LECTURE NOTES ON ERGODIC THEORY

DAVAR KHOSHNEVISAN

ABSTRACT. Some course notes on ergodic theory. These are for Math 7880-1 ("Topics in Probability"), taught at the Deparment of Mathematics at the University of Utah during the Spring semester of 2005.

They are constantly being updated and corrected. Read them at your own risk.

The authors thanks the National Science Foundation for their continued support.

1. STATIONARY PROCESSES

Stationary processes are one possible generalization of i.i.d. random sequences. They arise in a number of various areas in probability, ergodic theory, analysis, metric number theory, etc.

Definition 1.1. A stochastic process $\{X_n\}_{n\in\mathbf{Z}}$ is *stationary* if for all integers ℓ and $n_1 \le n_2 \le \cdots \le n_k$, the distribution of $(X_{n_1},\ldots,X_{n_k})$ is the same as that of $(X_{\ell+n_1},\ldots,X_{\ell+n_k})$.

Typically, we have a process indexed by $\mathbf{Z}_+ = \{0,1,2\ldots\}$ and not all of $\mathbf{Z}$. In that case, the preceding definition persists, but we insist that ℓ and n_1 are non-negative, so that everything is well defined. Any stationary process X, indexed by $\mathbf{Z}_+$, can usually be extended uniquely to a stationary process indexed by $\mathbf{Z}$. Here is how: Suppose $\mathrm{P}\{X_n \in [0,1]\} = 1$ for all $n \ge 0$. Define μ_k to be the following probability measure on $[0,1]^{\{-k,\ldots,k\}}$:

$$\begin{aligned}&\mu_k\left(A_{-k}\times\cdots\times A_{-1}\times A_0\times A_1\times\cdots\times A_k\right)\\&:= \mathrm{P}\left\{X_0\in A_{-k}, X_1\in A_{-k+1},\ldots,A_{2k}\in A_k\right\},\end{aligned}$$

for all Borel-measurable $A_i \subseteq [0,1]$ $(i\in\mathbf{Z})$. By the stationarity of the X_n's, $\{\mu_k\}_{k=0}^{\infty}$ is a consistent family of probability measures. Thanks to the Kolmogorov extension theorem there is a unique extension μ_∞ to all of $[0,1]^{\mathbf{Z}}$. Define $Y_n(\omega) = \omega_n$ for all $\omega \in [0,1]^{\mathbf{Z}}$ and $n \in \mathbf{Z}$. It follows that for all $k \ge 0$, $(Y_{-k}, Y_{-k+1},\ldots)$ has the same distribution as $(X_0, X_1,\ldots)$, and the entire process Y is now stationary. So you can think of Y as the 'stationary extension' of X to all of $[0,1]^{\mathbf{Z}}$. A similar procedure can be carried out as long as the space S in which the X's live can be embedded in $[0,1]^{\mathbf{Z}}$. Urysohn's metrization theorem of general topology ensures that any locally compact Hausdorff topological space S will do. [This is as large a state space as one could possibly hope to run into.]

I will no longer distinguish between stationary processes indexed by $\mathbf{Z}_+$ and those indexed by $\mathbf{Z}$.

Example 1.2.

(1) All i.i.d. sequences are stationary by default.

(2) If $\{X_n\}_{n\in\mathbf{Z}}$ is i.i.d. and $\{a_n\}_{n\in\mathbf{Z}}$ is non-random, we can define

$$Y_n := \sum_{j\in\mathbf{Z}} a_{n-j}X_j \qquad {}^{\forall} n\in\mathbf{Z},$$

provided that the sum is a.s. convergent.[1] The process Y is then stationary; it is called a *moving average process.*

[1]According to the Kolmogorov two-series theorem, this sum is a.s. convergent if $\mathrm{E}[X_0]=0$, $\mathrm{Var}(X_0)<\infty<$ and $\sum_{j\in\mathbf{Z}}|a_j|^2<\infty$.

2. The Canonical Probability Space

Let $(S, \mathscr{S}, \mathrm{P})$ be a probability space, and define Ω to be the collection of all infinite sequences $(\cdots, \omega_{-1}, \omega_0, \omega_1, \ldots)$ where $\omega_j \in S$. Define $\mathscr{F}$ to be the Borel σ-algebra $\mathscr{S}^{\mathbf{Z}}$, all the time keeping in mind that this is generated by all open sets in $S^{\mathbf{Z}}$, and the latter is endowed with the product topology.

Consider the stochastic process,

$$X_n(\omega) := \omega_n \qquad \forall n \in \mathbf{Z}. \tag{2.1}$$

Now consider the map $T : \Omega \to \Omega$ defined by

$$T : (\ldots, \omega_{-1}, \omega_0, \omega_1, \ldots) \to (\ldots, \omega_0, \omega_1, \omega_2, \ldots) \qquad \forall \omega \in \Omega.$$

That is, $(T\omega)_n := \omega_{n+1}$ for all $n \in \mathbf{Z}$ and $\omega \in \Omega$. Therefore, $T\omega$ is a coordinatewise "shift to the right" of ω. It is invertible, and its inverse is a shift to the left; i.e.,

$$(T^{-1}\omega)_n := \omega_{n-1} \qquad \forall n \in \mathbf{Z}. \tag{2.2}$$

Note that we have

$$X_n(\omega) = \omega_n = X_0(T^n\omega) \qquad \forall \omega \in \Omega,\ n \in \mathbf{Z}, \tag{2.3}$$

where $T^0\omega := \omega$, and $T^n\omega := T(T^{n-1}\omega)$ for all $n \neq 0$.

Definition 2.1. Suppose μ is a probability measure on $(\Omega, \mathscr{F})$. A map $\tau : \Omega \to \Omega$ is said to be *measure preserving* (equivalently, τ preserves μ) if:

1. τ is measurable;
2. τ is one-to-one and onto;
3. The distribution of τ under μ is μ itself; i.e., $\mu\{\tau \in A\} = \mu(A)$ for all $A \in \mathscr{F}$.

Proposition 2.2. *The process X defined by* (2.3) *is stationary if and only if T is measure preserving.*

Proof. We have seen already that T is one-to-one and onto. We even identified the inverse T^{-1}; consult (2.2). Suppose, first, that T is measure preserving. Then,

$$\begin{aligned} \mathrm{P}\left\{X_{n_j} \in A_j \ \forall 1 \le j \le k\right\} &= \mathrm{P}\left\{\left(X_0 \circ T^{n_j}\right) \in A_j \ \forall 1 \le j \le k\right\} \\ &= \mathrm{P}\left\{T^{n_j} \in X_0^{-1} A_j \ \forall 1 \le j \le k\right\} \\ &= \mathrm{P}\left(T^{-\ell}\left\{T^{n_j} \in Z_0^{-1} A_j \ \forall 1 \le j \le k\right\}\right), \end{aligned}$$

because T is measure preserving. Therefore,

$$\mathrm{P}\left\{X_{n_j} \in A_j \ \forall 1 \le j \le k\right\} = \mathrm{P}\left\{T^{\ell+n_j} \in X_0^{-1} A_j \ \forall 1 \le j \le k\right\}.$$

The stationarity of X follows from this. The converse is proved similarly. $\square$

It might help to think back on how one proves that i.i.d. random variables with a given distribution exist. The central portion of the proof, invariably, comes down to showing that when S is reasonable (i.e., $S \simeq [0,1]^{\mathbf{Z}}$), then we can take for our probability space the preceding $(\Omega, \mathscr{F})$, and define the X_n's by (2.1). Therefore,

(2.4)
In all cases of interest, our process can be assumed to have the form (2.1).

This means that the analysis of stationary processes is *equivalent* to the analysis of masure-preserving transformations. From now on, th eprocess X is assumed to be of the form given by (2.3). (Once again, if not, then in all cases of interest we can constrct such an X by possibly having to enlarge the underlying probability space.)

3. THE BASIC ERGODIC THEOREM

Definition 3.1. A set $A \in \mathscr{F}$ is *invariant* if $T^{-1}A = A$.

Let $\mathscr{I}$ denote the collection of all invariant $A \in \mathscr{F}$. This is called the *invariant* σ-algebra. The name is motivated by the first half of the following.

Lemma 3.2. *$\mathscr{I}$ is a σ-algebra. In addition, if X is $\mathscr{I}$-measurable, then $X = X \circ T$ a.s.*

Proof. Evidently, $T^{-1}\varnothing = \{T^{-1}\omega : \omega \in \varnothing\} = \varnothing$. Therefore, $\varnothing \in \mathscr{I}$. If $A \in \mathscr{I}$ then $T^{-1}(A^c) = (T^{-1}A)^c$, so $A^c \in \mathscr{I}$. Finally, if $A_n \in \mathscr{I}$ $(n = 1, 2, \ldots)$, then $T^{-1}(\cup_{n=1}^{\infty} A_n) = \cup_{n=1}^{\infty} T^{-1}A_n = \cup_{n=1}^{\infty} A_n \in \mathscr{I}$. This proves that $\mathscr{I}$ is a σ-algebra.

Next, suppose X is $\mathscr{I}$-measurable. Define

$$X_n(\omega) := \sum_{k=-\infty}^{\infty} \left(\frac{k}{2^n}\right) \mathbf{1}_{X^{-1}[k/2^n,(k+1)2^{-n})}(\omega).$$

Because $X_n = X_n \circ T$ a.s. and $X_n \uparrow X$ (pointwise), it follows that $X = X \circ T$ a.s. $\square$

Theorem 3.3 (The individual ergodic theorem; Birkhoff)**.** *If $X_1 \in L^1(\mathrm{P})$ then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} X_j = \mathrm{E}(X_0 \,|\, \mathscr{I}) \quad \textit{a.s. and in } L^1(\mathrm{P}).$$

Equivalently, if $X_0 \in L^1(\mathrm{P})$ then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} X_0\left(T^j \omega\right) = \mathrm{E}\left(X_0 \,|\, \mathscr{I}\right)(\omega) \qquad \text{a.s. and in } L^1(\mathrm{P})$$

The proof of Theorem 3.3 rests on the following inequality.

Theorem 3.4 (The Hopf Maximal Ergodic Lemma). *Let* $S_n(\omega) := \sum_{j=0}^{n-1} X_j(\omega)$ *and* $M_n(\omega) := \max\{0, S_1(\omega), \ldots, S_n(\omega)\}$. *Then, as long as* $X_0 \in L^1(\mathrm{P})$,

$$\mathrm{E}[X_0; M_n > 0] \geq 0.$$

Before proving this, let us see what it entails. Choose and fix $\lambda > 0$, and define

$$\begin{aligned} X_n'(\omega) &:= X_n(\omega) - \lambda \\ S_n'(\omega) &:= \sum_{j=0}^{n-1} X_j'(\omega) \\ M_n'(\omega) &:= \max\left\{0, S_1'(\omega), \ldots, S_n'(\omega)\right\}. \end{aligned}$$

Note that

$$\begin{aligned} M_n'(\omega) &= \max\{0, S_1(\omega) - \lambda, S_2(\omega) - 2\lambda, \ldots, S_n(\omega) - n\lambda\} \\ &= \max_{1\leq j\leq n} \left[S_j(\omega) - \lambda j\right]_+. \end{aligned}$$

In particular,

$$M_n'(\omega) > 0 \qquad \Longleftrightarrow \qquad \max_{1\leq j\leq n} \left(\frac{S_j(\omega)}{j}\right) > \lambda.$$

Because $\{X_n'\}$ is a stationary process with $X_0' \in L^1(\mathrm{P})$, we can then apply the Hopf maximal ergodic lemma to $\{X_j'\}$ and obtain the bound, $\mathrm{E}[X_0'; M_n' > 0] \geq 0$. Equivalently,

$$\mathrm{P}\left\{\max_{1\leq j\leq n} \left(\frac{S_j}{j}\right) > \lambda\right\} \leq \frac{1}{\lambda} \mathrm{E}\left[X_0; \max_{1\leq j\leq n} \left(\frac{S_j}{j}\right) > \lambda\right].$$

Next replace X_n' by $X_n'' := -X_n - \lambda$. This too is stationary, and we find that

$$\mathrm{P}\left\{\max_{1\leq j\leq n} \left(\frac{-S_j}{j}\right) > \lambda\right\} \leq \frac{1}{\lambda} \mathrm{E}\left[-X_0; \max_{1\leq j\leq n} \left(\frac{-S_j}{j}\right) > \lambda\right].$$

Add the preceding two displays, and bound $\pm X_0$ by $|X_0|$ in the expectations, to find that

$$\mathrm{P}\left\{\max_{1\leq j\leq n} \left|\frac{S_j}{j}\right| > \lambda\right\} \leq \frac{1}{\lambda} \mathrm{E}\left[|X_0|; \max_{1\leq j\leq n} \left|\frac{S_j}{j}\right| > \lambda\right].$$

Exercise 3.5 (N. Wiener)**.** Prove that if $X_0 \in L^p(\mathrm{P})$ for some $p > 1$ then $\sup_{n\geq 1} |S_n/n| \in L^p(\mathrm{P})$.[2] Use this and Birkhoff's theorem (Theorem 3.3) to prove that if $X_0 \in L^p(\mathrm{P})$ for some $p > 1$, then $(S_n/n) \to \mathrm{E}(X_0|\mathscr{I})$ in $L^p(\mathrm{P})$ as well.

Proof of Theorem 3.4. For all $1 \leq j \leq n$, $M_n(T\omega) \geq S_j(T\omega)$. Therefore, $X_0(\omega) + M_n(T\omega) \geq X_0(\omega) + S_j(T\omega) = S_{j+1}(\omega)$. That is,

$$X_0 \geq \max_{1\leq j\leq n+1} S_j - M_n \circ T \geq \max_{1\leq j\leq n} S_j - M_n \circ T. \tag{3.1}$$

If $M_n(\omega) > 0$ then $M_n(\omega) = \max_{1\leq j\leq n} S_j(\omega)$. Therefore, multiply (3.1) by $\mathbf{1}_{\{M_n>0\}}$ and take expectations to find that

$$\begin{aligned} \mathrm{E}\left[X_0; M_n > 0\right] &\geq \mathrm{E}\left[M_n; M_n > 0\right] - \mathrm{E}\left[M_n \circ T; M_n > 0\right] \\ &= \mathrm{E}[M_n] - \mathrm{E}[M_n \circ T; M_n > 0] \\ &\geq \mathrm{E}[M_n] - \mathrm{E}[M_n \circ T]. \end{aligned}$$

The last line uses only the fact that $M_n \circ T \geq 0$, pointwise. Because T is measure-preserving, $\mathrm{E}[M_n \circ T] = \mathrm{E}[M_n]$, whence the lemma follows. □

Exercise 3.6. Derive the following following strengthening of Theorem 3.4: For all $A \in \mathscr{I}$,

$$\mathrm{E}\left[X_0; M_n > 0; A\right] \geq 0.$$

(HINT: Consider the process $X_n' := X_n \mathbf{1}_A$.)

Proof of Birkhoff's Theorem. Let $X_n' := X_n - \mathrm{E}(X_0|\mathscr{I}) - \varepsilon$ where $\varepsilon > 0$ is held fixed and non-random. Evidently, $\{X_n'\}$ is a stationary process too. Apply Hopf to it, using obvious notation, and find that for all invariant sets A,

$$\mathrm{E}[X_0'; M_n' > 0; A] \geq 0.$$

Cf. Exercise 3.6. Since $M_n' \leq M_{n+1}'$, $\lim_n M_n' = \sup_n M_n'$ exists but may be infinite. By the dominated convergence theorem,

$$\mathrm{E}\left[X_0'; \lim_n M_n' > 0; A\right] \geq 0.$$

We can apply this to the set

$$A := \left\{\omega \in \Omega : \limsup_{n\to\infty} \frac{S_n'(\omega)}{n} > 0\right\}.$$

Because $\limsup_n (S_n(\omega)/n) = \limsup_n (S_n'(T\omega)/n)$, A is an invariant set, and we find that

$$\mathrm{E}\left[X_0'; \lim_n M_n' > 0; \limsup_{n\to\infty} \frac{S_n'}{n} > 0\right] \geq 0.$$

[2] This is false for $p = 1$. However, for a bigger challenge, try proving that if $\mathrm{E}\{|X_0| \log_+ |X_0|\} < \infty$ then $\sup_{n\geq 1} |S_n/n| \in L^1(\mathrm{P})$. This "$L \log L$" condition cannot be improved upon (theorem of D. Burkholder).

But it is easy to see that

$$\left\{\omega \in \Omega : \limsup_{n\to\infty} \frac{S'_n(\omega)}{n} > 0\right\} \subseteq \left\{\omega \in \Omega : \lim_{n\to\infty} M'_n(\omega) > 0\right\}.$$

Thus, by the towering property of conditional expectations and the preceding discussion,

$$\mathrm{E}\left[\mathrm{E}\left(X'_0 \,\middle|\, \mathscr{I}\right); \limsup_{n\to\infty} \frac{S'_n}{n} > 0\right] = \mathrm{E}\left[X'_0; \limsup_{n\to\infty} \frac{S'_n}{n} > 0\right] \geq 0.$$

But $\mathrm{E}(X'_0|\mathscr{I}) = -\varepsilon < 0$. This proves that $\limsup_n (S'_n/n) \leq 0$ a.s. Equivalently, we have proved that

$$\limsup_{n\to\infty} \frac{S_n}{n} \leq \mathrm{E}(X_0 \,|\, \mathscr{I}) + \varepsilon \qquad \text{a.s.}$$

Since $\varepsilon > 0$ is arbitrary, we find that $\limsup_n (S_n/n) \leq \mathrm{E}(X_0|\mathscr{I})$ a.s. Apply this inequality to $\{-X_j\}$ to find that $\lim_n (S_n/n) = \mathrm{E}(X_0|\mathscr{I})$ a.s., which is the a.s.-portion of the Birkhoff theorem.

The second half on L^1-convergence follows a familiar path. For all $\nu > 0$ define $X_n^\nu = X_n \mathbf{1}_{\{|X_n| \leq \nu\}}$ and $S_n^\nu = \sum_{j=0}^{n-1} X_j^\nu$. Because $\{X_j^\nu\}_j$ is stationary, our proof thus far shows that $\lim_n (S_n^\nu/n) = \mathrm{E}(X_0^\nu|\mathscr{I})$ a.s. Because $|S_n^\nu/n| \leq \nu$ a.s., the bounded convergence theorem implies that (S_n^ν/n) converges to $\mathrm{E}(X_0^\nu|\mathscr{I})$ also in $L^1(\mathrm{P})$. Now,

$$\frac{S_n}{n} - \mathrm{E}(X_0|\mathscr{I}) = \underbrace{\left[\frac{S_n^\nu}{n} - \mathrm{E}(X_0^\nu|\mathscr{I})\right]}_{T_1} + \underbrace{\left[\frac{S_n - S_n^\nu}{n}\right]}_{T_2} + \underbrace{\mathrm{E}\left[X_0^\nu - X_0 \,\middle|\, \mathscr{I}\right]}_{T_3}.$$

We have proven that as $n \to \infty$, $T_1 \to 0$ a.s. and in $L^1(\mathrm{P})$. Let us concentrate on T_2 and T_2 then. But

$$\|T_2\|_1 \leq \mathrm{E}\{|X_0|; |X_0| > \nu\}$$
$$\|T_3\|_1 \leq \mathrm{E}\{|X_0|; |X_0| > \nu\}.$$

Therefore,

$$\limsup_{n\to\infty} \left\| \frac{S_n}{n} - \mathrm{E}(X_0 \,|\, \mathscr{I}) \right\|_1 \leq 2\mathrm{E}\{|X_0|; |X_0| > \nu\}.$$

Because the preceding is valid for all $\nu > 0$, we can subsequently let $\nu \to \infty$ to finish. □

4. Ergodicity

Definition 4.1. We say that $\{X_n\}$ [and/or T] is *ergodic* if $\mathscr{I}$ is a trivial σ-algebra. That is, the probability of all invariant sets is zero or one.

When T is ergodic, $\mathrm{E}(X_0|\mathscr{I}) = \mathrm{E}[X_0]$ is non-random. In particular, (S_n/n) converges to the non-random quantity $\mathrm{E}[X_0]$ as long as the latter is defined absolutely.

Example 4.2. If $\{X_n\}_{n=1}^{\infty}$ are i.i.d. then $\mathscr{I}$ is ergodic. Here is why: Take an invariant set $A \in \mathscr{I}$, and note that $A \in \sigma(\{X_n\}_{n=1}^{\infty})$. But $A = T^{-1}A$ (a.s.), and the latter is in $\sigma(\{X_n\}_{n=2}^{\infty})$ (why?). Iterated this to find that $A \in \sigma(\{X_n\}_{n=k}^{\infty})$ for all $k \geq 1$. In particular, A is measurable with respect to the tail σ-algebra of the X_n's, and the latter is trivial by the Kolmogorov zero-one law.

Example 4.3 (H. Weyl)**.** Set $\Omega = [0, 2\pi)$, $\mathscr{F} := \mathscr{B}(\Omega)$, and P :=the uniform distribution on $(\Omega, \mathscr{F})$. Choose and fix $\theta \in [0, 2\pi)$, and define $T\omega := (\omega + \theta)$ mod 2π. That is, T is "rotation by θ." It is clear that T preserves P. The interesting feature, here, is the following:

Fact: $\mathscr{I}$ is ergodic iff $\theta/(2\pi)$ is irrational.

To prove this, first suppose $\theta = 2\pi m/n$ where $m < n$ are both positive integers. Then we can choose a neighborhood A of 0 such that $\{T^k A\}_{k=0}^{\infty}$ are disjoint. (Picture?). Define $B := \cup_{k=0}^{n} T^k A$. A little thought [and/or the right picture] proves that B is invariant. But if $\mathrm{P}(A) > 0$, then $\mathrm{P}(B) = (n+1)\mathrm{P}(A)$, which is positive. If, in addition, we chose A sufficiently small, then $\mathrm{P}(B) \in (0,1)$, whence the non-ergodicity of T.

To finish this example, we prove the converse. Namely, that if $\theta/(2\pi)$ is irrational, then T is ergodic. The proof is essentially due to H. Weyl.

Choose and fix an invariant $A \in \mathscr{I}$, and define $f := \mathbf{1}_A$. Our goal is to prove that $\mathrm{P}(A) \in \{0, 1\}$. But we can develop f as a Fourier series

$$f(\omega) \sim \sum_{k=-\infty}^{\infty} \widehat{f}_k e^{-ik\omega},$$

where '$\sim$' denotes convergence in $L^2(0, 2\pi)$, and the $\widehat{f}_k$'s are the Fourier coefficients of f; i.e., $\widehat{f}_k := \int_0^{2\pi} f(x) \exp(ixk)\, dx = \int_A \exp(ixk)\, dx$ ($k \in \mathbf{Z}$). Because f is $\mathscr{I}$-measurable, Lemma 3.2 tells us that $f(\omega) = f(T\omega)$ for almost all $\omega \in [0, 2\pi)$. Therefore, the following holds for almost all $\omega \in [0, 2\pi)$:

$$\sum_{k=-\infty}^{\infty} \widehat{f}_k e^{-ik\omega} = \sum_{k=-\infty}^{\infty} \widehat{f}_k e^{-ikT\omega} = \sum_{k=-\infty}^{\infty} \widehat{f}_k e^{-ik\theta} e^{-ik\omega}.$$

Because the Fourier coefficients of a Fourier series are defined uniquely, this proves that $\widehat{f}_k = \widehat{f}_k \exp(-ik\theta)$ for all $k \in \mathbf{Z}$. Since $\theta/(2\pi)$ is irrational, $\exp(-ik\theta)$ cannot be one for any integer $k \neq 0$. Therefore, $\widehat{f}_k \equiv 0$ for all $k \neq 0$. This, in turn, means that $f(\omega) = \widehat{f}_0$ = a constant, a.s., which does the job.

Example 4.4 (Continued Fractions)**.** Let $\Omega = [0,1)$, $\mathscr{F} = \mathscr{B}(\Omega)$, and define G to be the so-called "Gauss law." That is, for all Borel sets A,

$$\mathrm{G}(A) := \frac{1}{\ln 2}\int_A \frac{dx}{1+x}.$$

Define T to be

$$T\omega := \frac{1}{\omega} - \left[\frac{1}{\omega}\right] \qquad \forall \omega \in [0,1),$$

where $[x]$ denotes the greatest integers $\le x$. Because $T\omega = (1/\omega) \pmod 1$,[3] this is the fractional portion of $(1/\omega)$.

It turns out that T preserves G. It is enough to prove that for all $0 < a < b < 1$,

$$\mathrm{G}(a,b) = \mathrm{G}\left(T^{-1}(a,b)\right).$$

Now,

$$\begin{aligned}
T^{-1}(a,b) &= \{\omega :\ a < T\omega < b\} \\
&= \left\{\omega :\ a < \frac{1}{\omega} - \left[\frac{1}{\omega}\right] < b\right\} \\
&= \bigcup_{n=1}^{\infty} \left\{\omega :\ n < \frac{1}{\omega} < n+1\ ,\ a < \frac{1}{\omega} - n < b\right\} \qquad \text{a.s.},
\end{aligned}$$

where the union is a disjoint one. It follows that

$$T^{-1}(a,b) = \bigcup_{n=1}^{\infty} \left(\frac{1}{b+n}, \frac{1}{a+n}\right),$$

where the union is a disjoint one. Therefore,

$$\begin{aligned}
G\left(T^{-1}(a,b)\right) &= \frac{1}{\ln 2}\sum_{n=1}^{\infty} \int_{1/(b+n)}^{1(a+n)} \frac{dx}{1+x} \\
&= \frac{1}{\ln 2}\sum_{n=1}^{\infty} \left\{\ln\left[\frac{a+n+1}{b+n+1}\right] - \ln\left[\frac{a+n}{b+n}\right]\right\} \\
&= \frac{1}{\ln 2}\ln\left[\frac{b+1}{a+1}\right] = \frac{1}{\ln 2}\int_a^b \frac{dx}{1+x} = \mathrm{G}(a,b).
\end{aligned}$$

This, and a monotone class theorem, together prove that G is preserved by T. In fact, it turns out that T is ergodic. [Ergodicity is not so easy to prove. We will demonstrate it later in §5, page 11.] Instead, let us prove that T is a kind of "shift."

If $\omega \in (0,1)$, then we define, iteratively,

$$a_0(\omega) := [1/\omega],\ a_1(\omega) := a_0(T\omega), \cdots,\ a_n(\omega) := a_{n-1}(T\omega), \ldots.$$

---

[3]Also, sometimes written as $\{1/\omega\}$.

Now, define

$$\phi_n(\omega) := \cfrac{1}{a_0(\omega) + \cfrac{1}{a_1(\omega) + \cdots + \cfrac{1}{a_{n-1}(\omega) + \cfrac{1}{a_n(\omega)}}}}. \tag{4.1}$$

If $\omega \in (0,1)$ is rational, then for all n large, $\phi_n(\omega) = \omega$. This is false when ω is irrational, but it is not too hard to see that $\lim_{n\to\infty} \phi_n(\omega) = \omega$. In any event, we can represent every $\omega \in [0,1)$ as

$$\omega = [a_0(\omega), a_1(\omega), \ldots].$$

The collection $(a_0(\omega), a_1(\omega), \ldots)$ are the so-called *digits of the continued fraction expansion* of ω. It follows that T is a **measure-preserving shift** on the continued fraction digits of all irrational numbers in $[0,1)$.

Example 4.5 (H. Poincaré)**.** Define the "recurrence set," R_A of A as follows:

$$R_A := \left\{\omega \in \Omega : \ T^n\omega \in A \text{ for infinitely-many } n\text{'s}\right\}$$

Evidently, R_A is an invariant set. By the ergodic theorem (Theorem 3.3), together yield

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathbf{1}_A\left(T^j\omega\right) = \mathrm{P}(A \,|\, \mathscr{I}) \qquad \text{a.s.}$$

But by the definition of R_A, the left-hand side is zero a.s. on R_A^c. Therefore, so is the right-hand side. This, and the already-mentioned fact that $R_A^c \in \mathscr{I}$ together yield the following:

$$0 = \mathrm{E}\left[\mathrm{P}(A \,|\, \mathscr{I}); R_A^c\right] = \mathrm{E}\left[\mathrm{P}(A \cap R_A^c \,|\, \mathscr{I})\right] = \mathrm{P}(A \cap R_A^c). \tag{4.2}$$

This is due to Poincaré; it states that for almost all of its starting points in A, the stochastic process $X_n(\omega) := T^n\omega$ returns to A infinitely often.

The preceding example has a surprising, but immediate, consequence.

Corollary 4.6 (Poincaré)**.** *If T is ergodic and $\mathrm{P}(A) > 0$, then A is "recurrent" in the sense that, with probability one, $T^n\omega \in A$ for infinitely-many n's.*

Proof. We are seeking to prove that if $\mathrm{P}(A) > 0$ then $\mathrm{P}(R_A) = 1$. Because R_A is invariant and T is ergodic, $\mathrm{P}(R_A)$ is zero or one. If $\mathrm{P}(R_A)$ were zero, then (4.2) would imply that $\mathrm{P}(A) = 0$, and this would result in a contradiction. □

Example 4.7 (Kesten, Spitzer, Whitman)**.** Let $\{X_n\}_{n=1}^\infty$ be i.i.d., taking values in $\mathbf{R}^d$, and $S_n := X_1 + \cdots + X_n$. The *range* of the random walk $\{S_n\}_{n=1}^\infty$ is defined as

$$R_n := \#\{S_1, \ldots, S_n\}.$$

The Kesten-Spitzer-Whitman theorem states that a.s. and in $L^1(\mathrm{P})$,

$$\lim_{n\to\infty} \frac{R_n}{n} = \mathrm{P}\{S_k \neq 0 \ \text{ for all } k \geq 1\}. \tag{4.3}$$

Construct, as before, $X_n(\omega) := \omega_n$, where $\omega = (\omega_1, \omega_2, \ldots) \in \Omega$. Therefore, $S_n(\omega) = \sum_{j=0}^{n-1} X_0(T^j\omega)$, where T denotes the usual shift; i.e., $T\omega = (\omega_2, \omega_3, \ldots)$. Then we prove (4.3) in two steps. First note that

$$\begin{aligned} R_n(\omega) &\geq \# \left\{1 \leq k \leq n : \ S_l(\omega) \neq S_k(\omega) \quad \forall\, l > k\right\} \\ &= \sum_{k=1}^n \mathbf{1}_E(T^k\omega), \quad \text{where} \\ E &:= \left\{\omega \in \Omega : \ S_k(\omega) \neq 0 \quad \forall\, k \geq 1\right\}. \end{aligned}$$

Our goal is to prove that $\lim_{n\to\infty}(R_n/n) = \mathrm{P}(E\,|\,\mathscr{I})$ a.s. Because $\mathscr{I}$ is trivial (Example 4.2), this proves the claim (4.3).

Apply the ergodic theorem to find that $\liminf_{n\to\infty}(R_n/n) \geq \mathrm{P}(E\,|\,\mathscr{I})$ a.s. This is half of our goal. To derive the converse let

$$E_m := \{\omega \in \Omega : \ S_1(\omega) \neq 0, \ldots, S_m(\omega) \neq 0\} \qquad (m = 1, 2, \ldots).$$

Then, for any fixed integer $m \geq 1$,

$$\begin{aligned} R_n(\omega) &\leq m + \#\left\{1 \leq j \leq n - m : \ S_l(\omega) \neq S_j(\omega) \quad \forall\, j < l \leq j + m\right\} \\ &= m + \sum_{j=1}^{n-m} \mathbf{1}_{E_m}(T^j\omega). \end{aligned}$$

This proves that $\limsup_{n\to\infty}(R_n/n) \leq \mathrm{P}(E_m\,|\,\mathscr{I})$ a.s. Let $m \uparrow \infty$ and appeal to the monotone convergence theorem to finish.

5. More on Contiunued Fractions

We now go back and prove the claimed ergodicity in Example 4.4. In fact, we will do a little more.

5.1. **A Rational Approximation to Irrationals.**

Let us take a detour in some classical analytic/metric number theory. Throughout, the notation of Example 4.4 is used.

Define for all $n \geq 1$ and all irrational $\omega \in \Omega$,

$$\begin{aligned} p_{-1}(\omega) &= 1, \qquad p_0(\omega) = 0, \cdots, \ p_n(\omega) = a_{n-1}(\omega)p_{n-1}(\omega) + p_{n-2}(\omega) \\ q_{-1}(\omega) &= 0, \qquad q_0(\omega) = 1, \cdots, \ q_n(\omega) = a_{n-1}(\omega)q_{n-1}(\omega) + q_{n-2}(\omega). \end{aligned}$$

One can check, using induction only, that

(5.1)
$$p_{n-1}q_n - p_n q_{n-1} = (-1)^n \qquad \forall n \geq 0,$$
$$\cfrac{1}{a_0 + \cfrac{1}{a_1 + \cdots + \cfrac{1}{a_{n-1} + r}}} = \frac{p_n + r q_{n-1}}{q_n + r q_{n-1}} \qquad \forall r \in [0,1],\ n \geq 1.$$

In particular, for all irrational $\omega \in \Omega$,

(5.2)
$$\frac{p_n}{q_n} = \cfrac{1}{a_0 + \cfrac{1}{a_1 + \cdots + \cfrac{1}{a_{n-1}}}}, \text{ and}$$
$$\omega = \frac{p_n(\omega) + T^n\omega \cdot q_{n-1}(\omega)}{q_n(\omega) + T^n\omega \cdot q_{n-1}(\omega)}.$$

[I am making some fuss about irrationals because ω is rational if and only if $a_n(\omega)$ is infinite for some finite integer $n \cdots$ try it.][4] Therefore, based on what we know already,

$$\lim_{n\to\infty} \frac{p_n(\omega)}{q_n(\omega)} = \omega,$$

for all irrational (and, trivially, all rational) $\omega \in \Omega$. One advantage of this setup is that we can get estimates on this particular rational approximation of ω, and the rate turns out to be good uniformly over all irrational ω's.

Proposition 5.1. *For all irrational $\omega \in \Omega$ and all integers $n \geq 2$,*

$$\left| \omega - \frac{p_n(\omega)}{q_n(\omega)} \right| \leq \frac{\sqrt{2}}{2^n}.$$

Proof. According to (5.2),

$$\left| \omega - \frac{p_n(\omega)}{q_n(\omega)} \right| = \frac{1}{q_n(\omega)\left[\frac{q_n(\omega)}{T^n\omega} + q_{n-1}(\omega) \right]}.$$

But $(T^n\omega)^{-1} \geq a_n(\omega)$, by virtue of construction. Therefore,

$$\left| \omega - \frac{p_n(\omega)}{q_n(\omega)} \right| \leq \frac{1}{q_n(\omega)\left[a_n(\omega) q_n(\omega) + q_{n-1}(\omega) \right]} = \frac{1}{q_n(\omega) q_{n+1}(\omega)}.$$

[4] The second assertion requires also the observation that, in continued-fraction notation, $\omega = [a_0(\omega), a_1(\omega), \cdots, a_{n-1}(\omega) + T^n\omega]$.

To finish, we will prove that for all $n \geq 2$,

$$q_n \geq 2^{(n-1)/2}. \tag{5.3}$$

We can use the definition of q_n to write $q_n = a_n q_{n-1} + q_{n-2} \geq q_{n-1} + q_{n-2}$. Since $q_m \geq 0$ for all $m \geq 0$, this proves that $q_n \geq q_{n-1}$. In particular, $q_n \geq 2q_{n-2}$. If n is even, then

$$q_n \geq 2q_{n-2} \geq 4q_{n-4} \geq 8q_{n-6} \geq \cdots \geq 2^j q_{n-2j} \geq \cdots \geq 2^{n/2}.$$

This certainly yields (5.3). If n is odd, then $q_n \geq q_{n-1} \geq 2^{(n-1)/2}$ by what we just proved in the even case. Equation (5.3) follows in general. □

5.2. **Ergodicity.** We now change our viewpoint slightly. Given a sequence $\{b_i\}_{i=0}^\infty$ of positive integers, we can write—as we did for $\omega \in \Omega$—the continued fraction expansion,

$$\rho_n(r) = \rho_n(r, \{b\}) := \cfrac{1}{b_0 + \cfrac{1}{b_1 + \cfrac{1}{b_2 + \cdots + \cfrac{1}{b_{n-1} + r}}}} = \frac{p_n + r p_{n-1}}{q_n + r q_{n-1}},$$

where now $p_n = p_n(\{b_i\}_{i=0}^\infty)$ and $q_n = q_n(\{b_i\}_{i=0}^\infty)$. I will alternate between the two notations freely. So, in particular, we can think of ρ_n also as

(5.4)

$$\rho_n(r, \omega) := \cfrac{1}{a_0(\omega) + \cfrac{1}{a_1(\omega) + \cdots + \cfrac{1}{a_{n-1(\omega)} + r}}} \qquad \forall n \geq 1, \omega \in \Omega, r \in [0,1].$$

Three important features of ρ_n: (i) When n is even, ρ_n is increasing; (ii) when n is odd, ρ_n is decreasing; and (iii) for all $\omega \in \Omega$, $\omega = \rho_n(T^n \omega)$.

For any sequence $\{b_i\}_{i=0}^\infty$ of positive integers, define

$$I_n(\{b\}) := \left\{\omega \in \Omega : \; a_i(\omega) = b_i \; \forall i = 0, \ldots, n-1\right\}.$$

Because our continued-fraction expansions are both uniquely determining, as well as uniquely determined, $I_n(\{b\})$ is an interval. In fact, depending on the parity of n, $I_n(\{b\})$ is either $[\rho_n(0), \rho_n(1)]$ or $[\rho_n(1), \rho_n(0)]$. In any case, we have seen—using the older viewpoint—that the length of $I_n(\{b\})$ is no more than $2^{-(n-1)/2}$; cf. Proposition 5.1. Therefore, the collection of all such intervals generates $\mathscr{F}$. Moreover, it follows that $\{\mathscr{F}_n\}_{n=1}^\infty$ is a filtration where $\mathscr{F}_n$ denotes the σ-algebra generated by $I_n(\{b\})$ as the sequence $\{b\}$ is varied.

For future use, let us note that $\mathscr{F}_n$ is the σ-algebra generated by the function $\omega \mapsto \rho_n(r,\omega)$. Therefore,

$$a_0, \dots, a_{n-1} \text{ are } \mathscr{F}_n\text{-measurable for all } n \geq 1. \tag{5.5}$$

For all $0 \leq x < y \leq 1$, $n \geq 1$, and for a fixed determined sequence $\{b\}$ of positive integers, consider

$$\begin{aligned}
\mathrm{P}\left\{T^n \in (x,y]\,,\ I_n(\{b\})\right\} &= \left|\rho_n(x,\{b\}) - \rho_n(y,\{b\})\right| \\
&= \left|\frac{p_n + yp_{n-1}}{q_n + yq_{n-1}} - \frac{p_n + xp_{n-1}}{q_n + xq_{n-1}}\right| \\
&= (y-x)\left|\frac{p_{n-1}q_n - p_n q_{n-1}}{(q_n + yq_{n-1})(q_n + xq_{n-1})}\right| \\
&= \frac{y-x}{(q_n + yq_{n-1})(q_n + xq_{n-1})};
\end{aligned}$$

see the second equation in (5.1). Apply this with $x = 0$ and $y = 1$ to deduce that $\mathrm{P}(I_n(\{b\})) = (q_n(q_n + q_{n-1}))^{-1}$. Therefore,

$$\begin{aligned}
\mathrm{P}\left(x < T^n \leq y \,\middle|\, I_n(\{b\})\right) &= (y-x)\frac{q_n(q_n + q_{n-1})}{(q_n + yq_{n-1})(q_n + xq_{n-1})} \\
&\geq (y-x)\frac{q_n}{(q_n + q_{n-1})} \geq \frac{y-x}{2}.
\end{aligned}$$

Since $\mathscr{F}_n$ is countably-generated, this means that for all $0 \leq x < y \leq 1$,

$$\mathrm{P}\left(x < T^n \leq y \,\middle|\, \mathscr{F}_n\right) \geq \frac{y-x}{2} \qquad \text{a.s.[P]}.$$

Now apply a montone-class argument to deduce the following.

Lemma 5.2. *For all $n \geq 1$, $0 \leq x < y \leq 1$, and all $A \in \mathscr{F}$,*

$$\mathrm{P}\left(T^n \in A \,\middle|\, \mathscr{F}_n\right) \geq \frac{1}{2}\mathrm{P}(A) \qquad \textit{a.s.}[\mathrm{P}].$$

Thus, if A is an invariant set, $\{T^n \in A\} = T^{-n}A = A$ a.s.[P], and we find that

$$\mathrm{P}(A\,|\,\mathscr{F}_n) \geq \frac{1}{2}\mathrm{P}(A).$$

Let $n \to \infty$, and appeal to Lévy's martingale convergence theorem to find that

$$\mathbf{1}_A \geq \frac{1}{2}\mathrm{P}(A) \qquad \text{a.s.[P]}.$$

[We have used the already-mentioned fact that $\vee_n \mathscr{F}_n = \mathscr{F}$.] Thus, whenever $\mathrm{P}(A) > 0$ $\mathbf{1}_A = 1$ a.s., and so $\mathrm{P}(A) = 1$. This proves that all invariant sets have P-measure zero or one.

Recall the gauss measure G from Example 4.4. Note that for all $A \in \mathscr{F}$,

$$\frac{\mathrm{P}(A)}{2\ln 2} \le \mathrm{G}(A) \le \frac{\mathrm{P}(A)}{\ln 2}. \tag{5.6}$$

Therefore, any G-null set is also P-null, and any G-full set is also P-full. Because of the Birkhoff theorem, we have proven the following.

Proposition 5.3. *T is ergodic under either measure* P *or* G*. In particular, under either measure,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f\left(T^j \omega\right) = \int f\, d\mathrm{G} \qquad a.s.,$$

valid for all $f \in L^1(\mathrm{G}) = L^1(\mathrm{P})$.

5.3. **Applications to Metric Number Theory.** For our first application choose and fix a positive integer k, and define $f(\omega) := \mathbf{1}_{\{a_0=k\}}(\omega)$. To this we apply Proposition 5.3 to find that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathbf{1}_{\{a_j=k\}}(\omega) = \mathrm{G}\{a_0 = k\}, \tag{5.7}$$

valid almost surely P and/or G. But

$$\begin{aligned} \mathrm{G}\{a_0 = k\} &= \frac{1}{\ln 2} \int_{\{a_0=k\}} \frac{\mathrm{P}(d\omega)}{1+\omega} = \frac{1}{\ln 2} \int_{1/(k+1)}^{1/k} \frac{d\omega}{1+\omega} \\ &= \frac{2}{\ln 2} \ln\left(\frac{k+1}{k(k+2)}\right). \end{aligned}$$

Thus, we have the following result about the asymptotic distribution of continued-fraction digits of almost all numbers: Outside one null set [P and/or G] of ω's,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathbf{1}_{\{a_j=k\}}(\omega) = \frac{2}{\ln 2} \ln\left(\frac{k+1}{k(k+2)}\right) \qquad \forall k \ge 1. \tag{5.8}$$

Exercise 5.4. Prove that almost surely, $\lim_{n\to\infty} n^{-1} \sum_{i=0}^{n-1} a_i(\omega) = \infty$. (Hint: Truncate and then apply Birkhoff.)

For our next application we need a result from martingale theory.

Theorem 5.5 (Lévy's Borel–Cantelli Lemma)**.** *Let $\{\mathscr{A}_n\}_{n=0}^\infty$ denote a filtration of σ-algebras on some probability space $(A, \mathscr{A}, \mathrm{Q})$. Suppose $\{A_n\}_{n=0}^\infty$ is a sequence of sets such that $A_n \in \mathscr{A}_n$ ($n \ge 0$). Then $\sum_{n=1}^\infty \mathbf{1}_{A_n} < \infty$ almost surely if $\sum_{n=1}^\infty \mathrm{Q}(A_n \mid \mathscr{A}_{n-1}) < \infty$ a.s., and conversely.*

Proof. Let $M_N := \sum_{n=1}^N \mathbf{1}_{A_n} - \sum_{n=1}^N \mathrm{Q}(A_n \mid \mathscr{A}_{n-1})$ to find that $\{M_N\}_{N=1}^\infty$ is a mean-zero martingale. In fact, M is the martingale portion of the Doob decomposition of $N \mapsto \sum_{n=1}^N \mathbf{1}_{A_n}$.

For every $\lambda > 0$ define $T_\lambda := \inf\{N \geq 1 : \; M_N < -\lambda\}$ where $\inf\varnothing := \infty$. Note that $\sup_k |M_k - M_{k-1}| \leq 2$. Therefore,

$$M_{T_\lambda \wedge N} \geq M_{(T_\lambda - 1)\wedge N} - 2 \geq -\lambda - 2 \qquad {}^\forall N \geq 1.$$

Therefore, $\lambda + 2 + M_{T_\lambda \wedge N}$ defines a non-negative martingale (indexed by N), which we know converges a.s. [martingale convergence theorem]. Therefore, $\lim_{N\to\infty} M_N$ exists and is finite, a.s. on $\{T_\lambda = \infty\}$. Take the union over a countable set of λ's that tend to ∞ to find that

$$\lim_{N\to\infty} M_N \text{ exists a.s. on } \left\{\inf_N M_N > -\infty\right\}.$$

Apply this also to $-M$ to deduce that

$$\lim_{N\to\infty} M_N \text{ exists a.s. on } \left\{\inf_N M_N > -\infty \text{ or } \sup_N M_N < \infty\right\}.$$

In other words, with probability one, either $\lim_{N\to\infty} M_N$ exists and is finite, or $\sup_N M_N = \infty$ and $\inf_N M_N = -\infty$. This proves the result in disguise. □

For our next application, let $\{E_n\}_{n=0}^\infty$ denote a sequence of measurable subsets of Ω. If $\sum_{n=0}^\infty \mathrm{G}\{a_0 \in E_n\} < \infty$ then $\sum_{n=0}^\infty \mathrm{G}\{a_n \in E_n\} < \infty$ by stationarity. Therefore, according to the Borel–Cantelli lemma,

$$\mathrm{G}\left\{a_n \in E_n \text{ for infinitely-many } n\text{'s}\right\} = 0.$$

On the other hand, if $\sum_{n=0}^\infty \mathrm{G}\{a_0 \in E_n\} = \infty$, then $\sum_{n=0}^\infty \mathrm{P}\{a_n \in E_n\} = \infty$ by (5.6) and stationarity. Therefore, Lemma 5.2 implies that $\sum_n \mathrm{P}(a_n \in E_n \,|\, \mathscr{F}_n) = \infty$ a.s. This, (5.5), and Lévy's Borel–Cantelli lemma (Theorem 5.5) together yield the following:

$$\mathrm{G}\left\{a_n \in E_n \text{ for infinitely-many } n\text{'s}\right\} = 0 \qquad \text{if } \sum_n \mathrm{G}\{a_0 \in E_n\} < \infty;$$
$$\mathrm{G}\left\{a_n \in E_n \text{ for infinitely-many } n\text{'s}\right\} = 1 \qquad \text{if } \sum_n \mathrm{G}\{a_0 \in E_n\} = \infty.$$

The same is true if we replace G by P everywhere; cf. (5.6). Now suppose $\lambda_n \uparrow \infty$ are positive, non-random integers. Then,

$$\mathrm{P}\{a_0 \geq \lambda_n\} = \sum_{i=\lambda_n}^\infty \mathrm{P}\{a_0 = i\} = \sum_{i=\lambda_n}^\infty \left(\frac{1}{i} - \frac{1}{i+1}\right) = \sum_{i=\lambda_n}^\infty \frac{1}{i(i+1)}.$$

This proves that there exist $C > c > 0$ such that for all $n \geq 1$,

$$\frac{c}{\lambda_n} \leq \mathrm{P}\{a_0 \geq \lambda_n\} \leq \frac{C}{\lambda_n} \qquad {}^\forall n \geq 1.$$

Consequently,

$$a_n < \lambda_n \text{ for all but a finite number of } n\text{'s if } \sum_n \frac{1}{\lambda_n} < \infty;$$

$$a_n \geq \lambda_n \text{ for all but a finite number of } n\text{'s if } \sum_n \frac{1}{\lambda_n} = \infty.$$

But $\sum_n \lambda_n^{-1} < \infty$ if and only if $\sum_n (\kappa\lambda_n)^{-1} < \infty$ for any (and all) $\kappa > 0$. Therefore, we have proven the following: Almost surely [P and/or G],

$$\limsup_{n\to\infty} \left(\frac{a_n}{\lambda_n}\right) = \begin{cases} 0 & \text{if} \sum_n \frac{1}{\lambda_n} < \infty, \\ \infty & \text{if} \sum_n \frac{1}{\lambda_n} = \infty. \end{cases} \tag{5.9}$$

For example, with probability one,

$$\lim_{n\to\infty} \frac{a_n}{n} = \infty \quad \text{but} \quad \limsup_{n\to\infty} \frac{a_n}{n(\ln n)^2} = 0.$$

Exercise 5.6. We have seen that Lebesgue-almost all numbers $\omega \in (0,1]$ have the property that infinitely-many of their continued-fraction digits $(a_0(\omega), a_1(\omega), \ldots)$ attain any predescribed integer λ. What if λ is time-varying? To be more precise, let $\{\lambda_n\}_{n=1}^{\infty}$ denote a sequence of positive, non-random integers that increase without bound. Prove that $\mathrm{P}\{a_n = \lambda_n \text{ for infinitely-many } n\} = 0$ or 1. Find an analytic characterization for each case; your description should be solely in terms of the λ_n's. Use this to answer the following question: For exactly what values of $\alpha > 0$ is $\mathrm{P}\{a_n = \lfloor n^{\alpha} \rfloor \text{ for infinitely-many } n\text{'s}\} = 1$?

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF UTAH, 155 S. 1400 E. SALT LAKE CITY, UT 84112–0090

E-mail address: davar@math.utah.edu

URL: http://www.math.utah.edu/~davar