RANDOMNESS AND NON-ERGODIC SYSTEMS

JOHANNA N.Y. FRANKLIN AND HENRY TOWSNER

ABSTRACT. We characterize the points that satisfy Birkhoff's ergodic theorem under certain computability conditions in terms of algorithmic randomness. First, we use the method of cutting and stacking to show that if an element x of the Cantor space is not Martin-Löf random, there is a computable measure-preserving transformation and a computable set that witness that x is not typical with respect to the ergodic theorem, which gives us the converse of a theorem by V'yugin. We further show that if x is weakly 2-random, then it satisfies the ergodic theorem for all computable measure-preserving transformations and all lower semi-computable functions.

1. INTRODUCTION

Random points are typical with respect to measure in that they have no measure-theoretically rare properties of a certain kind, while ergodic theorems describe regular measure-theoretic behavior. There has been a great deal of interest in the connection between these two kinds of regularity recently. We begin by defining the basic concepts in each field and then describe the ways in which they are related. Then we present our results on the relationship between algorithmic randomness and the satisfaction of Birkhoff's ergodic theorem for computable measure-preserving transformations with respect to computable (and then lower semi-computable) functions. Those more familiar with ergodic theory than computability theory might find it useful to first read Section 7, a brief discussion of the notion of algorithmic randomness in the context of ergodic theory.

1.1. **Algorithmic randomness in computable probability spaces.**

Computability theorists seek to calibrate the computational strength of subsets of ω, the nonnegative integers. This calibration is accomplished using the notion of a Turing machine, which can be informally viewed as an idealized computer program (for a general introduction to computability theory, see [22, 23, 28]). We present the main concepts we will need here.

A subset A of the natural numbers ω is *computably enumerable*, or c.e., if it is the domain of some Turing machine P, that is, the set of numbers

Date: September 4, 2013.

2000 *Mathematics Subject Classification.* Primary 03D32; Secondary 37A25.

Key words and phrases. Algorithmic randomness, Martin-Lof random, dynamical system, ergodic theorem, upcrossing.

Partially supported by NSF grant DMS-1340666.

that, when input into P, will result in the program halting and returning an output. One can understand the origin of this terminology intuitively: a set is c.e. if we can generate it by running some Turing machine on more and more inputs and enumerating inputs into our set if the machine returns answers for them. Note that this means that we can list the c.e. sets: since we can enumerate the computer programs $\langle P_i \rangle$, we can enumerate their domains. A set A is *computable* if both it and its complement are c.e., and we say that a function $f : \omega \to \omega$ is computable if there is a Turing machine P whose domain is ω such that for all n, $f(n) = P(n)$. These concepts lead us to the last definition we will need: that of an *effectively c.e. sequence*. A sequence of c.e. sets is said to be effectively c.e. if there is a computable function f such that the n^{th} set in the sequence is the $f(n)^{th}$ c.e. set. We note without ceremony that we can consider c.e. and computable sets of objects other than natural numbers. For instance, there is a computable bijection between ω and the set of finite binary strings $2^{<\omega}$, and we will often speak of an c.e. subset of $2^{<\omega}$.

Subsets of ω are often identified in a natural way with infinite binary sequences, or *reals*: a set A corresponds to the infinite binary sequence whose $(n+1)^{st}$ bit is 1 if and only if n is in A. This is the approach that is most often taken in algorithmic randomness, since some definitions of randomness, such as the initial-segment complexity and the betting strategy definitions, are more naturally phrased in terms of infinite binary sequences than subsets of ω.

We recall the standard notations for sequences that will be used in this paper. We write 2^ω for the set of infinite binary sequences, that is, the set of functions from ω to $\{0, 1\}$. As mentioned above, we write $2^{<\omega}$ for the set of finite binary sequences, that is, functions from $[0, n)$ to $\{0, 1\}$ for some n. We sometimes write finite sequences in the form $\sigma = \langle s_0, \ldots, s_{n-1} \rangle$, where σ is the sequence with $\sigma(i) = s_i$ for all $i < n$.

If x is a finite or infinite sequence then $|x| \in \omega \cup \{\infty\}$ is the length of the sequence, and if $n \leq |x|$ then $x \upharpoonright n$ is the initial segment of x of length n (that is, the restriction of x, as a function, to the domain $[0, n)$). We write $x \sqsubseteq y$ if $|x| \leq |y|$ and $y \upharpoonright |x| = x$ (that is, if x is an initial segment of y) and $x \sqsubset y$ if $x \sqsubseteq y$ and $|x| < |y|$ (that is, if x is a proper initial segment of y). When σ is a finite sequence, $\sigma^\frown y$ is the concatenation of σ with y—that is, $\sigma^\frown y$ is the sequence with $(\sigma^\frown y)(i) = \sigma(i)$ for $i < |\sigma|$ and $(\sigma^\frown y)(i) = y(i - |\sigma|)$ for $i \geq |\sigma|$. If $\sigma \in 2^{<\omega}$ then $[\sigma] \subseteq 2^\omega$ is $\{x \in 2^\omega \mid \sigma \sqsubset x\}$, the set of infinite sequences extending σ, and if $V \subseteq 2^{<\omega}$ then $[V] = \bigcup_{\sigma \in V} [\sigma]$. We call $[\sigma]$ an *interval*. We say V is *prefix-free* if whenever $\sigma, \tau \in V$, $\sigma \sqsubseteq \tau$ implies $\sigma = \tau$.

We will usually use Greek letters such as $\sigma, \tau, \upsilon, \rho, \eta, \zeta, \nu, \theta$ for finite sequences and Roman letters such as x, y for infinite sequences.

For a general reference on algorithmic randomness, see [8, 9, 21]. We will confine our attention to the Cantor space 2^ω with the Lebesgue measure λ. In light of Hoyrup and Rojas' theorem that any computable probability space is isomorphic to the Cantor space in both the computable and measure-theoretic senses [15], there is no loss of generality in restricting to this case.

We can now present Martin-Löf's original definition of randomness [20].

Definition 1.1. An effectively c.e. sequence $\langle V_i \rangle$ of subsets of $2^{<\omega}$ is a *Martin-Löf test* if $\lambda([V_i]) \leq 2^{-i}$ for every i. If $x \in 2^\omega$, we say that x is *Martin-Löf random* if for every Martin-Löf test $\langle V_i \rangle$, $x \notin \cap_i [V_i]$.

It is easy to see that $\lambda(\cap_i [V_i]) = 0$ for any Martin-Löf test, and since there are only countably many Martin-Löf tests, almost every point is Martin-Löf random.

In Section 6, we will also consider weakly 2-random elements of the Cantor space. Weak 2-randomness is a strictly stronger notion than Martin-Löf randomness and is part of the hierarchy introduced by Kurtz in [19].

Definition 1.2. An effectively c.e. sequence $\langle V_i \rangle$ of subsets of $2^{<\omega}$ is a *generalized Martin-Löf test* if $\lim_{n\to\infty} \lambda([V_i]) = 0$. If $x \in 2^\omega$, we say that x is *weakly 2-random* if for every generalized Martin-Löf test $\langle V_i \rangle$, $x \notin \cap_i [V_i]$.

1.2. **Ergodic theory.** Now we discuss ergodic theory in the general context of an arbitrary probability space before transferring it to the context of a computable probability space. The following definitions can be found in [14].

Definition 1.3. Suppose (X, μ) is a probability space, and let $T : X \to X$ be a measurable transformation.

1) T is *measure preserving* if for all measurable $A \subseteq X$, $\mu(T^{-1}(A)) = \mu(A)$.
2) A measurable set $A \subseteq X$ is *invariant* under T if $T^{-1}(A) = A$ modulo a set of measure 0.
3) T is *ergodic* if it is measure preserving and every T-invariant measurable subset of X has measure 0 or measure 1.

One of the most fundamental theorems in ergodic theory is Birkhoff's Ergodic Theorem:

Birkhoff's Ergodic Theorem. *[4] Suppose that (X, μ) is a probability space and $T : X \to X$ is measure preserving. Then for any $f \in L_1(X)$ and almost every $x \in X$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i<n} f(T^i(x))$$

converges. Furthermore, if T is ergodic then for almost every x this limit is equal to $\int f \, d\mu$.

If we restrict ourselves to a countable collection of functions, this theorem gives a natural notion of randomness—a point is random if it satisfies the conclusion of the ergodic theorem for all functions in that collection. In a computable measure space, we can take the collection of sets defined by a computability-theoretic property and attempt to classify this notion in terms of algorithmic randomness. In particular, we are interested in the following property:

Definition 1.4. Let (X, μ) be a computable probability space, and let $T : X \to X$ be a measure-preserving transformation. Let $\mathcal{F}$ be a collection of functions in $L_1(X)$. A point $x \in X$ is a *weak Birkhoff point* for T with respect to $\mathcal{F}$ if for every $f \in \mathcal{F}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i<n} f(T^i(x))$$

converges. x is a *Birkhoff point* for T with respect to $\mathcal{F}$ if additionally

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i<n} f(T^i(x)) = \int f \, d\mu.$$

The definition of a Birkhoff point is only appropriate when T is ergodic; when T is nonergodic, the appropriate notion is that of a weak Birkhoff point.

There are two natural dimensions to consider: the ergodic-theoretic behavior of T and the algorithmic complexity of $\mathcal{C}$. The case where T is ergodic has been largely settled.

A point is Martin-Löf random if and only if the point is Birkhoff for all computable ergodic transformations with respect to lower semi-computable functions [2, 11]. The proof goes by way of a second theorem of ergodic theory:

Poincaré Recurrence Theorem ([24], Chapter 26)**.** *Suppose that (X, μ) is a probability space and $T : X \to X$ is measure preserving. Then for all $E \subseteq X$ of positive measure and for almost all $x \in X$, $T^n(x) \in E$ for infinitely many n.*

In short, the Poincaré Recurrence Theorem says that an ergodic transformation T returns almost every point to every set of positive measure repeatedly, and Birkhoff's Ergodic Theorem says that it will do so with a well-defined frequency in the limit.

A point $x \in X$ is a *Poincaré point* for T with respect to $\mathcal{C}$ if for every $E \in \mathcal{C}$ with positive measure, $T^n(x) \in E$ for infinitely many n. In [18], Kučera proved that a point in the Cantor space is Martin-Löf random if and only if it is a Poincaré point for the shift operator with respect to effectively closed sets. Later, Bienvenu, Day, Mezhirov, and Shen generalized this result and showed that in any computable probability space, a point is Martin-Löf random if and only if it is a Poincaré point for computable ergodic transformations with respect to effectively closed sets [2]. The proof that Martin-Löf random points are Poincaré proceeds by showing that a point which is Poincaré for any computable ergodic transformation with respect to effectively closed sets must also be a Birkhoff point for computable ergodic transformations with respect to lower semi-computable functions [2, 11].

	Transformations	
Sets	Ergodic	Nonergodic
Computable	Schnorr [13]	Martin-Löf [30]+Theorem 4.4
Lower semi-computable	Martin-Löf [2, 11]	?

TABLE 1. Randomness notions and ergodicity

Similarly, Gács, Hoyrup, and Rojas have shown that if a point fails to be *Schnorr random* then there is a computable ergodic transformation where the point fails to be Birkhoff for a bounded computable function [13]. (We say that x is Schnorr random if $x \notin \cap_i [V_i]$ for all Martin-Löf tests $\langle V_i \rangle$ where $\lambda([V_i]) = 2^{-i}$ for all i; Schnorr randomness is a strictly weaker notion than Martin-Löf randomness [26].) In fact, the transformation they construct has a stronger property—it is *weakly mixing*—and they show, conversely, that in a computable weakly mixing transformation every Schnorr random point is Birkhoff. However Rojas has pointed out [25] that this latter result can be strengthened: even in a computable ergodic transformation, every Schnorr random point is Birkhoff for every bounded computable function. Since this last result has not appeared in print, we include it in Section 5 for completeness. Combining these results, we see that a point is *Schnorr random* if and only if the point is Birkhoff for all computable ergodic transformations with respect to computable functions [13].

In this paper, we consider the analogous situations when T is nonergodic. V'yugin [30] has shown that if $x \in 2^\omega$ is Martin-Löf random then x is weakly Birkhoff for any (not necessarily ergodic) computable measure-preserving transformation T with respect to computable functions. Our main result is the converse: that if x is not Martin-Löf random then x is not weakly Birkhoff for some particular transformation T with respect to computable functions (in fact, with respect to computable sets).

These results are summarized in Table 1.

This says that a point is weakly Birkhoff for the specified family of computable transformations with respect to the specified collection of functions if and only if it is random in the sense found in the corresponding cell of the table.

We also begin an analysis of the remaining space in the table; we give an analog of V'yugin's result, showing that if x is weakly 2-random then x is a weak Birkhoff point for all computable measure-preserving transformations with respect to lower semi-computable functions.

The next two sections will be dedicated to a discussion of the techniques we will use in our construction. Section 2 contains a description of the type of partial transformations we will use to construct the transformation T mentioned above, and Section 3 discusses our methods for building new partial transformations that extend other such transformations. We combine the material from these two sections to prove our main theorem in Section 4, while Section 6 contains a further extension of our work and some speculative material on a more relaxed form of upcrossings. Section 7 is a general discussion of algorithmic randomness intended for ergodic theorists.

2. Definitions and Diagrams

We will build computable transformations $\widehat{T} : 2^\omega \to 2^\omega$ using computable functions $T : 2^{<\omega} \to 2^{<\omega}$ such that (1) $\sigma \sqsubseteq \tau$ implies $T(\sigma) \sqsubseteq T(\tau)$ and (2) $\widehat{T}(x) = \lim_{n\to\infty} T(x \upharpoonright n)$ is defined and infinite for all $x \in 2^\omega$ outside an F_σ set with measure 0.

We will approximate such a $\widehat{T}$ by *partial transformations*:

Definition 2.1. A *partial transformation* is a total computable function $T : 2^{<\omega} \to 2^{<\omega}$ such that if $\sigma \sqsubseteq \tau$ then $T(\sigma) \sqsubseteq T(\tau)$. We write $T \sqsubseteq T'$ if for all σ, $T(\sigma) \sqsubseteq T'(\sigma)$. If $T_0 \sqsubseteq T_1 \sqsubseteq \cdots \sqsubseteq T_n \sqsubseteq \cdots$ is a sequence of partial transformations, there is a natural limit $T : 2^{<\omega} \to 2^{<\omega}$ given by $T(\sigma) = \lim_n T_n(\sigma)$.

We say a computable transformation $\widehat{T} : 2^\omega \to 2^\omega$ *extends* T if for every σ, $\widehat{T}([\sigma]) \subseteq [T(\sigma)]$.

The "partial" refers to the fact that we may have $\lim_{n\to\infty} T(x \upharpoonright n)$ be finite for many or all points. If $T_0 \sqsubseteq \cdots$ is a uniformly computable sequence of partial transformations with $T = \lim_n T_n$ and for almost every x the limit $\lim_n |T(x \upharpoonright n)| = \infty$, then the transformation $\widehat{T}(x) = \lim_n T(x \upharpoonright n)$ is a computable transformation.

We will be exclusively interested in partial transformations which are described finitely in a very specific way:

Definition 2.2. A partial transformation T is *proper* if there are finite sets T_-, T_+ such that:

- $T_- \cup T_+$ is prefix-free,
- $\cup_{\sigma\in T_-\cup T_+}[\sigma] = 2^\omega$,
- If there is a $\tau \sqsubseteq \sigma$ such that $\tau \in T_-$ then $T(\sigma) = T(\tau)$,
- If $\sigma = \tau^\frown\rho$ with $\tau \in T_+$ then $T(\sigma) = T(\tau)^\frown\rho$,
- If $\sigma \in T_-$ then $|T(\sigma)| < |\sigma|$,
- If $\sigma \in T_+$ then $|T(\sigma)| = |\sigma|$,
- If $\sigma \in T_+$ and $\sigma \neq \tau \in T_+ \cup T_-$ then $T(\tau) \not\sqsupseteq T(\sigma)$,
- If $\sigma \in T_-$ and $\tau \in T_+ \cup T_-$ then $T(\tau) \not\sqsupset T(\sigma)$.

We say σ is *determined* in T if there is some $\tau \sqsubseteq \sigma$ with $\tau \in T_- \cup T_+$.

When T is a proper partial transformation, we write T_-, T_+ for some canonically chosen pair of sets witnessing this fact.

In practice, we will always describe a proper transformation T by describing $T \upharpoonright T_- \cup T_+$ for some particular choice of T_-, T_+, so there is always a canonical choice of T_- and T_+. Note that if σ is determined then $T(\sigma)$ is

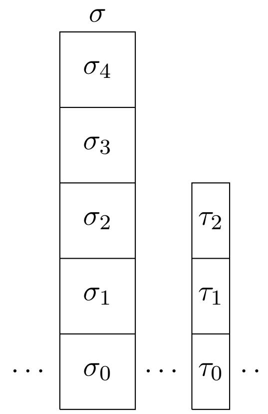

FIGURE 1. A typical diagram

uniquely defined by the values of T on the finitely many subsequences in $T_- \cup T_+$. The roles of T_- and T_+ will be clearer when we introduce a diagrammatic notion for describing transformations. For now, note that once we have $\tau \in T_+$, we have entirely determined $\widehat{T} \upharpoonright [\tau]$ for any $\widehat{T}$ extending T: $\widehat{T}(\tau^\frown x) = T(\tau)^\frown x$. The requirement that, for such a τ, $|\tau| = |T(\tau)|$ helps ensure that the resulting transformation is measure preserving.

Throughout this paper, all our partial transformations will be proper.

We will use the method of cutting and stacking, which was introduced by Chacon to produce dynamical systems with specific combinatorial properties [6, 7][1]. This method was recently introduced into the study of the computability properties of ergodic theorems by V'yugin [29]. One tries to construct a dynamical system, usually on the real interval $[0, 1]$, by specifying the transformation in stages. At a given stage, the interval has been "cut" into a finite number of components, some of which have been "stacked" into "towers" or "ladders." A tower is read upwards, so the interval on the bottom level is mapped by the transformation to the level above, and from that level to the level above that. On the top level of a tower, the transformation is not yet defined. To produce the next stage, the towers are cut into smaller towers and further stacked. By manipulating the order in which the components are stacked, specific properties of the transformation can be enforced. This method has been extensively used in ergodic theory and probability theory to construct examples with specific properties (some overviews of the area are [12, 17, 27]).

A typical diagram is shown in Figure 1. This figure represents that $|\sigma| < |\sigma_0| = |\sigma_1| = |\sigma_2| = |\sigma_3| = |\sigma_4|$ and that for all υ, $T(\sigma_i{}^\frown\upsilon) = \sigma_{i+1}{}^\frown\upsilon$ for $i < 4$, $T(\sigma_4{}^\frown\upsilon) = \sigma$, and similarly $T(\tau_i{}^\frown\upsilon) = T(\tau_{i+1}{}^\frown\upsilon)$ for $i < 2$ while $T(\tau_2) = \langle\rangle$. Although it is not essential to interpret the diagrams, we will

[1]Actually, according to [12], the method was first used several decades earlier by von Neumann and Kakutani, but not published until later [16].

try to be consistent about the scale of blocks; in Figure 1, the relative width of the blocks suggests that $|\tau_i| = |\sigma_i| + 1$—that is, $\lambda([\tau_i]) = \lambda([\sigma_i])/2$; the height of a block does not represent anything, so we draw each block with the same height. The only relevant dimensions are the widths of the blocks and the numbers of blocks in the towers.

In general, a block represents a subset of 2^ω of the form $[\sigma]$ for some sequence σ; by placing the block corresponding to $[\sigma]$ on top of the block corresponding to $[\tau]$, we are indicating that $\tau \in T_+$ and $T(\tau) = \sigma$—that is, in the transformation we construct extending T, $T([\tau]) = [\sigma]$. (We must, therefore, have $|\sigma| = |\tau|$.) By placing some sequence σ' with $|\sigma'| < |\sigma|$ on top of the block corresponding to $[\sigma]$, we are indicating that $\sigma \in T_-$ and $T(\sigma) = \sigma'$—that is, in the transformation $\widehat{T}$ we construct extending T, $\widehat{T}([\sigma]) \sqsubseteq [\sigma']$.

The roles of T_- and T_+ in the specification of a proper transformation are now clearer: the elements of $T_- \cup T_+$ are the particular blocks labeled in a given diagram; the elements $\tau \in T_+$ are those blocks which have another block on top, and therefore we have already defined the value of any extension $\widehat{T}$ of T on every element of $[\tau]$. The elements $\tau \in T_-$ are topmost blocks of some tower, for which we have (at most) partial information about the ultimate behavior of $\widehat{T}$ on $[\tau]$.

Definition 2.3. We say τ is *blocked* if there is any σ such that $T(\sigma) \sqsupseteq \tau$. Otherwise we say τ is *unblocked*.

An *open loop* in a partial transformation T is a sequence $\sigma_0, \dots, \sigma_n$ such that:

- $|\sigma_0| = |\sigma_1| = \cdots = |\sigma_n|$,
- $T(\sigma_i) = \sigma_{i+1}$ for $i < n$,
- $T(\sigma_n) \sqsubset \sigma_0$,
- σ_0 is unblocked.

We refer to σ_0 as the *initial element* of the open loop $\sigma_0, \dots, \sigma_n$ and σ_n as the *final element*.

The *width* of an open loop is the value $2^{-|\sigma_i|}$.

We say T is *partitioned into open loops* if for every determined σ there is an open loop $\sigma_0, \dots, \sigma_n$ in T with $\sigma = \sigma_i$ for some i. (In a proper transformation such an open loop must be unique.) In such a transformation we write $\mathcal{L}_T(\sigma)$ for the open loop $\sigma_0, \dots, \sigma_n$ such that for some i, $\sigma = \sigma_i$. We write $L_T(\sigma)$ for $n + 1$, the length of the open loop containing σ.

(We are interested in open loops to preclude the possibility that $T(\sigma_n) = \sigma_0$, since we are not interested in—indeed, will not allow the existence of—"closed" loops.) Diagrammatically, the requirement that T be partitioned into open loops is represented by requiring that any sequence written above a tower of blocks is a subsequence of the sequence at the bottom of that tower. (For instance, in Figure 1, we require that $\sigma \sqsubset \sigma_0$.)

One of the benefits of proper transformations partioned into open loops is that they ensure that our transformation is measure preserving:

Lemma 2.4. *Let $T_0 \sqsubseteq \ldots \sqsubseteq T_n \sqsubseteq \ldots$ be a sequence of proper partial transformations partitioned into open loops, let $T = \lim_n T_n$, and let $\widehat{T}(x) = \lim_{n\to\infty} T(x \upharpoonright n)$. Suppose that $|\widehat{T}(x)|$ is infinite outside a set of measure 0. Then $\widehat{T}$ is measure preserving.*

Proof. It suffices to show that for every σ and every $\epsilon > 0$,

$$\left|\lambda(\widehat{T}^{-1}([\sigma])) - \lambda([\sigma])\right| < \epsilon.$$

Fix σ and $\epsilon > 0$ and choose n large enough that if U is the set of $\tau \in T_{n,-}$ with $|T_n(\tau)| < |\sigma|$, $\sum_{\tau\in U} \lambda([\tau]) < \epsilon$.

If there is a $\tau \in T_{n,+}$ with $T_n(\tau) \sqsubseteq \sigma$ then there is a ρ with $|\tau^\frown\rho| = |\sigma|$ and $T_n(\tau^\frown\rho) = \sigma$ and therefore $\widehat{T}([\tau^\frown\rho]) = [\sigma]$. Since for $m \geq n$ and $\upsilon \in T_{m,+} \cup T_{m,-}$ with $\upsilon \not\sqsupseteq \tau$ we have $T_m(\upsilon) \not\sqsupseteq T_n(\tau)$, we therefore have $\widehat{T}^{-1}([\sigma]) = [\tau^\frown\rho]$ and are done.

Otherwise let $\sigma_0, \ldots, \sigma_k$ be such that each $\sigma_i \in T_{n,+} \cup T_{n,-}$ and $\bigcup_{i\leq k}[\sigma_i] = [\sigma]$. Each σ_i belongs to an open loop. Let I be the set of $i \leq k$ such that σ_i is the initial element of $\mathcal{L}_{T_n}(\sigma_i)$. If $i \notin I$ then there is a τ_i with $|\tau_i| = |\sigma_i|$ and $T_n(\tau_i) = \sigma_i$, and therefore $\widehat{T}^{-1}([\sigma_i]) = [\tau_i]$.

If $i \in I$, let υ_i be the final element of $\mathcal{L}_{T_n}(\sigma_i)$, so $|\upsilon_i| = |\sigma_i|$ but $T_n(\upsilon_i) \sqsubset \sigma_i$. Let $I' \subseteq I$ be those i such that $|T_n(\upsilon_i)| < |\sigma|$; note that if $i \in I \setminus I'$ then $\sigma \sqsubseteq T_n(\upsilon_i) \sqsubset \sigma_i$, so $[\upsilon_i] \subseteq \widehat{T}^{-1}([\sigma])$. Also, $\lambda(\bigcup_{i\in I'}[\upsilon_i]) < \epsilon$ and

$$\widehat{T}^{-1}([\sigma]) \supseteq \bigcup_{i\notin I}[\tau_i] \cup \bigcup_{i\in I\setminus I'}[\upsilon_i]$$

and so

$$\begin{aligned}
\lambda(\widehat{T}^{-1}([\sigma])) &\geq \sum_{i\notin I}\lambda([\tau_i]) + \sum_{i\in I\setminus I'}\lambda([\upsilon_i]) \\
&= \sum_{i\notin I}\lambda([\tau_i]) + \sum_{i\in I}\lambda([\upsilon_i]) - \sum_{i\in I'}\lambda([\upsilon_i]) \\
&= \sum_{i\notin I}\lambda([\sigma_i]) + \sum_{i\in I}\lambda([\sigma_i]) - \sum_{i\in I'}\lambda([\upsilon_i]) \\
&= \sum_{i\leq k}\lambda([\sigma_i]) - \sum_{i\in I'}\lambda([\upsilon_i]) \\
&> \lambda([\sigma]) - \epsilon.
\end{aligned}$$

Now consider any $\tau \in T_{n,-}$. If $|T_n(\tau)| \geq |\sigma|$ then we have either $T_n(\tau) \sqsupseteq \sigma$, in which case $\tau \in \{\tau_i\} \cup \{\upsilon_i\}$, or $T_n(\tau) \not\sqsupseteq \sigma$, in which case $T^{-1}([\sigma]) \cap [\tau] = \emptyset$. Therefore we have

$$\widehat{T}^{-1}([\sigma]) \subseteq \bigcup_{i\notin I}[\tau_i] \cup \bigcup_{i\in I\setminus I'}[\upsilon_i] \cup \bigcup_{\tau\in U}[\tau]$$

and so

$$\lambda(\widehat{T}^{-1}([\sigma])) \leq \sum_{i\leq k}\lambda([\sigma_i]) + \sum_{\tau\in U}\lambda([\tau])$$
$$< \lambda([\sigma]) + \epsilon,$$

completing the proof. □

A similar argument shows that also $\lambda(\widehat{T}(A)) = \lambda(A)$ for all A, but we do not need this fact.

Definition 2.5. When $A \subseteq 2^\omega$ and $\sigma \in 2^{<\omega}$, we write $\sigma \in A$ (σ is in A) if $[\sigma] \subseteq A$. Similarly, we say a sequence $\sigma_0, \dots, \sigma_k$ is in A if for each $i \leq k$, σ_i is in A.

We say σ *avoids* A if $[\sigma] \cap A = \emptyset$. Similarly we say a sequence $\sigma_0, \dots, \sigma_n$ avoids A if for each $i \leq k$, σ_i avoids A.

So σ is in A iff σ avoids $2^\omega \setminus A$.

Definition 2.6. An *escape sequence* for σ_0 in T is a sequence $\sigma_1, \dots, \sigma_n$ such that:

- $|\sigma_1| = |\sigma_2| = \dots = |\sigma_n|$,
- For all $0 \leq i < n$, $\sigma_{i+1} \sqsupseteq T(\sigma_i)$,
- If σ_{i+1} is blocked then $\sigma_{i+1} = T(\sigma_i)$,
- $T(\sigma_n) = \langle\rangle$,
- All σ_i are determined.

We say T is *escapable* if for every determined σ with $|T(\sigma)| < |\sigma|$, there is an escape sequence for σ. If $A, B \subseteq 2^\omega$, we say T is A, B*-escapable* if for every determined σ in A with $|T(\sigma)| < |\sigma|$, there is an escape sequence for σ in B.

An escape sequence for σ_0 is *reduced* if (1) $\sigma_i \sqsupseteq T(\sigma_j)$ implies that either $i \leq j+1$ or σ_i is blocked, and (2) if $i < n$, then $T(\sigma_i) \neq \langle\rangle$.

This is the first of many places where we restrict consideration to determined σ with $|T(\sigma)| < |\sigma|$. Note that this is the same as restricting to those σ such that there is some $\tau \sqsubseteq \sigma$ with $\tau \in T_-$. When $T(\sigma_0) = \langle\rangle$, the empty sequence is a valid escape sequence for σ_0 (and the unique reduced escape sequence).

Escapability preserves the option of extending T in such a way that we can eventually map $[\sigma_0]$ to anything not already in the image of another sequence (although it may require many applications of T). A, B-escapability will be useful at intermediate steps of our construction; typically we want to know that we have A, B-escapability so that we can manipulate portions outside of B with interfering with escapability. $\emptyset, \emptyset$-escapable is the same as escapable.

Lemma 2.7. *Every escape sequence for σ contains a reduced subsequence for σ.*

Proof. We proceed by induction on the length of the sequence. It suffices to show that if $\sigma_1, \dots, \sigma_n$ is a nonreduced escape sequence then there is a proper subsequence which is also an escape sequence for σ_0. If for some $i > j+1$, $\sigma_i \sqsupseteq T(\sigma_j)$ with σ_i unblocked, then $\sigma_1, \dots, \sigma_j, \sigma_i, \dots, \sigma_n$ is also an escape sequence. If for some $i < n$, $T(\sigma_i) = \langle\rangle$ then $\sigma_1, \dots, \sigma_i$ is also an escape sequence. □

Clearly if a sequence is in B, any subsequence is as well.

The following lemma is immediate from the definition of an escape sequence.

Lemma 2.8. *If $\sigma_1, \dots, \sigma_n$ is an escape sequence for σ_0 in T then for every ρ, $\sigma_1 {}^\frown \rho, \dots, \sigma_n {}^\frown \rho$ is also an escape sequence.*

Note that if $\sigma_1, \dots, \sigma_n$ is in some set B, so is $\sigma_1 {}^\frown \rho, \dots, \sigma_n {}^\frown \rho$.

If $\tau_1, \dots, \tau_n$ is an escape sequence, it is possible to choose an extension $\widehat{T}$ of T and $x \in [\tau_0]$ so that $T^i(x) \in [\tau_i]$ for all $i \leq n$. If $\sigma_0, \dots, \sigma_k$ is an open loop in T then in *every* extension $\widehat{T}$ of T, whenever $x \in [\sigma_0]$, we have $T^i(x) \in [\sigma_i]$ for $i \leq k$. Moreover, because $\widehat{T}$ is measure preserving, if $y \in [\sigma_{i+1}]$ then $T^{-1}(y) \in [\sigma_i]$. The next lemma shows that these properties interact—an escape sequence can only "enter" an open loop at the beginning, and if this happens, the escape sequence must then traverse the whole open loop in order.

Lemma 2.9. *Suppose $\sigma_0, \dots, \sigma_k$ is an open loop in T consisting of determined elements, $\tau_1, \dots, \tau_n$ is a reduced escape sequence for τ_0, and $|T(\tau_0)| < |\tau_0|$. Then one of the following occurs:*

- *$\tau_0 \sqsupseteq \sigma_k$ and for $j > 0$, $\tau_j \notin \cup_{i \leq k} [\sigma_i]$,*
- *There is a unique $j > 0$ such that for all $i \leq k$, $\tau_{j+i} \sqsupseteq \sigma_i$,*
- *For all j, $\tau_j \notin \cup_{i \leq k} [\sigma_i]$.*

Proof. First, suppose some $\tau_j \sqsupseteq \sigma_i$. Let j be least such that this is the case. If $j = 0$ then since $|T(\tau_0)| < |\tau_0|$, we must have $i = k$.

Suppose $j \neq 0$. If $i \neq 0$ then since σ_i is blocked, we must have $\tau_{j-1} \sqsubseteq \sigma_{i-1}$, which is impossible by our choice of j. So there is a $j > 0$ with $\tau_j \sqsupseteq \sigma_0$. Since each σ_i satisfies $T(\sigma_i) = \sigma_{i+1}$ with $|T(\sigma_i)| = |\sigma_{i+1}|$ and T is proper, for each $i \leq k$, we must have $\tau_{j+i} \sqsupseteq \sigma_i$.

So we have shown that if $\tau_j \sqsupseteq \sigma_i$ for some i, j with $j > 0$ then we have a complete copy of the open loop in our escape sequence. We now show that if $j < j'$ and $\tau_j \sqsupseteq \sigma_k$ then we cannot have $\tau_{j'} \sqsupseteq \sigma_i$; this shows both the second half of the first case and the uniqueness in the second case. For suppose we had $\tau_j \sqsupseteq \sigma_k$, $j' > j$, and $\tau_{j'} \sqsupseteq \sigma_i$. By the previous paragraph, we may assume $i = k$. But since σ_k is determined and $|T(\sigma_k)| < \sigma_k$, we have $T(\tau_j) = T(\sigma_k) = T(\tau_{j'})$. This means $T(\tau_j) \sqsubseteq \tau_{j'+1}$.

Note that $\tau_{j'+1}$ cannot be blocked: we have $|T(\tau_{j'})| = |T(\sigma_k)| < |\sigma_k| \leq |\tau_{j'}| = |\tau_{j'+1}|$ since $j' > j \geq 0$, so $T(\tau_{j'}) \neq \tau_{j'+1}$. So we must have $j' + 1 \leq j + 1$, contradicting the assumption that $j < j'$. □

3. Working with Transformations

We will work exclusively with partial transformations with a certain list of properties which, for purposes of this paper, we call "useful" partial transformations. In this section we describe some basic operations which can be used to manipulate useful transformations. While these operations are ultimately motivated by the construction in the next section, they also provide some intuition for why useful transformations deserve their name.

Definition 3.1. A partial transformation T is *useful* if:

- T is proper,
- T is partitioned into open loops, and
- T is escapable.

The next lemma illustrates one of the advantages of always working with open loops: we can always modify a useful partial transformation by replacing a open loop with a new open loop of the same total measure but arbitrarily small width.

Lemma 3.2 (Thinning Loops)**.** *Let T be a useful partial transformation, let $\sigma_0, \dots, \sigma_k$ be an open loop of determined elements and let $\epsilon = 2^{-n}$ be smaller than the width of this open loop. Then there is a useful $T' \sqsupseteq T$ such that*

> *There is an open loop $\tau_0, \dots, \tau_{k'}$ in T' of width ϵ such that $\cup_{j \leq k'}[\tau_i] = \cup_{i \leq k}[\sigma_i]$.*

Furthermore,

- *If $\tau \notin \cup_{i \leq k'}[\tau_i]$ is determined in T' then $T'(\tau) = T(\tau)$ and $L_T(\tau) = L_{T'}(\tau)$,*
- *If T is A, B-escapable and $\sigma_0, \dots, \sigma_k$ is in B then T' is A, B-escapable as well,*
- *If T is A, B-escapable and $\sigma_0, \dots, \sigma_k$ avoids B then T' is A, B-escapable as well.*

Proof. Figure 2 illustrates this lemma. Formally, let the width of $\sigma_0, \dots, \sigma_k$ be 2^{-m} with $m \leq n$. For any υ with $|\upsilon| = n - m$, by $\upsilon + 1$ we mean the result of viewing υ as a sequence mod 2 and adding 1 to it, so $010 + 1 = 011$ while $011 + 1 = 100$.

Define $T' \sqsupseteq T$ by:

- If $\tau = \sigma_k {}^\frown \rho$ where $|\rho| = n - m$ and υ is not all 1's then $T'(\tau) = \sigma_0 {}^\frown (\rho + 1)$,
- Otherwise $T'(\tau) = T(\tau)$.

Since this is our first construction of this kind, we point out that this is an operation on the description of T as a proper partial transformation: we

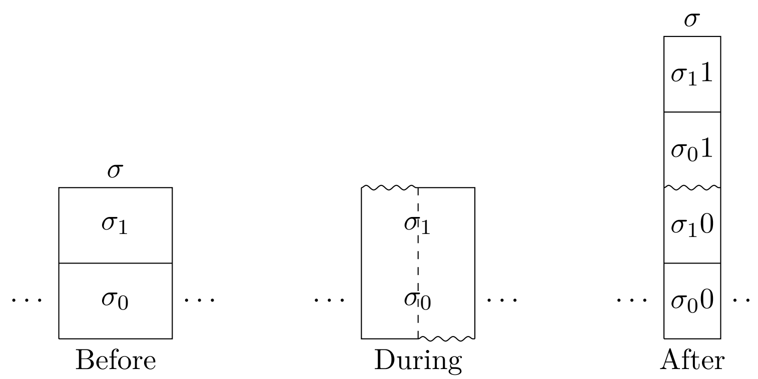

FIGURE 2. Thinning Loops, Lemma 3.2

have $\sigma_0, \ldots, \sigma_{k-1} \in T_+$ but $\sigma_k \in T_-$. We define

$$\begin{aligned} T'_+ = T_+ &\setminus \{\sigma_i \mid i < k\} \\ &\cup \{\sigma_i{}^\frown \upsilon \mid i < k, |\upsilon| = n - m\} \\ &\cup \{\sigma_k{}^\frown \upsilon \mid |\upsilon| = n - m, \upsilon \neq \langle 1, \ldots, 1\rangle\} \end{aligned}$$

and

$$T'_- = T_- \setminus \{\sigma_k\} \cup \{\sigma_k{}^\frown \langle 1, \ldots, 1\rangle\}.$$

We have explained how to define T' on all elements of $T'_+ \setminus T_+$ and $T'_- \setminus T_-$, and when σ is determined in T but not T' we set $T'(\sigma) = T(\sigma)$ and so have completely specified the new transformation T'.

Propriety and the fact that T' is partitioned into open loops are trivial.

To see escapability, consider some υ_0 determined in T' such that $|T'(\upsilon_0)| < |\upsilon_0|$ and fix a reduced escape sequence $\upsilon_1, \ldots, \upsilon_r$ for υ_0 in T. If $\upsilon_0 \in \cup[\sigma_i]$ then by Lemma 2.9 we have $\upsilon_0 \sqsupseteq \sigma_k$ and no other element of the escape sequence belongs to $\cup[\sigma_i]$, and therefore $\upsilon_1, \ldots, \upsilon_r$ is an escape sequence in T' as well.

If $\upsilon_0 \notin \cup[\sigma_i]$ but there is a $j > 0$ such that $\upsilon_{j+i} \sqsupseteq \sigma_i$ for $i \leq k$ then by Lemma 2.8 we may assume that for $j > 0$, $|\upsilon_j| \geq n$. Then since both $T(\upsilon_{j-1}) \sqsubseteq \upsilon_j$ and $\sigma_0 \sqsubseteq \upsilon_j$, $T(\upsilon_{j-1})$ and σ_0 are comparable. Since σ_0 is unblocked, we must have $T(\upsilon_{j-1}) \sqsubset \sigma_0$, and therefore for any ρ of suitable length, the sequence

$$\upsilon_1, \ldots, \upsilon_{j-1}, \sigma_0{}^\frown \langle 0, \ldots, 0\rangle^\frown \rho, \ldots, \sigma_k{}^\frown \langle 1, \ldots, 1\rangle^\frown \rho, \upsilon_{j+k+1}, \ldots, \upsilon_n$$

is an escape sequence for υ_0 in T'.

To see that we preserve A, B-escapability, if υ_0 is in A then we could have chosen the original escape sequence $\upsilon_1, \ldots, \upsilon_r$ in B, and therefore (since either $\sigma_0, \ldots, \sigma_k$ is in B or avoids B), the same argument shows that in T' there is an escape sequence for υ_0 in B. □

Remark 3.3. In the previous lemma, we actually have slightly more control over escape sequences: for any set B such that $\sigma_0, \ldots, \sigma_k$ avoids B, any escape sequence in T is an escape sequence in T'. In particular, in this situation we do not change the lengths of escape sequences.

We also need a modified version of the above lemma where instead of wanting $\cup_{j\leq k'}[\tau_i] = \cup_{i\leq k}[\sigma_i]$ we want to have a small amount of the original open loop left alone.

Lemma 3.4. *Let T be a useful partial transformation, let $\sigma_0, \dots, \sigma_k$ be an open loop of determined elements and let $\epsilon = 2^{-n}$ be smaller than the width of this open loop. Then there is a useful $T' \sqsupseteq T$ such that*

> *There is an open loop $\tau_0, \dots, \tau_{k'}$ in T' of width ϵ such that $\lambda(\bigcup_{i\leq k}[\sigma_i] \setminus \bigcup_{j\leq k'}[\tau_i]) = \epsilon \cdot (k+1)$.*

Furthermore:

- *If $\tau \notin \cup_{i\leq k'}[\tau_i]$ is determined in T' then $T'(\tau) = T(\tau)$ and $L_T(\tau) = L_{T'}(\tau)$,*
- *If T is A, B-escapable and $\sigma_0, \dots, \sigma_k$ is in B then T' is $A, B\setminus\bigcup_{i\leq k'}[\tau_i]$-escapable,*
- *If T is A, B-escapable and $\sigma_0, \dots, \sigma_k$ avoids B then T' is A, B-escapable.*

Proof. We proceed exactly as above except that we replace the first clause in the definition of T' with

> If $\tau = \sigma_k{}^\frown\upsilon$ where $|\upsilon| = n - m$ and υ is neither all 1's nor all 1's with a single 0 at the end then $T'(\tau) = \sigma_0{}^\frown(\upsilon + 1)$.

Equivalently, we place both $\sigma_k{}^\frown\langle 1, \dots, 1, 0\rangle$ and $\sigma_k{}^\frown\langle 1, \dots, 1\rangle$ in T'_- and all other extensions of σ_k in T'_+.

We check the stronger escapability condition. If υ_0 is in A and is determined in T' with $|T'(\upsilon_0)| < |\upsilon_0|$, take a reduced escape sequence $\upsilon_1, \dots, \upsilon_r$ for υ_0 in T in B. By Lemma 3.2, we may assume $|\upsilon_1| \geq n$. The only non-trivial case is if $\upsilon_0 \notin \bigcup[\sigma_i]$ but there is a $j > 0$ such that $\upsilon_{j+i} \sqsupseteq \sigma_i$ for $i \leq k$. Let $\upsilon_{j+i} = \sigma_i{}^\frown\rho$ (note that ρ does not depend on i). For each $i \leq k$, we may replace υ_{j+i} with $\sigma_i{}^\frown\langle 1, \dots, 1\rangle^\frown\rho'$ for any ρ' of appropriate length; since $T(\upsilon_{j-1}) \sqsubseteq \sigma_0$ (because σ_0 is determined and T is proper), this remains an escape sequence, and the modified escape sequence avoids $B \setminus \bigcup_{i\leq k'}[\tau_i]$. □

Remark 3.5. As before, the escape sequences in T' promised by the last two conditions in this lemma always have the same length as the ones in T.

The next lemma illustrates the use of escape sequences: we take some σ_0 and an escape sequence in T and extend T to a new partial transformation T' with the property that $\sigma_0 \in T'_+$ and the open loop $\tau_0, \dots, \tau_{k'}$ in T' which contains σ_0 has the property that $T'(\tau_{k'}) = \langle\rangle$. In other words, we can arrange for any σ_0 to belong to a tower which has $\langle\rangle$ on top.

Lemma 3.6 (Escape)**.** *Let T be a useful partial transformation, let σ_0 be determined with $|T(\sigma_0)| < |\sigma_0|$, and let $\sigma_1, \dots, \sigma_k$ be a reduced escape sequence for σ_0 such that $|\sigma_0| = |\sigma_1| + 1$. Then there is a useful $T' \sqsupseteq T$ such that*

There is an open loop $\tau_0, \dots, \tau_{k'}$ *in* T' *with* $[\sigma_0] \subseteq \bigcup_{i \le k'} [\tau_i]$ *such that* $T'(\tau_{k'}) = \langle\rangle$.

Furthermore,

- *If* $\tau \notin \cup_{i \le k'} [\tau_i]$ *is determined in* T' *then* $T'(\tau) = T(\tau)$ *and* $L_T(\tau) = L_{T'}(\tau)$,
- *If* T *is* A, B*-escapable where* $\tau_0, \dots, \tau_{k'}$ *is in* B *then* T' *is* A, B*-escapable,*
- *If* T *is* A, B*-escapable and* σ_0 *avoids* B *then* T' *is* $A \setminus \bigcup_{i \le k'} [\tau_i], B \setminus \bigcup_{i \le k'} [\tau_i]$*-escapable.*

Proof. Let T_-, T_+ witness that T is proper and extend T by defining $T'(\sigma_0) = \sigma_1 ^\frown \langle 0 \rangle$ and for each $i \in (0, k)$, $T'(\sigma_i ^\frown \langle 0 \rangle) = \sigma_{i+1} ^\frown \langle 0 \rangle$, and for all τ which do not extend some σ_i with $i < k$, $T'(\tau) = T(\tau)$.

To see that T' is proper, we need only check that if $\sigma \in T'_+$ and $\sigma \neq \tau \in T'_+ \cup T'_-$ then $T'(\tau) \not\sqsupseteq T'(\sigma)$. Clearly we need only check this for $T'(\sigma) = \sigma_i ^\frown \langle 0 \rangle$. Since the escape sequence was reduced, we cannot have $\sigma_i = \sigma_j$ for $i \neq j$, so we can restrict our attention to the τ such that $T'(\tau) = T(\tau)$. If σ_i was not blocked in T then there is no such τ, and if σ_i was blocked in T then already $T(\sigma_{i-1} ^\frown \langle 0 \rangle) = \sigma_i ^\frown \langle 0 \rangle$, and the claim follows since T was proper.

It is easy to see that T' remains partitioned into open loops.

We check that T' is escapable. Let υ_0 be given with $|T'(\upsilon_0)| < |\upsilon_0|$. Then the same was true in T, so υ_0 had an escape sequence $\upsilon_1, \dots, \upsilon_r$ in T. We may assume $|\upsilon_1| \geq |\sigma_1| + 1$. There are a few potential obstacles we need to deal with. First, it could be that for some i, $\upsilon_i \sqsupseteq \sigma_0$. Letting $\upsilon_i = \sigma_0 ^\frown \rho$, we must have that $\upsilon_1, \dots, \sigma_0 ^\frown \rho, \sigma_1 ^\frown \langle 0 \rangle ^\frown \rho, \dots, \sigma_k ^\frown \langle 0 \rangle ^\frown \rho$ is also an escape sequence for υ_0.

Suppose not. There could be some i and $j > 0$ such that $\upsilon_i \sqsupseteq \sigma_j ^\frown \langle 0 \rangle$. Let j be least such that this occurs. We cannot have $i = 0$ since $|T(\sigma_j ^\frown \langle 0 \rangle ^\frown \rho)| = |\sigma_j ^\frown \langle 0 \rangle ^\frown \rho|$ (unless $j = k$, in which case $T(\upsilon_0) = \langle\rangle$ and so the empty sequence is an escape sequence). If $T(\upsilon_{i-1}) \sqsupseteq \sigma_j ^\frown \langle 0 \rangle$ then σ_j was blocked in T, so $T(\sigma_{j-1}) = \sigma_j$, and therefore $\upsilon_{i-1} \sqsupseteq \sigma_{j-1} ^\frown \langle 0 \rangle$, contradicting the leastness of j. In particular, $\sigma_j ^\frown \langle 1 \rangle ^\frown \rho \sqsupseteq T(\upsilon_{i-1})$ as well. So whenever we have $\upsilon_i = \sigma_j ^\frown \langle 0 \rangle ^\frown \rho$, we replace it with $\upsilon' - i = \sigma_j ^\frown \langle 1 \rangle ^\frown \rho$, and the result is still an escape sequence.

Suppose T is A, B-escapable, $\tau_0, \dots, \tau_{k'}$ is in B, and υ_0 is in A. Then the argument just given, applied to an escape sequence in T in B, gives an escape sequence in T' in B.

Suppose T was A, B-escapable, $[\sigma_0]$ avoids B, and υ_0 is in A. Then, taking $\upsilon_1, \dots, \upsilon_r$ to be an escape sequence in T in B, we cannot have $\upsilon_i \sqsupseteq \sigma_0$, and so we are in the second case of the argument above, which gives an escape sequence in $B \setminus \bigcup_{i \le k'} [\tau_i]$. □

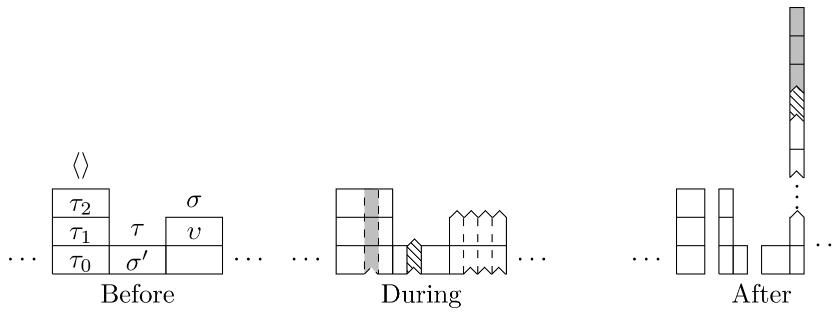

Figure 3.

Remark 3.7. Once again, the last two clauses actually ensure that when we have an escape sequence in T avoiding B, we actually have an escape sequence of the same length in T' avoiding B or $B \setminus \bigcup_{i \leq k'} [\tau_i]$.

The main building block of our construction will combine the steps given by these two lemmas as illustrated in Figure 3.

In the figure, we have $\sigma \sqsubset \sigma'$ and $\tau \sqsubset \tau_0$. We begin in a situation where we have an open loop (the one on the right in the "Before" diagram in Figure 3) and we wish to arrange the blocks so that the elements of that open loop belong to an open loop with $\langle\rangle$ on top such as the one on the left (so that we may later place any other open loop on top of it). The sequence $\sigma', \tau_0, \tau_1, \tau_2$ is an escape sequence for υ (and the end of an escape sequence for the block below υ). We could simply combine all these open loops—place σ' on top of υ and τ_0 on top of σ'—but we don't wish to do so, because we don't want to use up all of the escape sequence; we might be using $\sigma', \tau_0, \tau_1, \tau_2$ as part of an escape sequence for other elements as well, and we need some of it to remain.

So we apply Lemma 3.2 to the open loop on the right, replacing it with a much thinner open loop. Now we can apply Lemma 3.6, which takes subintervals from $\sigma', \tau_0, \tau_1, \tau_2$ and places them above the open loop on the right. Note that, by applying Lemma 3.2 with very small ϵ, we can make the total measure of the shaded portion as small as we like.

4. The Main Construction

Our main tool for causing the Birkhoff ergodic theorem to fail at a point is the notion of an upcrossing.

Definition 4.1. Given a measurable, measure-preserving, invertible $\widehat{T} : 2^\omega \to 2^\omega$, a point $x \in 2^\omega$, a measurable f, and rationals $\alpha < \beta$, an *upcrossing sequence* for α, β of length N is a sequence

$$0 \leq u_1 < v_1 < u_2 < v_2 < \cdots < u_N < v_N$$

such that for all $i \leq N$,

$$\frac{1}{u_i+1}\sum_{j=0}^{u_i} f(\widehat{T}^j x) < \alpha,\ \frac{1}{v_i+1}\sum_{j=0}^{v_i} f(\widehat{T}^j x) > \beta.$$

$\tau(x, f, \alpha, \beta)$ is the supremum of the lengths of upcrossing sequences for α, β.

By definition, Birkhoff's ergodic theorem fails at x exactly if $\tau(x, f, \alpha, \beta) = \infty$ for some $\alpha < \beta$. Our plan is to look at an Martin-Löf test $\langle V_j \rangle$ and, as sequences σ are enumerated into an appropriate V_j, ensure that the lower bound on $\tau(x, f, 1/2, 3/4)$ increases for each $x \in [\sigma]$.

While building a transformation as the limit of a sequence of partial transformations, we would like to be able to ensure at some finite stage in the sequence that certain points have many upcrossings. The following notion is the analog of an upcrossing sequence for a partial transformation.

Definition 4.2. Let T be a useful partial transformation, $A \subseteq 2^\omega$ and $\tau_0, \dots, \tau_n$ an open loop in T such that for each i either $[\tau_i] \subseteq A$ or $[\tau_i] \subseteq 2^\omega \setminus A$. Let $R = \{i \leq n \mid [\tau_i] \subseteq A\}$ and $\alpha < \beta$ rationals. A *τ_s-upcrossing sequence* for α, β of length N is a sequence

$$0 \leq u_1 < v_1 < u_2 < v_2 < \cdots < u_N < v_N \leq n$$

such that for all $i \leq N$,

$$\frac{1}{u_i+1}\sum_{j=s}^{u_i+s} \chi_R(j) < \alpha,\ \frac{1}{v_i+1}\sum_{j=s}^{v_i+s} \chi_R(j) > \beta.$$

Note that we shift indices in the sums over by s, since we begin with τ_s and count to later elements of the open loop.

Lemma 4.3. *Let T be a useful partial transformation, let $A \subseteq 2^\omega$, and $\tau_0, \dots, \tau_n$ an open loop in T such that for each i either $[\tau_i] \subseteq A$ or $[\tau_i] \subseteq 2^\omega \setminus A$. Let $0 \leq u_1 < v_1 < u_2 < v_2 < \cdots < u_N < v_N \leq n$ be a τ_s-upcrossing sequence for α, β.*

Then whenever $\widehat{T} : 2^\omega \to 2^\omega$ extends T, and $x \in [\tau_s]$, $u_1 < v_1 < \cdots < u_N < v_N$ is an upcrossing sequence for α, β in $\widehat{T}$ with the function χ_A.

We are finally ready to give our main construction.

Theorem 4.4. *Suppose $x \in 2^\omega$ is not Martin-Löf random. Then there is a computable set A and a computable transformation $\widehat{T} : 2^\omega \to 2^\omega$ such that x is not typical with respect to the ergodic theorem.*

Proof. Let $\langle V_j \rangle$ be a Martin-Löf test witnessing that x is not Martin-Löf random, so $x \in \cap_j [V_j]$ and $\lambda([V_j]) \leq 2^{-j}$. We write $V_j = \bigcup_{j,n} V_{j,n}$ where $(j, n) \mapsto V_{j,n}$ is computable. We will assume $V_{j,0} = \emptyset$ for all j and that $n < m$ implies $V_{j,n} \subseteq V_{j,m}$, and we refer to $V_{j,n+1} \setminus V_{j,n}$ as the portion of V_j enumerated at stage $n + 1$. It is convenient to assume that there is at most one element enumerated into *any* V_j at stage $n+1$; that is, $\sum_j |V_{j,n+1} \setminus V_{j,n}| \leq 1$, and so $\sum_j |V_{j,n}|$ is finite for any n.

We will construct an increasing sequence of useful partial transformations $T_0 \sqsubseteq T_1 \sqsubseteq T_2 \sqsubseteq \cdots$ so that setting $T = \lim_n T_n$, $\widehat{T}(x) = \lim_{n\to\infty} T(x \upharpoonright n)$ is the desired transformation. We first list the technical requirements on our induction; since they are rather elaborate, we will go through what the intended meanings are before describing the actual construction.

Inductive Specification: We now specify the properties that will be maintained at each stage of our construction. Since they are rather complicated, we recommend that the reader skip them at first; a detailed explanation of their intended meaning is given after.

We will define, as part of our construction of stage 0, a set A, a computable collection of finite sequences which we call *componential*, and a computable function d defined on componential σ.

At stage n we will have a partial transformation T_n, a partition $2^\omega = W^n \cup \bigcup_k (A^n_k \cup B^n_k)$ into components which are finite unions of intervals, a further partition $W^n = \bigcup_k W^n_k$ into finite unions of intervals, constants a^n_k, b^n_k, a function ρ^n, and for each $i < n/2$, a finite union of intervals $G_i \subseteq 2^\omega$, such that the following properties hold:

(1) *Structure of components*
- (1.a) T_n is useful,
- (1.b) $T_n \sqsubseteq T_{n+1}$,
- (1.c) $T_{n,+} \subseteq W^n$,
- (1.d) Each open loop in T_n belongs entirely to one component,
- (1.e) Each W^n_k is a union of intervals of the form $[\tau]$ with τ determined,
- (1.f) If σ is componential and determined then $[\sigma]$ is contained in a single component in T_n,
- (1.g) $A^n_k \subseteq A^{n+1}_k \cup W^{n+1}_0$,
- (1.h) $B^n_k \subseteq B^{n+1}_k \cup W^{n+1}_0$,
- (1.i) $W^n_k \subseteq W^{n+1}_k \cup W^{n+1}_{k+1}$,
- (1.j) For each k, T_n is A^n_k, A^n_k-escapable,
- (1.k) For each k, T_n is B^n_k, B^n_k-escapable,
- (1.l) T_n is W^n, W^n-escapable,

(2) *Management of upcrossings*
- (2.a) If $\sigma \in W^n_k$ is determined then in the open loop in T_n containing σ, there is an upcrossing sequence for σ of length k for $2^\omega \setminus A$ and $1/2, 3/4$,
- (2.b) The domain of ρ^n is the elements of W^n determined in T_n,
- (2.c) If $\sigma \sqsubseteq \tau$ are both determined and in W^n then $\rho^n(\sigma) = \rho^n(\tau)$,
- (2.d) If $\rho^n(\sigma) = \rho^n(\tau)$ and there is a k with $\sigma, \tau \in W^n_k$ then $L_{T_n}(\sigma) = L_{T_n}(\tau)$,
- (2.e) If $\sigma \in W^{n+1}_k \cap W^n_k$ is determined in T_{n+1} then
  - (2.e.i) $\rho^{n+1}(\sigma) = \rho^n(\sigma)$,
  - (2.e.ii) $L_{T_{n+1}}(\sigma) = L_{T_n}(\sigma)$,

(2.f) For each k, $0 \leq a^n_k < \lambda(A^n_k)$ and $0 \leq b^n_k < \lambda(B^n_k)$,
(2.g) For each k, let J^n_k be the image of $\rho^n \upharpoonright W^n_k$, and for each $j \in J^n_k$, let $l_{k,j}$ be the (by (2.d), necessarily unique) value of $L_{T_n}(\tau)$ for some $\tau \in W^n_k$; then
(2.g.i) $\sum_{j\in J^n_k} l_{k,j} \left(2^{-j} - \lambda([V_{j,(n-1)/2}])\right) \leq a^n_k$,
(2.g.ii) $4\sum_{j\in J^n_k} l_{k,j} \left(2^{-j} - \lambda([V_{j,(n-1)/2}])\right) \leq b^n_k$,
(2.h) If $\sigma \in W^n$ is determined in T_n then $[V_{\rho^n(\sigma),(n-1)/2}] \cap [\sigma] = \emptyset$,
(3) *Almost everywhere defined*
(3.a) For each $i < n/2$ and each $\sigma \in T_{n,+} \cup T_{n,-}$, σ is either in G_i or avoids G_i,
(3.b) For any $i_0, \dots, i_{k-1}$, $\lambda(\bigcap_{j\leq k} G_{i_j}) \leq 2^{-k}$,
(3.c) If $\sigma \in T_{n,-}$ then $|\sigma| \geq n/2$,
(3.d) For each componential $\sigma \in T_{n,-}$, $|T_n(\sigma)| \geq |\{i < n/2 \mid \sigma \text{ avoids } G_i\}| - d(\sigma)$.

Explanation of inductive clauses: Initially we will fix a partition into three regions, W, A, B. W will be a region known to contain x, and $W \cup B$ will be the set demonstrating the failure of the ergodic theorem for x. Our strategy will be that when we enumerate some τ into V_j for an appropriate j, we will arrange to add an upcrossing by first mapping every element of $[\tau]$ through A for a long time, ensuring that the ergodic average falls below 1/2. We will then have the transformation map those elements through B for a long time to bring the average back up to 3/4. We will do this to each element of $\cap_j[V_j]$ infinitely many times, ensuring that elements in this intersection are not typical.[2]

We will have further partitions $A = \bigcup A_k$ and $B = \bigcup B_k$; A_k is the portion of A reserved for creating the $k+1$-st upcrossing, and B_k the portion of B reserved for the same. We will start with $W^0_0 = W$, $A^0_k = A_k$, and $B^0_k = B_k$. At later stages, W^n_k will be the portion of W^n known to have at least k upcrossings (2.a). At a given step, we might move intervals from W^n_k to W^{n+1}_{k+1} (1.i) (because we have created a new upcrossing) and we might move intervals from A^n_k to W^{n+1}_0 (because it is part of a newly created upcrossing) and similarly for B^n_k (1.g),(1.h). In any other case, an interval in W^n_k is in W^{n+1}_k and similarly for A, B. (It is perhaps slightly confusing that we will use the term "component" to mean either W^n or A^n_k or B^n_k for some k, but that the W^n_k are not themselves components. However it will become clear that most properties—the behavior of loops and escapability, for instance—respect components in the sense in which we are using the term.)

When points are in A^n_k or B^n_k, they are simply waiting to (possibly) be used in the creation of an upcrossing, so they always belong to open loops

[2]It is not possible to ensure that every element of the set V_j receives j upcrossings, since this would imply that the theorem holds for every x which failed to be even Demuth random, which would contradict V'yugin's theorem.

with length 1 (1.c). All open loops with length longer than 1 wil be entirely within W^n (1.d). Similarly, escapability is handled componentwise (1.j)-(1.l). Combining these properties, the point is that there is simply no interaction between separate components except when we create upcrossings, and when we create an upcrossing, we put everything in W^n.

Suppose τ is enumerated into V_j, causing us to create a new upcrossing. We now need to look for extensions of τ to be enumerated into $\bigcup_{j'>j} V_{j'}$; specifically, we should choose a particular $j' > j$ and watch for an extension of τ to be enumerated into $V_{j'}$. We cannot simply choose $j' = j + 1$, since we may not have enough measure available in A^n_k. Instead we choose a new value for $\rho^n(\tau)$, and we will wait for extensions of τ in $V_{\rho^n(\tau)}$. (2.h) ensures that if we see some extension of τ get enumerated into $V_{\rho^n(\tau)}$, we will be forced to create a new upcrossing, and conversely (2.e.i) ensures that this is the only time we change $\rho^n(\tau)$.

When we assign $\rho^n(\tau) = j'$, we have to consider the worst case scenario: $\lambda([V_{j'}]) = 2^{j'}$ and $[V_{j'}] \subseteq [\tau]$. We therefore need enough measure in A^n_k and B^n_k to create the needed upcrossings. The amount of measure we needed is determined not only by the size of $[V_{j'}]$, but also by the length $L_{T_n}(\tau)$ of the loop containing τ. To simplify the calculations, we require that all sequences sharing a value of ρ^n have the same length L_{T_n} (2.d) and that this value does not change except when we create upcrossings (2.e.ii).

The values a^n_k and b^n_k keep track of the assigned portions of A^n_k and B^n_k respectively. We should never assign all of A^n_k, B^n_k because we need some space for upcrossings (2.f). When we assign $\rho^n(\tau) = j'$, we will increase a^n_k and b^n_k accordingly, and when we create an upcrossing we will use some of A^n_k, B^n_k, move the corresponding intervals into W^{n+1}, and decrease a^{n+1}_k, b^{n+1}_k accordingly (2.g.i),(2.g.ii).

Finally, we need to make sure that the transformation is defined almost everywhere. (3.d) ensures that if σ avoids many of the sets G_i then $T_n(\sigma)$ is large, and (3.b) ensures that most σ avoid many G_i. A technicality is that we want to extend T_n while respecting the fact that escape sequences are supposed to remain within components. We will call σ *componential* if $[\sigma]$ belongs to a single component in T_0; while there are infinitely many minimal componential elements (each A^0_k, for instance, instance, is an interval $[\tau]$, so τ is componential but no initial segment is), every x belongs to some $[\tau]$ with τ componential. $d(\sigma)$ will be the length of the smallest $\tau \sqsubseteq \sigma$ such that τ is componential. We will first make sure that $T_{n,-}$ is defined on longer and longer sequences (3.c), and then once we reach a componential portion (after at most $d(\sigma)$ steps), we will start making T_n longer and longer.

Stage 0: We now give the initial stage of our construction. We assume without loss of generality that x belongs to some interval W^0 with $\lambda(W^0) < 1$. For instance, we can suppose we know the first bit of x and let W^0 equal $[\langle 0\rangle]$ or $[\langle 1\rangle]$ as appropriate. Then, from the remaining measure, we take A^0_k and B^0_k to be intervals so that $\lambda(B^0_k) = 4\lambda(A^0_k)$ for all k. We set $W^0_0 = W^0$

and take T_0 to be the trivial transformation (i.e., $T_0(\sigma) = \langle\rangle$ for all σ). Choose j large enough that $2^{-j} < \lambda(A^0_0)$ and set $\rho^0(\sigma) = j$ for all σ. Set $a^0_0 = 2^{-j}$, $b^0_0 = 4 \cdot 2^{-j}$, and for $k > 0$, $a^0_k = b^0_k = 0$.

We define σ to be componential if either $[\sigma] \subseteq W^0$ or there is some k such that $[\sigma] \subseteq A^0_k$ or $[\sigma] \subseteq B^0_k$. For any componential σ, $d(\sigma)$ is the length of the smallest $\tau \sqsubseteq \sigma$ such that τ is componential.

Odd stages—ensuring T is total: Given T_n for an even n, we take steps at the odd stage $n+1$ to make sure that T is defined almost everywhere. We will define $T_{n+1,+} = T_{n,+}$ and

$$T_{n+1,-} = \cup_{\sigma\in T_{n,-}} \{\sigma^\frown\langle 0\rangle, \sigma^\frown\langle 1\rangle\}.$$

Consider some $\sigma \in T_{n,-}$. We set $T_{n+1}(\sigma^\frown\langle 0\rangle) = T_n(\sigma)$. If $[\sigma]$ is not componential then $T_{n+1}(\sigma^\frown\langle 1\rangle) = T_n(\sigma)$ as well.

If σ is componential, let τ be the initial element of the open loop containing σ, so $T_n(\sigma) \sqsubset \tau$. In particular, there is a $b \in \{0, 1\}$ such that $T_n(\sigma)^\frown\langle b\rangle \sqsubseteq \tau$, and we set $T_{n+1}(\sigma^\frown\langle 1\rangle) = T_n(\sigma)^\frown\langle b\rangle$. We define $W^{n+1} = W^n$, $A^{n+1}_k = A^n_k$, $B^{n+1}_k = B^n_k$, $a^{n+1}_k = a^n_k$, $b^{n+1}_k = b^n_k$ for all k. We define $\rho^{n+1}(\sigma) = \rho^n(\sigma)$ for determined $\sigma \in W^n$. We define $G_{n/2} = \bigcup_{\sigma\in T_{n,-}}[\sigma^\frown\langle 0\rangle]$.

Most of the inductive properties follow immediately from the inductive hypothesis since we changed neither the components nor the lengths of any open loops. To check (1.j)-(1.l), it suffices to show that whenever $\sigma \in T_{n,-}$ and $b \in \{0, 1\}$, there is an escape sequence for $\sigma_0^\frown\langle b\rangle$ in T_{n+1} belonging to the same component as σ_0. Given an escape sequence $\sigma_1, \dots, \sigma_k$ for σ_0 belonging to the same component, the only potential obstacle is that one or more σ_i has the form $\sigma_i^{-\frown}\langle 1\rangle^\frown\rho$ where $\sigma^- \in T_{n,-}$. If $[\sigma_i^-]$ is not componential, there is no obstacle. If σ_i^- is componential then everything in $[\sigma_i^-]$ must belong to the same component; let $\tau_0, \dots, \tau_r, \sigma_i^-$ be the open loop containing σ_i^- in T_n, and observe that

$$\sigma_1, \dots, \sigma_i^{-\frown}\langle 1\rangle, \tau_0{}^\frown\langle 1\rangle^\frown\rho, \dots, \tau_r{}^\frown\langle 0\rangle^\frown\rho, \sigma_{i+1}, \dots, \sigma_k$$

is an escape sequence for σ_0 in T_n, and since $[\sigma_i^-]$ belongs to this component and each open loop is contained in one component, this new escape sequence is contained entirely in the component of σ_0. After we have replaced each such $\sigma_i^{-\frown}\langle 1\rangle^\frown\rho$ in this way, we have an escape sequence for σ_0 in T_n which is also an escape sequence in T_{n+1} contained entirely in one component.

(3.a) holds for $i = n/2$ by construction. To see (3.b), consider the sequence $i_0, \dots, i_{k-1}, n/2$. By the inductive hypothesis, $\lambda(\bigcap_{j\leq k} G_{i_j}) \leq 2^{-k}$. Each $\sigma \in T_{n,+} \cup T_{n,-}$ is either in $\bigcap_{j\leq k} G_{i_j}$ or avoids $\bigcap_{j\leq k} G_{i_j}$, so $\bigcap_{j\leq k} G_{i_j} \cap G_{n/2}$ is exactly the union of those $[\sigma^\frown\langle 0\rangle]$ such that σ is in $\bigcap_{j\leq k} G_{i_j}$. In particular, whenever $[\sigma^\frown\langle 0\rangle] \subseteq \bigcap_{j\leq k} G_{i_j} \cap G_{n/2}$, we must have $[\sigma] \subseteq \bigcap_{j\leq k} G_{i_j}$, so $\lambda(\bigcap_{j\leq k} G_{i_j} \cap G_{n/2}) \leq 2^{-1}\lambda(\bigcap_{j\leq k} G_{i_j}) \leq 2^{-k-1}$.

If $\sigma^\frown\langle b\rangle \in T_{n+1,-}$ then $\sigma \in T_{n,-}$ and by the inductive hypothesis, $|\sigma| \geq n/2$, so $|\sigma^\frown\langle b\rangle| = |\sigma| + 1 \geq (n+1)/2$ since n is even, showing (3.c).

To see (3.d), consider some componential $\sigma^\frown\langle b\rangle \in T_{n+1,-}$. If σ was componential then we have $\sigma \in T_{n,-}$ and by the inductive hypothesis, $|T_n(\sigma)| \geq |\{i < n/2 \mid \sigma \text{ avoids } G_i\}| - d(\sigma)$. We have

$$\begin{aligned}|T_{n+1}(\sigma^\frown\langle 0\rangle)| &= |T_n(\sigma)| \\ &\geq |\{i < n/2 \mid \sigma \text{ avoids } G_i\}| - d(\sigma) \\ &= |\{i < (n+1)/2 \mid \sigma^\frown\langle 0\rangle \text{ avoids } G_i\}| - d(\sigma^\frown\langle 0\rangle)\end{aligned}$$

and

$$\begin{aligned}|T_{n+1}(\sigma^\frown\langle 1\rangle)| &= |T_n(\sigma)| + 1 \\ &\geq |\{i < n/2 \mid \sigma \text{ avoids } G_i\}| - d(\sigma) + 1 \\ &= |\{i < (n+1)/2 \mid \sigma^\frown\langle 1\rangle \text{ avoids } G_i\}| - d(\sigma^\frown\langle 1\rangle).\end{aligned}$$

If σ was not componential then $d(\sigma^\frown\langle b\rangle) = |\sigma^\frown\langle b\rangle| \geq (n+1)/2$, so

$$|T_{n+1}(\sigma^\frown\langle b\rangle)| \geq 0 \geq |\{i < (n+1)/2 \mid \sigma^\frown\langle b\rangle \text{ avoids } G_i\}| - d(\sigma^\frown\langle b\rangle).$$

Even stages—creating upcrossings; the construction: We now consider the real work. When $n = 2n' + 1$ is odd, we take steps at the even stage $n + 1$ to ensure that the ergodic theorem does not hold for any element of $\cap_j [V_j]$ by adding an additional upcrossing to some element enumerated into one of the V_js. As noted above, we assume that there is exactly one π in $\bigcup_j V_{j,n'} \setminus \bigcup_j \bigcup_{m<n'} V_{j,n'}$, so let $\hat{j}$ be such that $\pi \in V_{\hat{j},n'} \setminus \bigcup_{m<n'} V_{\hat{j},m}$. If π is not determined, we may partition $[\pi] = \bigcup_i [\pi_i]$ where each π_i is determined, and we will deal with each π_i separately.

We first consider how to deal with a single element. Let $\tau = \pi$ if π is determined or $\tau = \pi_i$ for some i is π is not determined. We will construct a transformation T_{n+1} and other objects (W_k^{n+1}, etc.) satisfying all the conditions above other than (2.h). (We will discuss (2.h) after.) If $\rho^n(\tau) \neq \hat{j}$, we will just take $T_{n+1} = T_n$ and similarly for all other objects.

So we consider the case where $\rho^n(\tau) = \hat{j}$. We have $\tau \in W_k^n$ for some k. We will extend the open loop containing τ to a longer open loop which will contain a long subsequence coming from A_k^n and then an even longer subsequence coming from B_k^n, thereby creating a new upcrossing in T_{n+1} for each element of $[\tau]$. We illustrate the intended arrangement in Figure 4. We must ensure that every point in $[\tau]$, a section of fixed total measure, receives a new upcrossing. We must do so while ensuring that the total measure of the portions of A_k^n used is strictly less than $L_{T_n}(\tau)\lambda([\tau]) + (\lambda(A_k^n) - a_k^n)$ and the total measure of the portions of B_k^n used is strictly less than $4L_{T_n}(\tau)\lambda([\tau]) + (\lambda(B_k^n) - b_k^n)$. Finally, the escape sequences will all have a fixed height, which we cannot expect to bound in advance. Our solution will be to thin all the parts other than the escape sequences until the entire tower is so narrow that we can afford the error introduced by the various escape sequences.

We now fix the following:

- Let $\tau_1, \ldots, \tau_t$ be the open loop containing τ,

- Identify a length e^t such that there is a reduced escape sequence for τ_t of length e^t contained in W^n,
- For some finite V and each $i \leq V$, let $\nu_i \in B^n_k$ be such that the ν_i are pairwise distinct, prefix-free, each ν_i is determined, and

$$\lambda(B^n_k) > \sum_{i\leq V}\lambda([\nu_i]) > 4\sum_{i\leq t}\lambda([\tau_i]) = 4L_{T_n}(\tau)\lambda([\tau]).$$

  Such ν_i must exist because $\lambda(B^n_k) > b^n_k \geq 4L_{T_n}(\tau)\lambda([\tau])$.
- For some finite U and each $i \leq U$, let $\upsilon_i \in A^n_k$ be such that the υ_i are pairwise distinct, prefix-free, each υ_i is determined, and

$$\min\{\lambda(A^n_k), \frac{1}{4}\sum_{i\leq V}\lambda([\nu_i])\} > \sum_{i\leq U}\lambda([\upsilon_i]) > \sum_{i\leq t}\lambda([\tau_i]) = L_{T_n}(\tau)\lambda([\tau]).$$

  Such υ_i must exist because $\lambda(A^n_k) > a^n_k \geq L_{T_n}(\tau)\lambda([\tau])$.
- For each $i \leq U$, identify an e^u_i such that there is a reduced escape sequence for υ_i of length e^u_i contained entirely in A^n_k, and set $e_u = \sum_{i\leq U} e^u_i$,
- For each $i \leq V$, identify an e^v_i such that there is a reduced escape sequence for ν_i of length e^v_i contained entirely in B^n_k, and set $e_v = \sum_{i\leq V} e^v_i$,

Choose N, N' sufficiently large (as determined by the argument to follow). We apply Lemma 3.2 to the open loop $\tau_1, \dots, \tau_t$ in T_n with $\epsilon = 2^{-(N+N')}$ to obtain a partial transformation $T^{0,0}_n$ where the width of the open loop containing τ is ϵ.

Note that each υ_i is an open loop in its own right, and similarly for each ν_i (since an open loop containing υ_i consists entirely of elements in A^n_k, and no element of A^n_k belongs to $T_{n,+}$).

Given $T^{0,i}_n$, let $T^{0,i+1}_n$ be the result of applying Lemma 3.4 to the open loop υ_i in $T^{0,i}_n$ with $\epsilon = 2^{-N}$. In $T^{0,i+1}_n$ there is a open loop $\upsilon'_{i,0}, \dots, \upsilon'_{i,u'_i}$ with width 2^{-N}, $\lambda([\upsilon_i] \setminus \bigcup_{j\leq u'_i}[\upsilon'_{i,j}]) = 2^{-N}$.

Let $T^{1,0}_n = T^{0,U+1}_n$; given $T^{1,i}_n$, let $T^{1,i+1}$ be the result of applying Lemma 3.2 to the open loop $\upsilon'_{i,0}, \dots, \upsilon'_{i,u'_i}$ in $T^{1,i}_n$ with $\epsilon = 2^{-(N+N')}$. In $T^{1,i+1}_n$ there is an open loop $\upsilon''_{i,0}, \dots, \upsilon''_{i,u''_i}$ with width $2^{-(N+N')}$ and $\lambda([\upsilon_i]\setminus\bigcup_{j\leq u''_i}[\upsilon''_{i,j}]) = 2^{-N}$.

Let $T^{2,0}_n = T^{1,U+1}_n$; given $T^{2,i}_n$, let $T^{2,i+1}_n$ be the result of applying Lemma 3.4 to the open loop ν_i in $T^{0,i}_n$ with $\epsilon = 2^{-N}$. In $T^{2,i+1}_n$ there is an open loop $\nu'_{i,0}, \dots, \nu'_{i,v'_i}$ with width 2^{-N} and $\lambda([\nu_i] \setminus \bigcup_{j\leq v'_i}[\nu'_{i,j}]) = 2^{-N}$.

Let $T^{3,0}_n = T^{2,V+1}_n$; given $T^{3,i}_n$, let $T^{3,i+1}_n$ be the result of applying Lemma 3.2 to the open loop $\nu'_{i,0}, \dots, \nu'_{i,v'_i}$ in $T^{3,i}_n$ with $\epsilon = 2^{-(N+N')}$. In $T^{3,i+1}_n$ there is an open loop $\nu''_{i,0}, \dots, \nu'_{i,v''_i}$ with width $2^{-(N+N')}$ and $\lambda([\nu_i]\setminus\bigcup_{j\leq v''_i}[\nu''_{i,j}]) = 2^{-N}$.

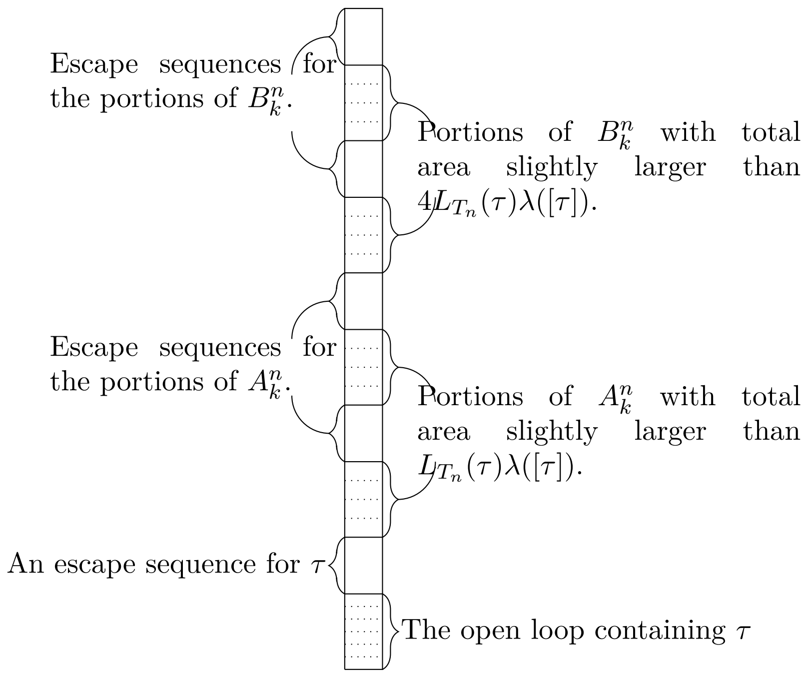

FIGURE 4. Construction of T'

We set $T_n^{4,0} = T_n^{3,V+1}$. For each $i \le U$, we choose an escape sequence for $\upsilon''_{i,u''_i}, \eta^i_1, \ldots, \eta^i_{e^u_i}$ contained entirely in $A^n_k \setminus \bigcup_{i\le U, j\le u''_i}[\upsilon''_{i,j}]$. To see that such an escape sequence exists, recall that an escape sequence for υ_i in A^n_k of the right length existed T_n, and since $\upsilon''_{i,u'_i} \sqsupseteq \upsilon_i$, the escape sequence was an escape sequence for υ''_{i,u'_i} as well. By Lemma 3.4 and Remark 3.5, the desired escape sequence existed in $T_n^{1,0}$, and its existence was preserved by the remaining steps. Further, if N and N' were chosen large enough, we may ensure that the collection of escape sequences $\{\bigcup_{j\le e^u_i}[\eta^i_j]\}$ is pairwise disjoint.

Given $T_n^{4,i}$, let $T_n^{4,i+1}$ be the result of applying Lemma 3.6 to the escape sequence $\eta^i_1, \ldots, \eta^i_{e^u_i}$ in $T_n^{4,i}$.

Let $T_n^{5,0} = T_n^{4,U+1}$. For each $i \le V$, we choose an escape sequence for $\nu''_{i,v''_i}, \theta^i_1, \ldots, \theta^i_{e^v_i}$. These escape sequences exist for the same reason as above. Given $T_n^{5,i}$, let $T_n^{5,i+1}$ be the result of applying Lemma 3.6 to the escape sequence $\nu''_{i,v''_i}, \theta^i_1, \ldots, \theta^i_{e^v_i}$ in $T_n^{5,i}$.

Let $T_n^6 = T_n^{5,V+1}$. Choose an escape sequence $\xi_1, \ldots, \xi_{e^t}$ for τ_t. This exists for the same reason as above. Let T_n^7 be the result of applying Lemma 3.6 to the escape sequence $\tau_t, \xi_1, \ldots, \xi_{e^t}$ in T_n^6.

$\bigcup_{i\le t}[\tau_i]$ is contained in an open loop in T_n^7 whose final element is $\xi_{e^t}{}^\frown\langle 0\rangle$. For each $i \le U$, $[\upsilon_i]$ is almost contained (except for a portion of measure $\le e_u 2^{-N}$) in an open loop with initial element $\upsilon''_{i,0}$ and final element $\eta^i_{e^u_i}{}^\frown\langle 0\rangle$.

For each $i \leq V$, $[\nu_i]$ is almost contained (except for a portion of measure $\leq e_v 2^{-N}$) in an open loop with initial element $\nu''_{i,0}$ and final element $\theta^i_{e^v_i} {}^\frown \langle 0 \rangle$. We define T_{n+1} by taking

$$T_{n+1,+} = T^7_{n,+} \cup \{\xi_{e^t} {}^\frown \langle 0 \rangle\} \cup \{\eta^i_{e^u_i} {}^\frown \langle 0 \rangle \mid i \leq U\} \cup \{\theta^i_{e^v_i} {}^\frown \langle 0 \rangle \mid i < V\}.$$

We define $T_{n+1}(\xi_{e^t} {}^\frown \langle 0 \rangle) = \upsilon''_{0,0}$; for $i < U$, $T_{n+1}(\eta^i_{e^u_i} {}^\frown \langle 0 \rangle) = \upsilon''_{i+1,0}$; $T_{n+1}(\eta^U_{e^u_U} {}^\frown \langle 0 \rangle) = \nu''_{0,0}$; for $i < V$, $T_{n+1}(\theta^i_{e^v_i} {}^\frown \langle 0 \rangle) = \nu''_{i+1,0}$.

The conclusion of all this is that in T_{n+1}, we have a large open loop which consists exactly of

$$\begin{aligned} S = & \bigcup_{i \leq t} [\tau_i] \\ & \cup \bigcup_{i \leq e^t} [\xi_i {}^\frown \langle 0 \rangle] \\ & \cup \bigcup_{i \leq U} \left[\bigcup_{j \leq u''_i} [\upsilon''_{i,j}] \cup \bigcup_{j \leq e^u_i} [\eta^i_j {}^\frown \langle 0 \rangle] \right] \\ & \cup \bigcup_{i \leq V} \left[\bigcup_{j \leq v''_i} [\nu''_{i,j}] \cup \bigcup_{j \leq e^v_i} [\theta^i_j {}^\frown \langle 0 \rangle] \right]. \end{aligned}$$

If σ is not in S then $T_{n+1}(\sigma) = T_n(\sigma)$.

For $k' \neq k$, we set $A^{n+1}_{k'} = A^n_{k'}$, $B^{n+1}_{k'} = B^n_{k'}$. We set

$$A^{n+1}_k = A^n_k \setminus S$$

and

$$B^{n+1}_k = B^n_k \setminus S.$$

For each $k' > 0$, we let

$$W^{n+1}_{k'+1} = (W^n_{k'+1} \setminus S) \cup (W^n_{k'} \cap S)$$

and

$$W^{n+1}_0 = (W^n_0 \setminus S) \cup (A^n_k \cap S) \cup (B^n_k \cap S).$$

We need to define ρ^{n+1}. If $\sigma \in W^{n+1} \setminus S$ is determined then set $\rho^{n+1}(\sigma) = \rho^n(\sigma)$. For each k' such that $W^{n+1}_{k'} \not\subseteq W^n_{k'}$, choose some $j_{k'}$ not in the image of ρ^n so that $2^{-j_{k'}}$ is very small relative to $\lambda(A^n_{k'}) - a^n_{k'}$ and $\lambda(B^n_{k'}) - b^n_{k'}$. For each determined $\sigma \in S \cap W^{n+1}_{k'}$, set $\rho^{n+1}(\sigma) = j_{k'}$. Let L be the length of S. For each k', if $W^{n+1}_{k'} \not\subseteq W^n_{k'}$, set $\delta_{k'} = L \cdot 2^{-j_{k'}}$, and otherwise set $\delta_{k'} = 0$. For $k \neq k'$, define

$$a^{n+1}_{k'} = a^n_{k'} + \delta_{k'}, \ b^{n+1}_{k'} = b^n_{k'} + 4\delta_{k'}.$$

Define

$$a^{n+1}_k = a^n_k - L_{T_n}(\tau)\lambda([\tau]) + \delta_k$$

and

$$b^{n+1}_k = b^n_k - 4L_{T_n}(\tau)\lambda([\tau]) + 4\delta_k.$$

Even stages—creating upcrossings; the verification: We now check that T_{n+1} satisfies the inductive conditions.

For (1.a), T_n^7 is useful since it was produced by a series of applications of Lemma 3.4, Lemma 3.2, and Lemma 3.6. T_{n+1} is proper by construction. It is partitioned into open loops since either $\sigma \notin S$, in which case σ belongs to the same open loop it did in T_n^7, or $\sigma \in S$, in which case S is the open loop containing σ. We will check escapability below, when we check conditions (1.j)-(1.l).

(1.d) is immediate—each open loop in T_{n+1} is either an open loop from T_n which remains in the same component, or the open loop S, which is in W^{n+1}. (1.e) is immediate from the inductive hypothesis and the construction. If σ is componential and determined in T_{n+1} then either σ was in W^0, so also in W^{n+1}, or in some A^0_k or B^0_k. If σ is not in A^{n+1}_k or B^{n+1}_k, respectively, it must be because some portion of $[\sigma]$ is in S, and so was moved to W^{n+1}. But S is a union of intervals in $T_{n+1,+} \cup T_{n+1,-}$, so if σ is determined and $[\sigma] \cap S \neq \emptyset$ then $[\sigma] \subseteq S$, so σ is in W^{n+1}. In either case, σ belongs to a single component, showing (1.f). (1.g)-(1.i) are immediate from the definition.

For $k' \neq k$, each $T^{i,j}_n$ and T^i_n is $A^n_{k'}, A^n_{k'}$- and $B^n_{k'}, B^n_{k'}$-escapable, and T_{n+1} is as well, since this property is preserved by each step of the construction. T_n was A^n_k, A^n_k-escapable, and therefore A^{n+1}_k, A^n_k-escapable (since $A^{n+1}_k \subseteq A^n_k$), and Lemma 3.4 ensures that $T^{1,0}_n$ is A^{n+1}_k, A^{n+1}_k-escapable. This is preserved by each remaining step, so T_{n+1} is A^{n+1}_k, A^{n+1}_k-escapable. A similar argument shows that $T^{3,0}_n$ is B^{n+1}_k, B^{n+1}_k-escapable, and so T_{n+1} is as well. This shows (1.j) and (1.k).

T^7_n is W^n, W^n-escapable since T_n was and this property is preserved by each step. If $\sigma \in W^{n+1} \setminus W^n$ with $|T_{n+1}(\sigma)| < |\sigma|$ then $\sigma \sqsupseteq \theta^V_{e^v_V} {}^\frown \langle 0 \rangle$, and therefore $T_{n+1}(\sigma) = \langle\rangle$. Otherwise $\sigma \in W^n$, so let $\sigma_0, \ldots, \sigma_r$ be an escape sequence in T^7_n in W^n for σ; we may assume $|\sigma_1| \geq |\tau'_0|$. If no $\sigma_i \in S$ then this is also an escape sequence in T_{n+1} in W^{n+1}. If some $\sigma_i \in S$ then by Lemma 2.9 there is an i with $\sigma_i \sqsupseteq \xi_{e^t}$ and therefore for some ρ

$$\sigma_0, \ldots, \sigma_i, \upsilon''_{0,0} {}^\frown \rho, \ldots, \nu''_{0,0} {}^\frown \rho, \ldots, \theta^V_{e^v_V} {}^\frown \langle 0 \rangle {}^\frown \rho$$

is an escape sequence in W^{n+1}. This shows (1.l).

For (2.a), consider some determined $\sigma \in W^{n+1}_{k'+1}$. If σ is not in S then $\sigma \in W^n_{k'+1}$ and the claim follows since it was true in T_n. The interesting case is when σ is in S; we first consider some points about the structure of the open loop S, which we may write $\zeta_0, \ldots, \zeta_z$. It is natural to divide S into three pieces, $S \cap W^n$, $S \cap A^n_k$, and $S \cap B^n_k$. There are $z_0 < z_1$ so that

$$\bigcup_{i \leq z_0} [\zeta_i] = S \cap W^n, \quad \bigcup_{z_0 < i \leq z_1} [\zeta_i] = S \cap A^n_k, \quad \bigcup_{z_1 < i \leq z} [\zeta_i] = S \cap B^n_k.$$

Furthermore, we have $\lambda(S \cap W^n) = L_{T_n}(\tau)\lambda(\tau) + e^t 2^{-N+N'}$. On the other hand $\lambda(S \cap A^n_k) = (1 - 2^{-N})\lambda(\bigcup_i [\upsilon_i]) + 2^{-N+N'} \sum_{i \leq U} e^i_u$. By choosing N small enough, we ensured that the first term was larger than $\lambda(S \cap W^n)$, so $\lambda(S \cap A^n_k) > \lambda(S \cap W^n)$. In particular, this means $z_1 > 2z_0$. Finally $\lambda(S \cap B^n_k) > (1-2^{-N})\lambda(\bigcup_i [\nu_i]) > 4(1-2^{-N})\lambda(\bigcup_i [\nu_i])$, so again by choosing N and N' large enough, we ensured that $\lambda(S \cap B^n_k) > 4\lambda(S \cap A^n_k)$. In particular, this means $z > 4(z_1 - z_0) > 4z_0$. We'll write $\alpha = \{i \mid [\zeta_i] \subseteq 2^\omega \setminus A\}$; in particular $(z_1, z] \subseteq \alpha$ and $(z_0, z_1] \cap \alpha = \emptyset$.

Now consider some σ in $S \cap W^{n+1}_{k'+1}$. Then $\sigma \in W^n_{k'}$, so $\sigma \in [\zeta_s]$ for some $s \leq z_0$. So it suffices to show that we add an upcrossing to ζ_s. In T_n, there was an upcrossing sequence for ζ_s of length k'—say, $0 \leq u_1 < v_1 < \cdots < u_{k'} < v_{k'}$. Then $0 \leq u_1 < v_1 < \cdots < u_{k'} < v_{k'}$ is an upcrossing sequence for ζ_s in T_{n+1}. We claim that $0 \leq u_1 < v_1 < \cdots < u_{k'} < v_{k'} < z_1 - s < z - s$ is an upcrossing sequence in T_{n+1} as well. This is because

$$\frac{1}{z_1 - s} \sum_{j=s}^{z_1} \chi_\alpha(j) \leq \frac{z_0 - s}{z_1 - s} \leq \frac{z_0}{z_1} < 1/2$$

while

$$\frac{1}{z - s} \sum_{j=s}^{z} \chi_\alpha(j) \geq \frac{z - z_1}{z - s} \geq \frac{z - z_0}{z} > 3/4.$$

(2.b)-(2.e.i) are immediate from the definition. For any $\eta \notin S$ determined in T^7_n, $L_{T^7_n}(\eta) = L_{T_n}(\eta)$ since this is preserved by each step in the construction of T^7_n. The passage from T^7_n to T_{n+1} only affects the loop S, so $L_{T_{n+1}}(\eta) = L_{T^7_n}(\eta) = L_{T_n}(\eta)$, giving (2.e.ii).

For any k' we have $a^n_{k'} < \lambda(A^n_{k'})$ and $A^{n+1}_{k'} = A^n_{k'} \setminus (S \cap A^n_{k'})$ (where $S \cap A^n_{k'} = \emptyset$ unless $k = k'$), so also $a^n_{k'} - \lambda(A^n_{k'}) < \lambda(A^{n+1}_{k'})$. The same holds for $b^n_{k'}$. Since we could choose the values $j_{k'}$ arbitrarily small, we can make them small enough that (2.f) holds.

We turn to (2.g.i). For $k' \neq k$, if there is a $\eta \in W^{n+1}_{k'} \setminus S$ then there was an $\eta' \in W^n_{k'}$ with $\eta' \sqsubseteq \eta$, $\rho^{n+1}(\eta) = \rho^n(\eta)$, and $L_{T_{n+1}}(\eta) = L_{T_n}(\eta)$. (Recall that L is the length of S.) Then

$$\begin{aligned}
\sum_{j \in J^{n+1}_k} l_{k,j}\left(2^{-j} - \lambda([V_{j,n'}]\right) &= \sum_{j \in J^{n+1}_k \setminus \{j_{k'}\}} l_{k,j}\left(2^{-j} - \lambda([V_{j,n'}]\right) \\
&\quad + l_{k',j_{k'}}\left(2^{-j_{k'}} - \lambda([V_{j_{k'},n'}]\right) \\
&\leq \sum_{j \in J^n_k} l_{k,j}\left(2^{-j} - \lambda([V_{j,n'-1}]\right) \\
&\quad + l_{k',j_{k'}}\left(2^{-j_{k'}} - \lambda([V_{j_{k'},n'}]\right) \\
&\leq a^n_{k'} + L2^{-j_{k'}} \\
&= a^{n+1}_{k'}.
\end{aligned}$$

For k we have to take into account that we have reduced a_k^n by $L_{T_n}(\tau)\lambda([\tau])$, and that this is compensated for by the fact that $\tau \in V_{j,n'}$:

$$\begin{aligned}\sum_{j\in J_k^{n+1}} l_{k,j}\left(2^{-j}-\lambda([V_{j,n'}]\right) \leq & \sum_{j\in J_k^{n+1}\setminus\{j_k\}} l_{k,j}\left(2^{-j}-\lambda([V_{j,n'-1}]\right)\\ & - l_{k,\hat{j}}\lambda([V_{\hat{j},n'}])\\ & + l_{k,j_k}\left(2^{-j_k}-\lambda([V_{j_k,n'}]\right)\\ \leq & \sum_{j\in J_k^n} l_{k,j}\left(2^{-j}-\lambda([V_{j,n'-1}]\right)\\ & - L_{T_n}(\tau)\lambda([\tau])\\ & + l_{k,j_k}\left(2^{-j_k}-\lambda([V_{j_k,n'}]\right)\\ \leq & \, a_k^n - L_{T_n}(\tau)\lambda([\tau]) + L2^{-j_k}\\ = & \, a_k^{n+1}.\end{aligned}$$

The argument for (2.g.ii) is identical.

Since n is odd, an integer $i < n/2$ iff $i < (n+1)/2$, so (3.a)-(3.d) follow immediately from the inductive hypothesis.

Even stages—creating upcrossings; undetermined elements: Recall that we enumerated π into V_j and consider the decomposition of π into determined elements π_i. Our last difficulty is ensuring (2.h).

We first observe that if π is determined, so $\tau = \pi$ in the construction above, then we have satisfied (2.h). First, in the $\rho^n(\pi) \neq \hat{j}$ case, no changes were made and (2.h) followed immediately from the inductive hypothesis. When $\rho^n(\pi) = \hat{j}$, if σ is not in S then $\rho^{n+1}(\sigma) = \rho^n(\sigma)$ and σ avoids π, so

$$[V_{\rho^{n+1}(\sigma),n/2}] \cap [\sigma] = [V_{\rho^n(\sigma),n/2}] \cap [\sigma] = [V_{\rho^n(\sigma),(n-1)/2}] \cap [\sigma] = \emptyset$$

since π was the only element in $\bigcup_j V_{j,(n-1)/2} \setminus \bigcup_j V_{j,(n-1)/2-1}$. For any element in S, $\rho^{n+1}(\sigma) \neq \rho^n(\sigma)$, and we chose a fresh value for $\rho^{n+1}(\sigma)$, so we may assume we chose $V_{\rho^{n+1}(\sigma),n/2}$ to be empty.

We now consider the general case where π is not determined. We have a partition $[\pi] = \bigcup_{i\leq r}[\pi_i]$ where each π_i is determined. We let $U_0 = \{i \mid \rho^n(\pi_i) = \hat{j}\}$. If $U_0 = \emptyset$ then we may simply take $T_{n+1} = T_n$, so assume U_0 is non-empty. Pick some $i_0 \in U_0$ such that the final element of the open loop containing π_{i_0} has an escape sequence disjoint from $\bigcup_{i\in U_0}[\pi_i]$. To find such an escape sequence, take any π_i with $i \in U_0$ and take an escape sequence $\upsilon_0, \dots, \upsilon_k$ for the final element of the open loop containing π_i; if this is not already such an escape sequence, let $d \leq k$ be greatest such that $\upsilon_d \sqsupseteq \tau_j$ for some $j \in U$, and let $d' \geq d$ be such that $\upsilon_{d'}$ is contained in the final element of the open loop containing π_j. Then $\upsilon_{d'+1}, \dots, \upsilon_k$ is an escape sequence for the final element of the open loop containing π_j which, since d was chosen greatest, is disjoint from $\bigcup_{i\in U_0}[\pi_i]$.

We may apply the construction above with $\tau = \pi_{i_0}$, giving a transformation $T_n^{8,0}$ (the transformation that we referred to as T_{n+1} above). For each $i \in U_0$, one of two things happens: either π_i is in the open loop containing π_{i_0} in T_n, in which case π_i may not be determined in $T_n^{8,0}$, but any determined σ in $[\pi_i]$ satisfies $\rho_n^{8,0}(\sigma) \neq \hat{\jmath}$; or π_i is not in the open loop containing π_{i_0}, in which case $[\pi_i]$ is disjoint from the escape sequence used in the construction of $T_n^{8,0}$, and so also disjoint from the open loop S, so π_i is still determined in $T_n^{8,0}$. Let $U_1 \subseteq U_0$ be the set of $i \in U_0$ such that π_i is not in the open loop containing π_{i_0} in T_n. Clearly $i_0 \in U_0 \setminus U_1$.

If $U_1 \neq \emptyset$, we repeat this argument, choosing an $i_1 \in U_1$ such that the final element of the open loop containing π_{i_1} has an escape sequence disjoint from $\bigcup_{i \in U_1}[\pi_i]$, and we apply the construction above to give $T_n^{8,1}$. We repeat this argument $r' \leq r$ times, giving $T_n^{8,r'-1}$ so that $U_{r'} = \emptyset$, and therefore every determined σ in $[\tau]$ satisfies $\rho_n^{8,r'-1}(\sigma) \neq \hat{\jmath}$. Then we may set $T_{n+1} = T_n^{8,r'-1}$. Each $T_n^{8,i}$ satisfies all inductive clauses except for (2.h), and further satisfies (2.h) for all σ avoiding π and also for all π_j with $j \notin U_i$. Since $U_{r'-1} = \emptyset$, T_{n+1} at last satisfies (2.h).

Checking the construction: We have completed the inductive construction. We now show that $\hat{T}$ is the desired transformation. Suppose $x \in \bigcap_j [V_j]$; we must show that $\hat{T}$ has infinitely many upcrossings on x. It suffices to show that $\hat{T}$ has at least k upcrossings for every k. Suppose not. Since $x \in W^0$ and each element of W_k^n has at least k upcrossings, let k be largest such that $x \in W_k^n$ for some k, and pick some large enough n so that $x \in W_k^n$. Let $\sigma \in T_{n,+} \cup T_{n,-}$ be such that $x \in [\sigma]$. Let $j = \rho^n(\sigma)$. Since $x \in \bigcap_j [V_j]$ but $[\sigma] \cap [V_{j,(n-1)/2}] = \emptyset$, there is some $m > n$ such that $x \in [V_{j,(m-1)/2}]$. Let $\tau \in T_{m,+} \cup T_{m,-}$ be such that $x \in [\tau]$. Since $x \in W_k^m$, $\rho^m(\tau) = \rho^n(\sigma)$ by (2.e.i), but this contradicts (2.h).

To see that $|\hat{T}(x)|$ is infinite except on an effective F_σ set of measure 0, we claim that if $|\hat{T}(x)|$ is finite then is an i_0 such that for every $i > i_0$, $x \in G_i$. Since $\lambda(\bigcap_{i > i_0} G_i) = 0$, we can also see that $\lambda(\bigcup_{i_0} \bigcap_{i > i_0} G_i) = 0$, so this shows that the set of such x has measure 0. Let $|\hat{T}(x)| = k$ be finite. Choose some n and some componential $\sigma \in T_{n,-}$ with $x \in [\sigma]$ and $|T_n(\sigma)| = k$. For each $m > n$, let $\sigma_m \in T_{m,-}$ be such that $x \in [\sigma_m]$. Since $|T_m(\sigma_m)| = k$ for all $m > n$, in particular at each odd m we must have $[\sigma_m] \subseteq G_{(m-1)/2}$. Therefore for all $i > (n-1)/2$, we have $x \in G_i$, as desired. □

5. The Ergodic Case

As promised above, we now present the strengthening of Gács, Hoyrup and Rojas' result to show that every Schnorr random point is Birkhoff for computable, bounded functions with computable ergodic transformations.

Theorem 5.1. *Let f be a bounded computable function and suppose T is an ergodic, computable, measure-preserving transformation. Then every Schnorr random point is Birkhoff for f.*

Proof. It is convenient to work with a *strong Borel-Cantelli* test: an effectively c.e. sequence $\langle V_i \rangle$ is a strong Borel-Cantelli test if $\sum_i \lambda([V_i]) < \infty$ is a computable real number. It is known [13] that x is Schnorr random iff for every strong Borel-Cantelli test $\langle V_i \rangle$, x is in only finitely many V_i.

We will write $A_n(x)$ for the function $\frac{1}{n}\sum_{i<n} f(T^i x)$. Our main tool is Theorem 5.3 of [1]:

> Let T and f be computable, let $f^*(x) = \lim_{n\to\infty} A_n(x)$, and suppose that $||f^*||_{L^2}$ is computable. Then for each $\epsilon > 0$, there is an $N(\epsilon)$, computable from $T, f, ||f^*||_{L^2}, \epsilon$ such that
> $$\mu(\{x \mid \max_{N(\epsilon)\le m} |A_m(x) - A_n(x)| > \epsilon\}) \le \epsilon.$$

Since T is ergodic, f^* is the function constantly equal, almost everywhere, to $c = \int f\, d\lambda$, so $||f^*||_{L^2}$ is just c, which, since f is bounded and computable, is itself a computable real number. Define a sequence $n_i = N(2^{-i-2})$. Then for each i we take

$$V_i = \{x \mid \max_{n_i \le m \le n_{i+1}} |c - A_m(x)| > 2^{-i}\}.$$

If $x \in [V_i]$ then either x is not a Birkhoff point—a set measure 0 by the Birkhoff ergodic theorem—or there is an $m' \ge n_i = N(2^{-i})$ with $|c - A_{m'}(x)| < 2^{-i-1}$. In the latter case we have $|A_m(x) - A_{m'}(x)| \ge 2^{-i-1}$ and therefore either $|A_n(x) - A_m(x)| \ge 2^{-i-2}$ or $|A_n(x) - A_{m'}(x)| \ge 2^{-i-2}$, and so $\max_{n_i \le m} |A_m(x) - A_n(x)| > 2^{-i-2}$. Therefore $\lambda([V_i]) \le 2^{-i-2}$.

Since the V_i are computable sets and $\lim_{j\to\infty} \sum_{i>j} \lambda([V_i]) = 0$ with a computable rate of convergence, $\sum_i \lambda([V_i])$ is a computable real number. Since $\sum_i \lambda([V_i]) \le \sum_i 2^{-i-2} = 1/2$, $\langle V_i \rangle$ is a strong Borel-Cantelli test.

If x is not a Birkhoff point then for some $\delta > 0$ there are infinitely many m with $|c - A_m(x)| > \epsilon$. Fix i with $2^{-i} < \delta$, and there are infinitely many $j > i$ with $x \in V_j$, so x fails to be Schnorr random. □

6. Upcrossings

Throughout this section, we will take T to be a computable, measure-preserving transformation.

Recall the following theorem of Bishop [5]:

Theorem 6.1.
$$\int \tau(x, f, \alpha, \beta)dx \le \frac{1}{\beta - \alpha} \int (f - \alpha)^+ dx.$$

This is easily used to derive the following special case of a theorem of V'yugin:

Theorem 6.2 ([30])**.** *If x is Martin-Löf random and f is computable then $\lim_{n\to\infty} \frac{1}{n+1} \sum_{j=0}^{n} f(T^j x)$ converges.*

Proof. Suppose $\lim_{n\to\infty}\frac{1}{n+1}\sum_{j=0}^{n} f(T^j x)$ does not converge. Then there exist $\alpha < \beta$ such that $\frac{1}{n+1}\sum_{j=0}^{n} f(T^j x)$ is infinitely often less than α and also infinitely often greater than β. Equivalently, $\tau(x, f, \alpha, \beta)$ is infinite. But observe that when f is computable, $\tau(x, f, \alpha, \beta)$ is lower semi-computable, so in particular,

$$V_n = \{x \mid \tau(x, f, \alpha, \beta) \geq n\}$$

is computably enumerable and $\mu(V_n) \leq \frac{1}{n(\beta-\alpha)}\int (f-\alpha)^+ dx$. Therefore an appropriate subsequence of $\langle V_n\rangle$ provides a Martin-Löf test, and $x \in \cap_n V_n$, so x is not Martin-Löf random. □

We now consider the case where f is lower semi-computable. We will have a sequence of uniformly computable increasing approximations $f_i \to f$, and we wish to bound the number of upcrossings in f. The difficulty is that $\tau(x, f_i, \alpha, \beta)$ is not monotonic in i: it might be that an upcrossing sequence for f_i ceases to be an upcrossing sequence for f_{i+1}.

In order to control this change, we need a suitable generalization of upcrossings, where we consider not only the upcrossings for f, but for all functions between f and $f + h$ where h is assumed to be small.

Definition 6.3. A *loose upcrossing sequence* at x for α, β, f, h is a sequence

$$0 \leq u_1 < v_1 < u_2 < v_2 < \cdots < u_N < v_N$$

such that for all $i \leq N$,

$$\frac{1}{u_i+1}\sum_{j=0}^{u_i} f(T^j x) < \alpha,\ \frac{1}{v_i+1}\sum_{j=0}^{v_i}(f+h)(T^j x) > \beta.$$

$\upsilon(x, f, h, \alpha, \beta)$ is the supremum of the lengths of loose upcrossing sequences for α, β, f, h.

Loose upcrossings are much more general than we really need, and so the analog of Bishop's theorem is correspondingly weak. For instance, consider the case where T is the identity transformation, $f = \chi_A$, and $h = \chi_B$ with A and B disjoint (so $f + h = \chi_{A\cup B}$). Then $\upsilon(x, f, h, \alpha, \beta) = \infty$ for every $x \in B$ whenever $0 < \alpha < \beta < 1$. Nonetheless, we are able to show the following:

Theorem 6.4. *Suppose $h \geq 0$, $\int h\, dx < \epsilon$ and $\beta - \alpha > \delta$. There is a set A with $\mu(A) < 4\epsilon/\delta$ such that*

$$\int_{X\setminus A} \upsilon(x, f, h, \alpha, \beta) dx$$

is finite.

Proof. By the usual pointwise ergodic theorem, there is an n and a set A' with $\mu(A') < 2\epsilon/\delta$ such that if $x \notin A'$ then for all $n', n'' \geq n$,

$$\left|\frac{1}{n'+1}\sum_{j=0}^{n'} h(T^j x) - \frac{1}{n''+1}\sum_{j=0}^{n''} h(T^j x)\right| < \delta/2.$$

Consider those $x \notin A'$ such that, for some $n' \geq n$,

$$\frac{1}{n'+1}\sum_{j=0}^{n'} h(T^j x) \geq \delta.$$

We call this set A''. Then for all $n' \geq n$, such an x satisfies

$$\frac{1}{n'+1}\sum_{j=0}^{n'} h(T^j x) \geq \delta/2,$$

and in particular,

$$\int_{A''} h\ dx \geq \delta\mu(A'')/2.$$

Therefore $\mu(A'') \leq 2\epsilon/\delta$. If we set $A = A' \cup A''$, we have $\mu(A) < 4\epsilon/\delta$.

Now suppose $x \notin A$. We claim that any loose upcrossing sequence for α, β, f, h with $n \leq u_1$ is already an upcrossing sequence for $\alpha, \beta - \delta$. If $n \leq u_1 < v_1 < \cdots < u_N < v_N$ is a loose upcrossing sequence, we automatically satisfy the condition on the u_i. For any v_i, we have

$$\beta < \frac{1}{v_i+1}\sum_{j=0}^{v_i}(f+h)(T^j x) = \frac{1}{v_i+1}\sum_{j=0}^{v_i} f(T^j x) + \frac{1}{v_i+1}\sum_{j=0}^{v_i} h(T^j x).$$

Since $\frac{1}{v_i+1}\sum_{j=0}^{v_i} h(T^j x) \leq \delta$, it follows that $\frac{1}{v_i+1}\sum_{j=0}^{v_i} f(T^j x) > \beta - \delta$ as desired. Therefore

$$\int_{X\setminus A} \upsilon(x, f, h, \alpha, \beta)dx \leq \mu(X \setminus A)\int_{X\setminus A} n + \tau(x, f, \alpha, \beta - \delta)dx$$

is bounded. □

Theorem 6.5. *If x is weakly 2-random and f is lower semi-computable then $\lim_{n\to\infty} \frac{1}{n+1}\sum_{j=0}^{n} f(T^j x)$ converges.*

Proof. Suppose $\lim_{n\to\infty} \frac{1}{n+1}\sum_{j=0}^{n} f(T^j x)$ does not converge. Then there exist $\alpha < \beta$ such that $\frac{1}{n+1}\sum_{j=0}^{n} f(T^j x)$ is infinitely often less than α and also infinitely often greater than β. Equivalently, $\tau(x, f, \alpha, \beta)$ is infinite. Let $f_n \to f$ be the sequence of computable functions approximating f from below.

For each n, we set

$$V_n = \{x \mid \exists m \geq n\ \upsilon(x, f_n, f_m - f_n, \alpha, \beta) \geq n\}.$$

By construction, $x \in \cap_n V_n$. To see that $V_{n+1} \subseteq V_n$, observe that if

$$\upsilon(x, f_{n+1}, f_m - f_{n+1}, \alpha, \beta) \geq n + 1$$

then there is a loose upcrossing sequence witnessing this, and it is easy to check (since the f_n are increasing) that this is also a loose upcrossing sequence witnessing

$$\upsilon(x, f_n, f_m - f_n, \alpha, \beta) \geq n + 1 > n.$$

We must show that $\mu(V_n) \to 0$. Fix $\delta < \beta - \alpha$ and let $\epsilon > 0$ be given. Choose n to be sufficiently large that $||f_n - f|| < \delta\epsilon/4$. Then, since the f_n approximate f from below, clearly $\upsilon(x, f_m, f_{m'} - f_m, \alpha, \beta) \leq \upsilon(x, f_m, f - f_m, \alpha, \beta)$ for any $m' \geq m$. By the previous theorem, there is a set A with $\mu(A) < \epsilon/2$ such that $\int_{X \setminus A} \upsilon(x, f_m, f - f_m, \alpha, \beta) dx$ is bounded. We may choose $n' \geq n$ sufficiently large that

$$B = \mu(\{x \notin A \mid \upsilon(x, f_m, f - f_m, \alpha, \beta) \geq n'\}) < \epsilon/2.$$

Then $V_{n'} \subseteq A \cup B$, so $\mu(V_{n'}) \leq \epsilon$. □

6.1. **Room for Improvement.** It is tempting to try to improve Theorem 6.4. The premises of that theorem are too general and the proof is oddly "half-constructive"—we mix the constructive and nonconstructive pointwise ergodic theorems. One would think that by tightening the assumptions and using Bishop's upcrossing version of the ergodic theorem in both places, we could prove something stronger.

In the next theorem, we describe an improved upcrossing property which, if provable, would lead to a substantial improvement to Theorem 6.5: balanced randomness would guarantee the existence of this limit. (Recall that a real is *balanced random* if it passes every balanced test, or sequence $\langle V_i \rangle$ of r.e. sets such that for every i, $\mu([V_i]) \leq 2^{-i}$ and $V_i = W_{f(i)}$ for some function f with a recursive approximation that has at most 2^n mind-changes for the value of $f(n)$ [10].) The property hypothesized seems implausibly strong, but we do not see an obvious route to ruling it out.

Theorem 6.6. *Suppose the following holds:*

> *Let f and $\epsilon > 0$ be given, and let $0 \leq h_0 \leq h_1 \leq \cdots \leq h_n$ be given with $||h_n||_{L^\infty} < \epsilon$. Then*
>
> $$\int_X \sup_n \tau(x, f + h_n, \alpha, \beta) dx < c(||f||_{L^\infty}, \epsilon)$$
>
> *where $c(||f||_{L^\infty}, \epsilon)$ is a computable bound depending only on $||f||_{L^\infty}$ and ϵ.*

Then whenever x is balanced random and f is lower semi-computable then $\lim_{n\to\infty} \frac{1}{n+1} \sum_{j=0}^{n} f(T^j x)$ converges.

Proof. We assume $||f||_{L^2} \leq 1$ (if not, we obtain this by scaling). Suppose $\lim_{n\to\infty} \frac{1}{n+1} \sum_{j=0}^{n} f(T^j x)$ does not converge. Then there exist $\alpha < \beta$ such that $\frac{1}{n+1} \sum_{j=0}^{n} f(T^j x)$ is infinitely often less than α and also infinitely often greater than β. Equivalently, $\tau(x, f, \alpha, \beta)$ is infinite. Let $f_n \to f$ be the sequence of computable functions approximating f from below.

We define the set

$$V_{(n,k)} = \{x \mid \exists m \geq n \ \tau(x, f_m, \alpha, \beta) \geq k\}.$$

We then define the function $g(n, n')$ to be least such that $\forall m \in [n, n'] \ ||f_{n'} - f_m|| < 2^{-n}$ and $g(n) = \lim_{n'} g(n, n')$. Since the sequence f_m converges to f from below, $g(n)$ is defined everywhere, and $|\{s \mid g(n, s+1) \neq g(n, s)\}| < 2^n$

for all n. Indeed, $g(n)$ is the least number such that $\forall m \geq g(n)$ $||f - f_m|| \leq 2^{-n}$.

Observe that $\mu(V_{(n,k)}) < \frac{c(||f||_{L^\infty}, 2^{-n})}{k}$. Choose $h(n)$ to be a computable function growing quickly enough that $\frac{c(||f||_{L^\infty}, 2^{-n})}{h(n)} \leq 2^{-n}$ for all n. If $x \in V_{(g(n+1),h(n+1))}$ then there is some $m \geq g(n+1)$ so that $\tau(x, f_m, \alpha, \beta) \geq h(n+1)$. Since $g(n+1) \geq g(n)$ and $h(n+1) \geq h(n)$, we also have that $x \in V_{(g(n),h(n))}$. Therefore $\langle V_{(g(n),h(n))} \rangle$ is a balanced test.

But since $\tau(x, f, \alpha, \beta)$ is infinite, we must have $x \in \cap V_{(g(n),h(n))}$. This contradicts the assumption that x is balanced random, so $\lim_{n\to\infty} \frac{1}{n+1} \sum_{j=0}^{n} f(T^j x)$ converges. □

In fact, the test $\langle V_{(g(n),h(n))} \rangle$ has an additional property: if $s_0 < s_1 < s_2$ with $g(n+1, s_0) \neq g(n+1, s_1) \neq g(n+1, s_2)$ then $g(n, s_0) \neq g(n, s_2)$. This means that $\langle V_{(g(n),h(n))} \rangle$ is actually an *Oberwolfach test* [3], and so we can weaken the assumption to x being Oberwolfach random.

7. Discussion for Ergodic Theorists

In the context of analytic questions like the ergodic theorem, matters of computability are mostly questions of continuity and uniformity: the computability of a given property usually turns on whether it depends in an appropriately uniform way on the inputs. Algorithmic randomness gives a precise way of characterizing how sensitive the ergodic theorem is to small changes in the underlying function.

The paradigm is to distinguish different sets of measure 0, viewed as an intersection $A = \cap_i A_i$, by characterizing how the sets A_i depend on the given data (in the case of the ergodic theorem, the function f). The two main types of algorithmic randomness that have been studied in this context thus far are Martin-Löf randomness and Schnorr randomness. In both cases, we ask that the sets A_i be unions $A_i = \cup_j A_{i,j}$ of sets where $A_{i,j}$ is determined based on a finite amount of information about the orbit of f (in particular, the dependence of $A_{i,j}$ on f and T should be continuous). (To put it another way, we ask that the set of exceptional points which violate the conclusion of the ergodic theorem be contained in a G_δ-set which depends on f in a uniform way.) The distinction between the two notions is that in Schnorr randomness, $\mu(A_i) = 2^{-i}$, while in Martin-Löf randomness, we only know $\mu(A_i) \leq 2^{-i}$. This means that in the Schnorr random case, a finite amount of information about the orbit of f suffices to limit the density of A_i outside of a small set (take J large enough that $\mu(\cup_{j\leq J} A_{i,j})$ is within ϵ of 2^{-i}; then no set disjoint from $\cup_{i\leq J} A_{i,j}$ contains more than ϵ of A_i). In the Martin-Löf random case, this is not possible: if $\mu(A_i) \leq 2^{-i} - \epsilon$, no finite amount of information about the orbit of f can rule out the possibility that some $A_{i,j}$ with very large j will add a set of new points of measure ϵ. In particular, while we can identify sets which do belong to A_i, finite information about the orbit of f does not tell us much about which points are not in A_i.

The two classes of functions discussed in this paper are the computable and the lower semi-computable ones; these are closely analogous to the continuous and lower semi-continuous functions. Unsurprisingly, both the passage from computable to lower semi-computable functions and the passage from ergodic to nonergodic transformations make it harder to finitely characterize points violating the conclusion of the ergodic theorem. Perhaps more surprising, both changes generate precisely the same result: if a point violates the conclusion of the ergodic theorem for a computable function with a nonergodic transformation, we can construct a lower semi-computable function with an ergodic transformation for which the point violates the conclusion of the ergodic theorem, and vice versa.

The remaining question is this: What happens when we make both changes? What characterizes the points which violate the conclusion of the ergodic theorem for lower-semi computable functions with nonergodic transformations? The answer is likely to turn on purely ergodic theoretic questions about the sensitivity of upcrossings, such as the hypothesis we use above.

Question 7.1. *Let (X,μ) be a metric space and let $T : X \to X$ be measure preserving. Let $\epsilon > 0$ be given. Is there a bound K (depending on T and on ϵ) such that for any f with $||f||_{L^\infty} \leq 1$ and any sequence $0 \leq h_0 \leq h_1 \leq \cdots \leq h_n$ with $||h_n||_{L^\infty} < \epsilon$,*

$$\int \sup_n \tau(x, f + h_n, \alpha, \beta)dx < K$$

where $\tau(x, g, \alpha, \beta)$ is the number of upcrossings from below α to above β starting with the point x?

References

[1] Jeremy Avigad, Philipp Gerhardy, and Henry Towsner. Local stability of ergodic averages. *Trans. Amer. Math. Soc.*, 362(1):261–288, 2010.

[2] Laurent Bienvenu, Adam Day, Ilya Mezhirov, and Alexander Shen. Ergodic-type characterizations of Martin-Löf randomness. In *6th Conference on Computability in Europe (CiE 2010)*, volume 6158 of *Lecture Notes in Comput. Sci.*, pages 49–58. Springer, Berlin, 2010.

[3] Laurent Bienvenu, Noam Greenberg, Antonín Kučera, André Nies, and Dan Turetsky. Balancing randomness. In Klaus Ambos-Spies, Rodney G. Downey, Steffen Lempp, and Wolfgang Merkle, editors, *Computability Theory*, number 08/2012 in Oberwolfach Reports. Mathematisches Forschungsinstitut Oberwolfach, 2012.

[4] George D. Birkhoff. Proof of the ergodic theorem. *Proc. Nat. Acad. Sci.*, 17(12):656–660, 1931.

[5] Errett Bishop. A constructive ergodic theorem. *J. Math. Mech.*, 17:631–639, 1967/1968.

[6] R. V. Chacon. Transformations having continuous spectrum. *J. Math. Mech.*, 16:399–415, 1966.

[7] R. V. Chacon. Weakly mixing transformations which are not strongly mixing. *Proc. Amer. Math. Soc.*, 22:559–562, 1969.

[8] Rod Downey, Denis R. Hirschfeldt, André Nies, and Sebastiaan A. Terwijn. Calibrating randomness. *Bull. Symbolic Logic*, 12(3):411–491, 2006.

[9] Rodney G. Downey and Denis R. Hirschfeldt. *Algorithmic Randomness and Complexity.* Springer, 2010.

[10] Santiago Figueira, Denis Hirschfeldt, Joseph S. Miller, Keng Meng Ng, and André Nies. Counting the changes of random Δ^0_2 sets. In *Programs, proofs, processes*, volume 6158 of *Lecture Notes in Comput. Sci.*, pages 162–171. Springer, Berlin, 2010.

[11] Johanna N.Y. Franklin, Noam Greenberg, Joseph S. Miller, and Keng Meng Ng. Martin-Löf random points satisfy Birkhoff's ergodic theorem for effectively closed sets. *Proc. Amer. Math. Soc.*, 140(10):3623–3628, 2012.

[12] Nathaniel A. Friedman. Replication and stacking in ergodic theory. *Amer. Math. Monthly*, 99(1):31–41, 1992.

[13] Peter Gács, Mathieu Hoyrup, and Cristóbal Rojas. Randomness on computable probability spaces—a dynamical point of view. *Theory Comput. Syst.*, 48(3):465–485, 2011.

[14] Paul R. Halmos. *Lectures on Ergodic Theory.* Chelsea Publishing Company, 1956.

[15] Mathieu Hoyrup and Cristóbal Rojas. Computability of probability measures and Martin-Löf randomness over metric spaces. *Inform. and Comput.*, 207(7):830–847, 2009.

[16] Shizuo Kakutani. Examples of ergodic measure preserving transformations which are weakly mixing but not strongly mixing. In *Recent advances in topological dynamics (Proc. Conf., Yale Univ., New Haven, Conn., 1972; in honor of Gustav Arnold Hedlund)*, pages 143–149. Lecture Notes in Math., Vol. 318. Springer, Berlin, 1973.

[17] Anatole Katok and Jean-Paul Thouvenot. Spectral properties and combinatorial constructions in ergodic theory. In *Handbook of dynamical systems. Vol. 1B*, pages 649–743. Elsevier B. V., Amsterdam, 2006.

[18] Antonín Kučera. Measure, Π^0_1-classes and complete extensions of PA. In *Recursion theory week (Oberwolfach, 1984)*, volume 1141 of *Lecture Notes in Math.*, pages 245–259. Springer, Berlin, 1985.

[19] Stuart Alan Kurtz. *Randomness and genericity in the degrees of unsolvability.* PhD thesis, University of Illinois, 1981.

[20] Per Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619, 1966.

[21] André Nies. *Computability and Randomness.* Clarendon Press, Oxford, 2009.

[22] Piergiorgio Odifreddi. *Classical Recursion Theory.* Number 125 in Studies in Logic and the Foundations of Mathematics. North-Holland, 1989.

[23] Piergiorgio Odifreddi. *Classical Recursion Theory, Volume II*. Number 143 in Studies in Logic and the Foundations of Mathematics. North-Holland, 1999.

[24] Henri Poincaré. *Les Méthodes Nouvelles de la Mécanique Céleste*. Dover Publications, Inc., 1957 edition, 1899.

[25] Cristóbal Rojas. Private communication.

[26] C.-P. Schnorr. *Zufälligkeit und Wahrscheinlichkeit*, volume 218 of *Lecture Notes in Mathematics*. Springer-Verlag, Heidelberg, 1971.

[27] Paul C. Shields. Cutting and stacking: a method for constructing stationary processes. *IEEE Trans. Inform. Theory*, 37(6):1605–1617, 1991.

[28] Robert I. Soare. *Recursively Enumerable Sets and Degrees*. Perspectives in Mathematical Logic. Springer-Verlag, 1987.

[29] V. V'yugin. On instability of the ergodic limit theorems with respect to small violations of algorithmic randomness. In *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, pages 1514–1518, 2011.

[30] V. V. V′yugin. Effective convergence in probability, and an ergodic theorem for individual random sequences. *Teor. Veroyatnost. i Primenen.*, 42(1):35–50, 1997.

Department of Mathematics, University of Connecticut U-3009, 196 Auditorium Road, Storrs, CT 06269-3009, USA
E-mail address: johanna.franklin@uconn.edu
URL: www.math.uconn.edu/~franklin

Department of Mathematics, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104-6395, USA
E-mail address: htowsner@math.upenn.edu
URL: http://www.math.upenn.edu/~htowsner