* A is the binary matrix with rows and columns indexed by $\mathcal{A}$ and entries

$$A(a,b) = \begin{cases} 1, & \text{if } a \to b \\ 0, & \text{otherwise.} \end{cases} \tag{3.4.1}$$

We define also for each $\beta \in \mathbb{R}$ the *weighted transition matrix* M_β with entries

$$M_\beta(a,b) = \begin{cases} e^{\beta\varphi(a)}, & \text{if } a \to b \\ 0, & \text{otherwise.} \end{cases} \tag{3.4.2}$$

Theorem 3.8 *(see, eg., [20]) Let Σ, φ, M_β be as above. Then the pressure of $\beta\varphi$ is given by the logarithm of the spectral radius of the weighted transition matrix*

$$\mathcal{P}(\beta\varphi) = \log \operatorname{rad}(M_\beta).$$

Theorem 3.9 *(see, eg., [20]) Let Σ, φ, M_β be as above. Then the equilibrium state for $\beta\varphi$ is the stationary Markov measure with stochastic matrix P_β and probability vector p_β given by*

$$P_\beta(a,b) = \frac{r(b)}{\lambda r(a)} M_\beta(a,b), \quad p_\beta(a) = r(a)l(a)$$

where λ, r, l (they also depend on β) are the largest eigenvalue and the strictly positive right and left eigenvectors given by the Perron Frobenius theorem, $M_\beta r = \lambda r$, $lM_\beta = \lambda l$, scaled in such a way that $\sum_{a\in\mathcal{A}} l(a)r(a) = 1$.

We remark that M_β is irreducible and aperiodic because Σ was assumed to be topologically mixing. That is why the Perron Frobenius theorem gives a strictly positive eigenvector.

Combining Theorem 3.8 with Theorem 3.5, we see that the problem of computing Gurevich entropy is reduced to the problem of minimizing the leading eigenvalue of the weighted transition matrix. When the vertex set $\mathcal{A}$ is small enough, we can sometimes solve this problem quickly and explicitly.
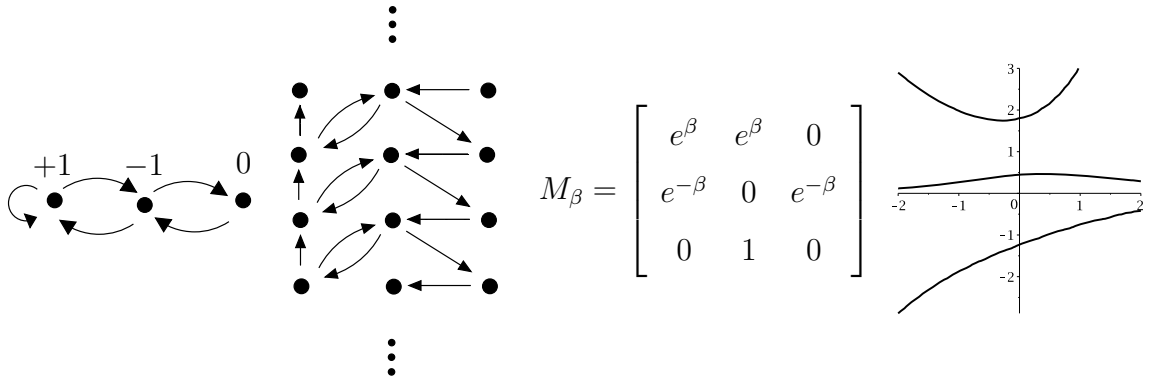

Figure 3.2. Entropy Calculations for a Countable Extension

For example, consider the countable extension $\hat{\Sigma}$ of the pair (Σ, φ) shown in Figure 3.2. The characteristic polynomial is $-\lambda^3 + e^\beta \lambda^2 + (1 + e^{-\beta})\lambda - 1$, and we wish to minimize the leading root of this polynomial as we allow β to vary. Set the characteristic polynomial equal to zero and look at the solution set in the (β, λ) plane. We can solve for e^β explicitly with the quadratic formula. The discriminant is $\lambda^6 - 2\lambda^4 - 2\lambda^3 + \lambda^2 - 2\lambda + 1$. The Gurevich entropy is the logarithm of the largest real root of this discriminant polynomial; the approximate value is $h_{Gur}(\hat{\Sigma}) \approx \log 1.7549$.

In light of Theorem 3.9, it is easy to see when the system Σ equipped with the maximal drift-free measure μ_0 and the partition by time-zero cylinders is a Bernoulli process.

Proposition 3.10 *Let $\hat{\Sigma}$ be the countable extension of (Σ, φ). Suppose that $\hat{\Sigma}$ is topologically transitive and Σ is topologically mixing. Then Σ equipped with the maximal drift-free measure μ_0 and the partition by time-zero cylinders is a Bernoulli process if and only if the transition graph of Σ is* complete *in the sense that for all $a, b \in \mathcal{A}$, there is an arrow $a \to b$.*

Proof If the transition graph is complete, then the weighted transition matrix of Equation 3.4.2 has all columns equal. Therefore its rank is one and the strictly positive eigenvector r is just the common column vector. It follows that the stochastic matrix P_β of Theorem 3.9 has all its rows equal, and so the corresponding Markov process is in fact Bernoulli. Conversely, if all rows of P_β are equal, then by transitivity P_β can contain no zeros, and hence the transition graph is complete. ∎

4. APPLICATIONS

Our findings are relevant in one-dimensional dynamics in the study of circle maps. If we have a piecewise monotone degree one map of the circle, a transitive lifting of this map to $\mathbb{R}$, and an appropriate Markov partition, then the induced symbolic dynamical system is a countable extension as defined in Chapter 3. This allows us to compute a certain entropy for our degree one lifting. Moreover, for every λ greater than or equal to the exponential of this entropy, we can construct a conjugacy to a map of constant slope λ.

4.1   Degree One Circle Maps with Markov Partitions

Throughout this chapter, we will assume that $F : \mathbb{R} \to \mathbb{R}$ is a transitive lifting of a degree one map of the circle $f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$. We assume also the existence of a Markov partition. Explicitly, we require that $[0, 1]$ (and also $\mathbb{R}/\mathbb{Z}$) is the union of a finite collection $\mathcal{V}$ of closed intervals (in the circle these are closed arcs) with pairwise disjoint interiors. Then $\mathbb{R}$ is the union of the countable collection $\mathcal{V} \times \mathbb{Z}$ of closed intervals with pairwise disjoint interiors given by setting (v, m) equal to the translation $v + m$ of the interval v by m units, $v \in \mathcal{V}, m \in \mathbb{Z}$. Moreover, we assume that for all $(v, m) \in \mathcal{V} \times \mathbb{Z}$, the restriction $F|_{(v,m)}$ is monotone and the image $F\big((v, m)\big)$ is a union of intervals from $\mathcal{V} \times \mathbb{Z}$. This is the Markov property – it means that if we study dynamics symbolically by coding points according to their itineraries, then we obtain subshifts of finite type. In fact, these symbolic systems will have the structure of countable extensions.

Let us give explicitly the construction of these symbolic systems. This construction will be easier to read with a concrete example in mind; the reader may wish to look ahead to Example 4.2. Corresponding to the circle map f we construct a labeled

directed graph G^* with vertex set $\mathcal{V}$. Corresponding to the lifting F we obtain an unlabeled directed graph $\hat{G}^*$ with vertices $\mathcal{V} \times \mathbb{Z}$. The dynamics of F determine where to draw the arrows and what labels to assign them, as follows:

$$v \xrightarrow{l} w \text{ in } G^* \quad \text{iff} \quad (v, m) \to (w, m + l) \text{ in } \hat{G}^* \quad \text{iff} \quad F\big((v, m)\big) \supseteq (w, m + l), \quad (4.1.1)$$

and by the degree one property, this definition does not depend on the choice of m. The notation $v \xrightarrow{l} w$ means there is an arrow a pointing from v to w with label l, and then for each $m \in \mathbb{Z}$ we denote by (a, m) the arrow $(v, m) \to (w, m + l)$. We use the symbol $\mathcal{A}$ for the set of arrows of G^*, $\varphi : \mathcal{A} \to \mathbb{Z}$ for the labels (on the arrows); the arrow set for $\hat{G}^*$ is $\mathcal{A} \times \mathbb{Z}$. We will also use the notation $init(a)$, $term(a)$ for the initial and terminal points of an arrow $a \in \mathcal{A}$.

In general, the graph G^* may have multiple arrows pointing between the same two vertices. This happens when an arc $v \in \mathcal{V}$ has an image under f that wraps around the circle multiple times. Let A^* denote the *transition matrix* for G^*; its vw-entry is the number of arrows pointing from v to w,

$$A^*(v, w) = \#\{a \in \mathcal{A} : init(a) = v, term(a) = w\}. \quad (4.1.2)$$

We also define for each $\beta \in \mathbb{R}$ the *weighted transition matrix* matrix M_β^* with entries

$$M_\beta^*(v, w) = \sum_{a:init(a)=v, term(a)=w} e^{\beta \varphi(a)}. \quad (4.1.3)$$

The graph $\hat{G}^*$ can have only single arrows; its transition matrix T^* is binary and is given by

$$T^*\big((v, m), (w, m+l)\big) = \begin{cases} 1, & \text{if } \exists a \in \mathcal{A} : init(a) = v, term(a) = w, \varphi(a) = l \\ 0, & \text{otherwise.} \end{cases} \quad (4.1.4)$$

We have not yet defined a countable extension – we still need to construct $\hat{\Sigma}$ and Σ. Because G^* can have multiple arrows, we must take the arrow set $\mathcal{A}$ as the alphabet for our shift space. In other words, we construct our chain from the *dual graph*. The dual graph to G^* uses the arrows of G^* as its vertices (they are still labeled by φ) and

allows a transition from a to b if and only if $init(b) = term(a)$. We write G without a star for this dual graph and (Σ, φ) for the corresponding topological Markov chain and observable function. Its transition matrix and weighted transition matrices have already been defined in equations 3.4.1 and 3.4.2. Similarly we write $\hat{G}$ for the dual graph to $\hat{G}^*$ and we write $\hat{\Sigma}$ for the corresponding countable-state topological Markov chain. Its transition matrix T is given by

$$T\big((a, m), (b, n)\big) = \begin{cases} 1, & \text{if } term(a) = init(b) \text{ and } \varphi(a) = n - m \\ 0, & \text{otherwise.} \end{cases} \tag{4.1.5}$$

Comparing the transition rules of equation 4.1.1 with the transition rules of equation 3.1.1, we see that $\hat{\Sigma}$ is the countable extension of (Σ, φ). This concludes our construction.

Remark 4.1 *The graphs G^*, $\hat{G}^*$ and their corresponding matrices form an unnecessary intermediate stage in the above construction. It is possible to eliminate this stage by replacing $\mathcal{V} \times \mathbb{Z}$ with the finer partition $(\mathcal{V} \times \mathbb{Z}) \vee F^{-1}(\mathcal{V} \times \mathbb{Z})$. Indeed, there is a natural identification of this refined partition with $\mathcal{A} \times \mathbb{Z}$; a nondegenerate interval of the form $(v, m) \cap F^{-1}\big((w, m + l)\big)$ corresponds to (a, m) where a is the labelled arrow $v \xrightarrow{l} w$. Thus, from this finer partition we can obtain G and $\hat{G}$ directly. Our choice to make such a long construction is motivated by applications. When we want theoretical results, we will use $\hat{\Sigma}$ and Σ because our theory of countable extensions applies. But when we want to make numerical calculations, we will use the intermediate stage, because we want our matrices to have as few rows and columns as possible. This point of view is justified by lemma 4.3.*

Example 4.2 *Let F be the piecewise affine "connect-the-dots" map with turning points $F(k) = k - 1$, $F(k + \frac{1}{2}) = k + 2$, $k \in \mathbb{Z}$ and Markov partition $\mathcal{V} \times \mathbb{Z}$ generated by $\mathcal{V} = \{[0, \frac{1}{2}], [\frac{1}{2}, 1]\}$. Figure 4.1 depicts F and the corresponding circle map f as well as several of the associated directed graphs and matrices.*
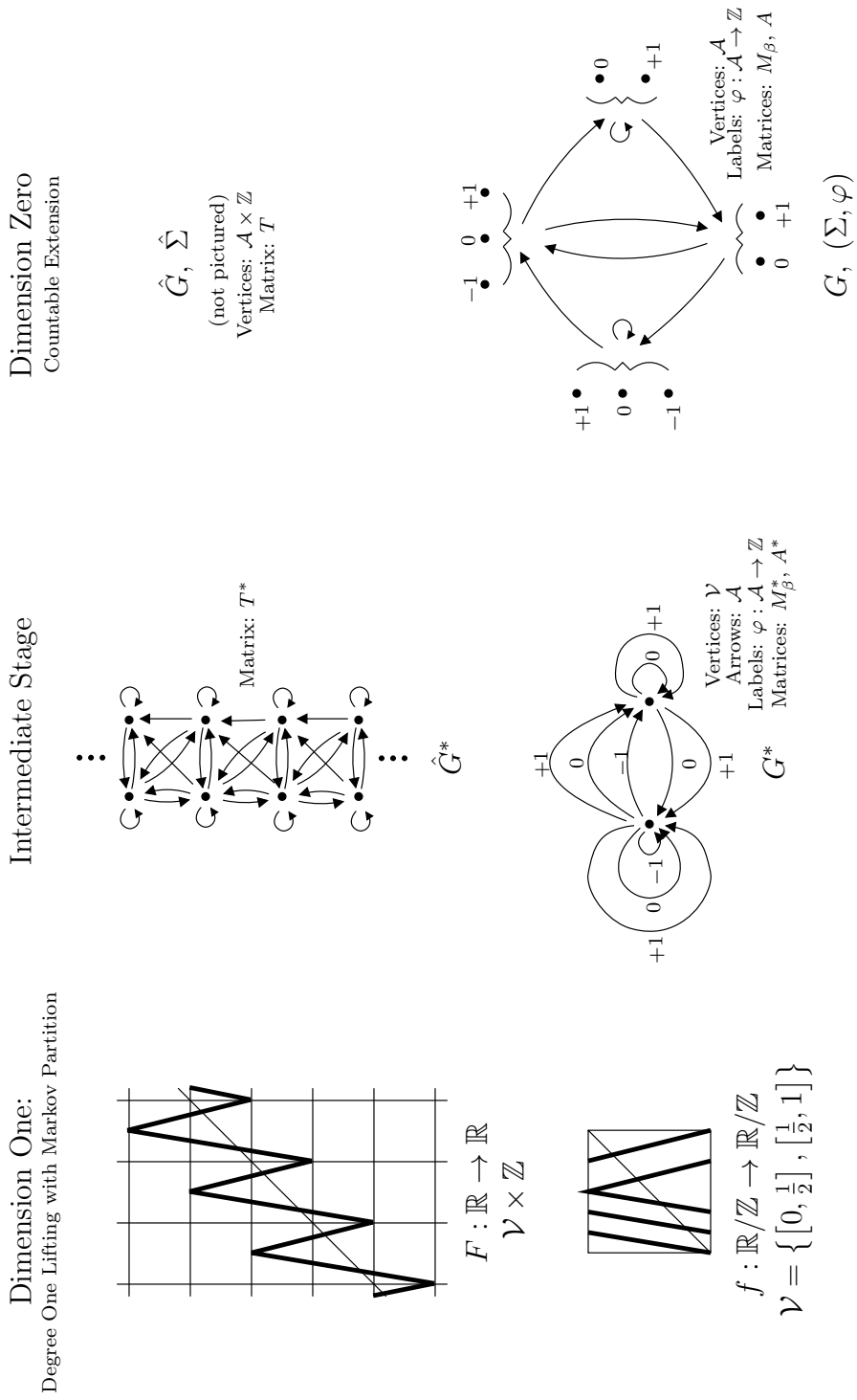
Dimension Zero
Countable Extension

$\hat{G}, \hat{\Sigma}$

(not pictured)
Vertices: $\mathcal{A} \times \mathbb{Z}$
Matrix: T

Vertices: $\mathcal{A}$
Labels: $\varphi : \mathcal{A} \to \mathbb{Z}$
Matrices: M_β, A

$G, (\Sigma, \varphi)$

Intermediate Stage

Matrix: T^*

$\hat{G}^*$

Vertices: $\mathcal{V}$
Arrows: $\mathcal{A}$
Labels: $\varphi : \mathcal{A} \to \mathbb{Z}$
Matrices: M_β^*, A^*

G^*

Dimension One:
Degree One Lifting with Markov Partition

$F : \mathbb{R} \to \mathbb{R}$
$\mathcal{V} \times \mathbb{Z}$

$f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$
$\mathcal{V} = \left\{ \left[0, \frac{1}{2}\right], \left[\frac{1}{2}, 1\right] \right\}$

Figure 4.1. Symbolic Systems Derived from a Markov Circle Map

Lemma 4.3 *Let $G^* = (\mathcal{V}, \mathcal{A}, \varphi)$ be a labeled directed graph, G its dual graph. Let M_β^*, M_β be the respective weighted transition matrices as defined in equations 4.1.3 and 3.4.2. If r^* is an eigenvector with $M_\beta^* r^* = \lambda r^*$, then the vector with entries*

$$r(a) = e^{\beta\varphi(a)} r(term(a))$$

satisfies $M_\beta r = \lambda r$. Moreover, all eigenvectors of M_β are obtained in this way from eigenvectors of M_β^*, with the possible exception that M_β may have additional eigenvectors with eigenvalue $\lambda = 0$.

Proof   The matrix M_β^* should be regarded as representing a linear operator on the space $\mathbb{R}^{\mathcal{V}}$. We will abuse notation and regard $\mathcal{V}$ not only as the index set for $\mathbb{R}^{\mathcal{V}}$, but also as the basis, so that the symbol v represents the column vector with a 1 in position v and zeros elsewhere. We may represent a column vector as a sum of coefficients times basis vectors, for example, $r^* = \sum_{w \in \mathcal{V}} r^*(w) w$. If we multiply M_β^* by a basis vector w, the result is the wth column vector of M_β^*,

$$M_\beta^* w = \sum_{b \in \mathcal{A}: term(b)=w} e^{\beta\varphi(b)} init(b). \tag{4.1.6}$$

In the same way, we regard $\mathcal{A}$ as both index set and basis for the linear space $\mathbb{R}^{\mathcal{A}}$ on which M_β acts. We have

$$M_\beta b = \sum_{a \in \mathcal{A}: term(a)=init(b)} e^{\beta\varphi(a)} a.$$

Consider the linear subspace $\mathcal{F} \subseteq \mathbb{R}^{\mathcal{A}}$ spanned by vectors of the form

$$f_w := \sum_{b \in \mathcal{A}: term(b)=w} e^{\beta\varphi(b)} b, \quad w \in \mathcal{V}.$$

Now we may write more simply $M_\beta b = f_{init(b)}$. This shows that the range of M_β is contained in the subspace $\mathcal{F}$, so if we wish to find eigenvectors with nonzero eigenvalues it suffices to look at the restriction $M_\beta|_{\mathcal{F}}$. By linearity we have

$$M_\beta f_w = \sum_{b \in \mathcal{A}: term(b)=w} e^{\beta\varphi(b)} f_{init(b)} \tag{4.1.7}$$

Comparing equations 4.1.6 and 4.1.7, we see that $M_\beta|_{\mathcal{F}}$ is conjugate to M_β^* by the linear isomorphism $w \mapsto f_w$. This isomorphism carries an eigenvector $r^* = \sum_{w \in \mathcal{V}} r^*(w)w$ to the eigenvector

$$r = \sum_{w \in \mathcal{V}} r^*(w)f_w = \sum_{b \in \mathcal{A}} r^*(term(b))e^{\beta\varphi(b)}b,$$

which agrees with the formula for r in the statement of the theorem. ∎

One more observation is necessary before we can apply the theory of countable extensions to transitive lifts of degree one circle maps. Under the assumption that F is transitive, it follows that all of the dynamical systems F, f, $\hat{\Sigma}$, and Σ are topologically mixing and that all of the matrices T^*, M_β^*, A^*, T, M_β, and A are irreducible and aperiodic. To see this, recall that the only way for a continuous map of the real line to be transitive but not mixing is if interchanges the intervals $(-\infty, c)$ with (c, ∞) for some $c \in \mathbb{R}$ (see [3, pp. 156-159]; the result is stated for interval maps but generalizes easily to maps on the real line). No such c can exist for a degree one lifting, and so F is topologically mixing. Then, using paths through interiors of partition elements of $\mathcal{V} \times \mathbb{Z}$ or of the finer partition $(\mathcal{V} \times \mathbb{Z}) \vee F^{-1}(V \times \mathbb{Z})$ we can produce the necessary paths in the various directed graphs to verify the rest of the claim.

4.2 Entropy

We wish to be able to calculate the entropy of the map $F : \mathbb{R} \to \mathbb{R}$. Since we are working in noncompact dynamics, we must specify which entropy we mean. One possibility is to compactify the dynamics by introducing fixed points at $+\infty$ and at $-\infty$. The extended map $\bar{F} : [-\infty, \infty] \to [-\infty, \infty]$ is continuous, and so has a well-defined topological entropy. Another possibility is to take the supremum of topological entropies over all compact invariant subsets. In fact, these two notions coincide and are equal to the Gurevich entropy of the corresponding symbolic system $\hat{\Sigma}$.

Theorem 4.4 *Let F be a transitive lifting of a degree one circle map with Markov partition $\mathcal{V} \times \mathbb{Z}$. Let $\hat{\Sigma}$ be the associated subshift of finite type. Then*

$$h(\bar{F}) = \sup\{h(F|_K) : K \subset \mathbb{R} \text{ compact, invariant}\} = h_{Gur}(\hat{\Sigma})$$

Proof Our proof will be terse, since the ideas are not new – similar ideas appear in [13] and [17]. Recall the characterization of Gurevich entropy as the supremum of entropies of finite subgraphs. Let $\hat{G}_n$ denote the subgraph of $\hat{G}$ formed from the vertices $\mathcal{A} \times [-n, n]$ and all arrows between these vertices. These subgraphs form an increasing sequence, and any other finite subgraph is contained in $\hat{G}_n$ for sufficiently large n. Therefore we may calculate $h_{Gur}(\hat{\Sigma})$ as the increasing limit $\lim_{n \to \infty} h(\hat{G}_n)$. Similarly, let K_n define the compact invariant set consisting of all points $X \in \mathbb{R}$ with forward orbit contained in $[-n, n + 1]$. The supremum of entropies over all compact invariant sets can be calculated along this sequence as $\lim_{n \to \infty} h(F|_{K_n})$. Now recognize that $\hat{G}_n$ encodes the symbolic dynamics of $F|_{K_n}$, and so $h(\hat{G}_n) = h(F|_{K_n})$ for all n. Therefore the limits are equal.

Next we show that $h(\bar{F})$ is equal to $\lim_{n \to \infty} h(F|_{K_n})$. The inequality $h(\bar{F}) \geq \lim_{n \to \infty} h(F|_{K_n})$ is obvious. Define "truncated" maps $F_n : [-\infty, \infty] \to [-\infty, \infty]$ by setting

$$F_n(X) = \begin{cases} n + 1, & \text{if } X \in [-n, n + 1] \text{ and } F(X) > n + 1 \\ -n, & \text{if } X \in [-n, n + 1] \text{ and } F(X) < -n \\ F(X) & \text{if } X \in [-n, n + 1] \text{ and } F(X) \in [-n, n + 1] \end{cases}$$

and then extend with $F_n|_{[n+1, \infty]}$ and $F_n|_{[-\infty, -n]}$ constant. The entropy of the truncated map F_n is at least as great as the entropy of the restricted map $F|_{K_n}$, because these two maps are identical on the compact invariant set K_n. To get the reverse inequality, notice that the dynamics of F_n are not substantially different from the dynamics of $F|_{K_n}$ in the following precise sense: each of the points $F(-n)$, and $F(n + 1)$ either belongs to K_n or is (pre)periodic, and every point in $[-\infty, \infty] \setminus K_n$ has a trajectory which eventually arrives at one of these two points. It follows from the Poincare

recurrence theorem that any invariant measure is supported on the union of K_n with perhaps one or two additional periodic orbits, and therefore by the variational principle F_n has entropy no greater than that of $F|_{K_n}$. Therefore $\lim_{n\to\infty} h(F_n) = \lim_{n\to\infty} h(F_{K_n})$.

The topological space $[-\infty, \infty]$ is homeomorphic to the interval, and, regarded as interval maps, the truncations F_n converge uniformly to $\bar{F}$. Topological entropy is a lower semicontinuous function on the space of interval maps with respect to the topology of uniform convergence [1, Theorem 4.5.2]. Therefore $h(\bar{F}) \leq \lim_{n\to\infty} h(F_n) = \lim_{n\to\infty} h(F|_{K_n})$. ∎

Theorem 4.4 allows us to apply the theory of countable extensions to compute the entropy of the transitive lifting of a degree one circle map.

Corollary 4.5 *Let F be a transitive lifting of a degree one circle map with Markov partition $\mathcal{V} \times \mathbb{Z}$, and $\bar{F}$ its continuous extension to $[-\infty, \infty]$. Let M_β^* be the associated weighted transition matrix defined in equation 4.1.3. Then $h(\bar{F}) = \log \min_{\beta \in \mathbb{R}} \mathrm{rad} M_\beta^*$.*

Proof Theorem 4.4 equates the entropy of F with the Gurevich entropy of $\hat{\Sigma}$. Theorem 3.5 equates this entropy with the minimum pressure of $\beta\varphi$. Theorem 3.8 equates the pressure of $\beta\varphi$ with the logarithm of the spectral radius of the matrix M_β. And Lemma 4.3 implies that M_β and M_β^* have the same spectral radii. ∎

We illustrate our results with several examples. We want to show what issues may arise in computations.

First, consider the map F from example 4.2. From the transition graphs, it is easy to verify the transitivity hypothesis. The matrix M_β^* has rank one, and so its spectral radius is equal to its trace. Thus, the problem of finding the entropy is reduced to minimizing $e^{-\beta} + 2 + 2e^\beta$ and taking a logarithm. Then, elementary calculus gives $h(\bar{F}) = \log(2 + 2\sqrt{2})$.

Next, let F be the piecewise-affine "connect-the-dots" map with critical points $F(k) = k - 2$, $F(k + \frac{1}{2}) = k + 2$, $k \in \mathbb{Z}$, again with the Markov partition $\mathcal{V} \times \mathbb{Z}$ where

$\mathcal{V} = \{[0, \frac{1}{2}], [\frac{1}{2}, 1]\}$. We can easily write down the weighted transition matrix M_β^*. It again has rank one, so that its spectral radius is just its trace, which we calculate to be

$$\operatorname{rad} M_\beta^* = 2e^{2\beta} + 2e^\beta + 2 + 2e^{-\beta} + 2e^{-2\beta} + e^{-3\beta}$$

To minimize this expression we set the derivative equal to zero, substitute $\mu = e^\beta$, and look for positive real solutions of the resulting quintic equation

$$4\mu^5 + 2\mu^4 - 2\mu^2 - 4\mu - 3 = 0.$$

There has to be exactly one positive real root, because $\operatorname{rad} M_\beta^*$ is the exponential of $\mathcal{P}(\beta\varphi)$ and is therefore convex with a unique minimum. Computations give this root as $\mu \approx 1.1138$ and the resulting value for the entropy is $h(\bar{F}) \approx \log(10.8403)$.

4.3   Constant Slope

Lemma 4.6 *If $M_\beta r = \lambda r$, then the nonnegative vector*

$$y\big((a, m)\big) = e^{\beta m} r(a)$$

satisfies $Ty = \lambda y$.

Proof   By hypothesis we have

$$\sum_{b \in \mathcal{A}} M_\beta(a, b) r(b) = \lambda r(a), \quad a \in \mathcal{A}.$$

Applying Equation 3.4.2 this becomes

$$\sum_{\substack{b \in \mathcal{A}: \\ init(b) = term(a)}} e^{\beta\varphi(a)} r(b) = \lambda r(a), \quad a \in \mathcal{A}.$$

Multiplying both sides of the equation by $e^{\beta m}$, $m \in \mathbb{Z}$ arbitrary, and applying the definition of y, we obtain

$$\sum_{\substack{b \in \mathcal{A}: \\ init(b) = term(a)}} y\big((b, m + \varphi(a))\big) = \lambda y\big((a, m)\big), \quad a \in \mathcal{A}, m \in \mathbb{Z}.$$

Applying Equation 4.1.5 we obtain the desired result

$$\sum_{(b,n)\in\mathcal{A}\times\mathbb{Z}} T\big((a,m),(b,n)\big)y\big((b,n)\big) = \lambda y\big((a,m)\big), \quad (a,m)\in\mathcal{A}\times\mathbb{Z}.$$

∎

Theorem 4.7 *Let F be a transitive lifting of a degree one circle map with Markov partition $\mathcal{V}\times\mathbb{Z}$, and let $\bar{F}:[-\infty,\infty]\to[-\infty,\infty]$ denote the extended map (with fixed points at $\pm\infty$). Fix $\lambda > 1$. Then $\bar{F}$ is conjugate to a map of constant slope λ on some interval $[a,b]\subseteq[-\infty,\infty]$ if and only if $\log\lambda \geq h(\bar{F})$.*

Proof Our proof applies the theory of countable extensions to the symbolic systems $\hat{\Sigma}$ and (Σ,φ) constructed from F in Section 4.1. We saw there that $\hat{\Sigma}$ is topologically transitive and Σ is topologically mixing, so that the theory of countable extensions applies in full force. We will need to use the weighted transition matrix M_β of (Σ,φ) given in Equation 3.4.2 and the infinite transition matrix T of $\hat{\Sigma}$ given in Equation 4.1.5. The theory of countable extensions will allow us to determine which positive numbers λ are eigenvalues for nonnegative eigenvectors of T.

We also appeal to our work in Chapter 2 on countably piecewise monotone and Markov maps. We choose to regard $\bar{F}$ as countably piecewise monotone and Markov with respect to the refined partition $(\mathcal{V}\times\mathbb{Z})\vee F^{-1}(\mathcal{V}\times\mathbb{Z})$. Remark 4.1 identifies this partition with $\mathcal{A}\times\mathbb{Z}$, and we see that the Markov transition matrix for $\bar{F}$ with respect to this partition is the same matrix T that encodes $\hat{\Sigma}$. The discussion at the end of Section 4.1 shows that $\bar{F}$ is topologically mixing. Applying Theorem 2.2, it suffices to prove the equivalence

$\log\lambda \geq h(\bar{F})$ iff T has a nonnegative eigenvector in $\mathbb{R}^{\mathcal{A}\times\mathbb{Z}}$ with eigenvalue λ.

Suppose that $\log\lambda \geq h(\bar{F})$. By Theorem 4.4, $\log\lambda \geq h_{Gur}(\hat{\Sigma})$. By Corollary 3.6 we can find β such that $\log\lambda = \mathcal{P}(\beta\varphi)$. By Theorem 3.8, λ is the spectral radius of the weighted transition matrix M_β. By the Perron Frobenius theorem, $M_\beta r = \lambda r$ for some strictly positive vector r. By Lemma 4.6, we can lift r to a nonnegative eigenvector $y\in\mathbb{R}^{\mathcal{A}\times\mathbb{Z}}$ for T with eigenvalue λ.

Conversely, suppose that T has a nonnegative eigenvector $y \in \mathbb{R}^{\mathcal{A} \times \mathbb{Z}}$ with eigenvalue λ. Then $T^n y = \lambda^n y$ for all $n \in \mathbb{N}$. Fix a state $I \in \mathcal{A} \times \mathbb{Z}$ such that $y(I) > 0$. By the definition of matrix multiplication and the nonnegativity of y we obtain the inequality

$$\lambda^n y(I) \geq (T^n)_{II}\, y(I), \qquad n \in \mathbb{N}.$$

Now recall the characterization of Gurevich entropy given by Equation 3.2.1. After taking the logarithm of both sides of our inequality, dividing by n, and letting n tend to infinity, it follows that $\log \lambda \geq h_{Gur}(\hat{\Sigma})$.

∎

Among all constant slope maps conjugate to $\bar{F}$, Theorem 4.7 characterizes which slopes can be realized. What can we say about the constant slope maps themselves? In light of Theorem 2.2, finding a constant slope map conjugate to $\bar{F}$ is the same as finding a nonnegative eigenvector for the matrix T acting on the linear space $\mathbb{R}^{\mathcal{A} \times \mathbb{Z}}$.

Conjecture 4.8 *Let F be a transitive lifting of a degree one circle map with Markov partition $\mathcal{V} \times \mathbb{Z}$. Let T be the infinite transition matrix of the corresponding countable extension $\hat{\Sigma}$. Let $d(\lambda)$ denote the dimension of the intersection of the nullspace of $T - \lambda I$ with the cone of nonnegative vectors in $\mathbb{R}^{\mathcal{A} \times \mathbb{Z}}$. We conjecture that*

$$d(\lambda) = \begin{cases} 0, & \text{if } 0 < \log \lambda < h(\bar{F}) \\ 1, & \text{if } \log \lambda = h(\bar{F}) \\ 2, & \text{if } \log \lambda > h(\bar{F}) \end{cases}$$

We give now partial evidence in support of this conjecture. Theorem 4.7 shows that $d(\lambda) = 0$ for $0 < \log \lambda < h(\bar{F})$ and $d(\lambda) \geq 1$ for $\log \lambda \geq h(\bar{F})$. If we read the proofs of Theorems 3.5 and 4.7 carefully, we see that $d(\lambda) \geq 2$ for $\lambda > \exp h(\bar{F})$. This is because the pressure function $\beta \mapsto \mathcal{P}(\beta \varphi)$ attains the value $\log \lambda$ for exactly two distinct values of β. This gives two weighted transition matrices M_β with the same spectral radius λ, which by Lemma 4.6 gives two linearly independent nonnegative eigenvectors for T (linear independence follows from the distinctness of β). Our

conjecture asserts that up to taking positive linear combinations, these are the only nonnegative eigenvectors for T.

Now we give some alternative computational techniques for finding the nullspace of $T - \lambda I$. Our computations give a finite upper bound for $d(\lambda)$ in terms of the cardinality of $\mathcal{A}$ and the maximum value of the displacement function φ. We make no attempt to sharpen this upper bound.

Fix λ with $\log \lambda \geq h_{Gur}(\hat{\Sigma})$. Let r denote the cardinality of $\mathcal{A}$ and m the maximum value of the displacement function φ. Form a finite submatrix S of the matrix $T - \lambda I$ taking the entries from rows $\mathcal{A} \times [l, l + 2mr]$ and columns $\mathcal{A} \times [l - m, l + 2mr + m]$; the result is independent of the choice of $l \in \mathbb{Z}$. By the definition of m, S contains all nonzero entries from rows $\mathcal{A} \times [l, l + 2mr]$ of $T - \lambda I$. Therefore if y is in the nullspace of $T - \lambda I$, then the projection of y on $\mathbb{R}^{\mathcal{A} \times [l - m, l + 2mr + m]}$ is in the nullspace of S.

Now form a square matrix R by taking rows $\mathcal{A} \times [l, l + 2mr]$ and columns $\mathcal{A} \times [l - m, l + 2mr + m]$ from T and adjoining mr rows of zeros at the top and mr rows of zeros at the bottom. It is the binary matrix corresponding to a finite subgraph of the transition graph of $\hat{\Sigma}$. Recall now the characterization of Gurevich entropy as the supremum of topological entropies over finite subgraphs. Theorem 3.7 tells us that $\hat{\Sigma}$ has no measure of maximal entropy. But the finite state subshift of the finite subgraph corresponding to R does have a measure attaining its topological entropy, namely, its Parry measure, (or else the Parry measure on some irreducible component) (see, eg., [13]). Therefore the topological entropy of this subgraph is strictly smaller than $h_{Gur}(\hat{\Sigma})$. But the topological entropy of this subgraph is also the spectral radius of R (see, eg., [1]). It follows that the spectral radius of R is smaller than λ, and therefore $R - \lambda I$ has full rank. But S is just $R - \lambda I$ with the upper and lower rm rows removed. Therefore S has full rank.

Apply Gauss-Jordan elimination to find the reduced row-echelon form of S. The number of columns without a leading one is $2mr$. It follows by the pigeonhole principle that for every $a \in \mathcal{A}$ there exists $k \in \{0, 1, \dots, 2mr\}$ such that the column with index $(a, l + k)$ contains a leading one. Thus, for any vector in the nullspace of S, we

can solve for entry $(a, l + k)$ in terms of some succeeding entries, and the number of succeeding entries required is less than N, where N is the number of columns of S. It follows that for any vector y in the nullspace of $T - \lambda I$, we can solve for entry $(a, l + k)$ as a function of the succeeding N entries. This is true for every $a \in \mathcal{A}$, and the integer l is completely arbitrary, and therefore we can solve for every entry of y in terms of the succeeding N entries. We may also apply Gauss-Jordan elimination working from the bottom right-hand corner of S to produce trailing ones instead of leading ones. It follows that we can solve for every entry of y in terms of the preceding N entries. We have shown that once we know N consecutive entries of a vector y in the nullspace of $T - \lambda I$, we can calculate all the remaining entries. Therefore the nullspace of $T - \lambda I$ in $\mathbb{R}^{\mathcal{A} \times \mathbb{Z}}$ has dimension at most N.

REFERENCES

REFERENCES

[1] Lluis Alsedà, Jaume Llibre and Michał Misiurewicz, *Combinatorial Dynamics and Entropy in Dimension One*, Second Ed., Adv. Ser. Nonlinear Dynam. 5, World Sci., Singapore, 2000.

[2] Lluis Alsedà and Michał Misiurewicz, *Semiconjugacy to a map of a constant slope*, in preparation.

[3] Louis Block and William Coppel, *Dynamics in One Dimension*, Lecture Notes in Math. 1513, Springer, Berlin, 1992.

[4] Louis Block and Ethan Coven *Topological Conjugacy and Transitivity for a Class of Piecewise Monotone Maps of the Interval*, Trans. Amer. Math. Soc. **300** (1987), 297–306.

[5] Alexander Blokh *On Sensitive Mappings of the Interval*, Russian Math. Surveys **37** (1982), 189–190.

[6] Jozef Bobok, *Semiconjugacy to a map of a constant slope*, Studia Math. **208** (2012), 213–228.

[7] Jozef Bobok, Private Communication. 2014.

[8] Rufus Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms* Lecture Notes in Mathematics 470, Springer, Berlin, 1975.

[9] Tomasz Downarowicz, *Entropy in Dynamical Systems*, Cambridge University Press, New York, 2011.

[10] Claude-Alain Faure, *A Short Proof of Lebesgue's Density Theorem*, Amer. Math. Monthly **109** (2002), 194–196.

[11] Boris Gurevich, *Topological Entropy of Enumerable markov Chains*, (Russian), Dokl. Akad. Nauk SSSR, **187** (1969), 715-718; English translation: Soviet Math. Dokl., **10(4)** (1969), 911-915.

[12] Boris Gurevich, *Shift entropy and Markov Measures in the Space of Paths of a Countable Graph* (Russian), Dokl. Akad. Nauk SSSR, **192** (1970), 963-965; English translation: Russian Math. Surveys, **53(2)** (1998), 245-344.

[13] Bruce Kitchens, *Symbolic Dynamics*, Springer-Verlag, Berlin, 1998.

[14] John Milnor and William Thurston, *On iterated maps of the interval*, Dynamical Systems, Lecture Notes in Math. 1342, Springer, Berlin, 1988, 465–563.

[15] Michał Misiurewicz and Samuel Roth, *No Semiconjugacy to a Map of Constant Slope*, Ergodic Theory and Dynamical Systems, Available on CJO2014. doi:10.1017/etds.2014.81. Print copy to appear.

[16] Michał Misiurewicz and Wiesław Szlenk, *Entropy of piecewise monotone mappings*, Studia Math. **67** (1980), 45–63.

[17] Michał Misiurewicz and Juan Tolosa, *Entropy of Snakes and the Restricted Variational Principle*, Ergod. Th. & Dynam. Sys. **12** (1992), 791–802.

[18] William Parry, *Symbolic dynamics and transformations of the unit interval*, Trans. Amer. Math. Soc. **122** (1966), 368–378.

[19] William Parry, *Entropy and Generators in Ergodic Theory*, W. A. Benjamin, Inc., New York, 1969.

[20] William Parry and Selim Tuncel, *Classification Problems in Ergodic Theory*, Cambridge University Press, Cambridge, 1982.

[21] David Ruelle, *Thermodynamic Formalism*, Second Edition, Cambridge University Press, Cambridge, 2004.

[22] Stephen Silverman *On Maps with Dense Orbits and the Definition of Chaos*, Rocky Mountain J. Math. **22** (1992), 353-375.

VITA

VITA

Upon receiving his Ph.D. from Purdue University, Samuel Roth will move to the Czech Republic to join the love of his life.