II. Ergodic measures for F

Objekttyp: **Chapter**

Zeitschrift: **L'Enseignement Mathématique**

Band (Jahr): **40 (1994)**

Heft 3-4: **L'ENSEIGNEMENT MATHÉMATIQUE**

PDF erstellt am: **22.09.2024**

where the first $n - 1$ digits on the right side of (5) are zeros and the number on the right is expressed in base 3 so that $\alpha_k = 0$ or 2. Consequently $|h^{-1}(x) - h^{-1}(y)| \leqslant 3^{-n-1} < \varepsilon$. Since h^{-1} is a continuous bijection and I_∞ is compact, it follows from a well-known theorem in topology that h is continuous and therefore h is a homeomorphism.

(b) Suppose $x \in I_\infty \cap \left[\frac{2}{3}, 1\right]$ and let the base 3 expansion of x be given by $x = 0.2\alpha_1\alpha_2\alpha_3\ldots$, thus $h(x) = (0, x_1, x_2, \ldots)$, where $x_i = 1 - \frac{\alpha_i}{2}$. Then $F(x) = x - \frac{2}{3}$ has base 3 expansion given by $F(x) = 0.0\alpha_1\alpha_2\alpha_3\ldots$. Therefore $h(F(x)) = (1, x_1, x_2, \ldots)$ which is the same as $h(x) + \mathbf{1}$. If $x = 0$, then $h(F(0)) = (0, 0, 0, \ldots) = h(0) + \mathbf{1}$. If $x \in I_\infty \cap \left(0, \frac{1}{3}\right]$, then $x \in \left[\frac{2}{3^i}, \frac{1}{3^{i-1}}\right)$ for some $i \geqslant 2$. A brief calculation shows that

$$F(x) = x + 0.\overbrace{2\ldots2}^{i-2}1\overset{\overset{i}{\downarrow}}{1}\bar{0} \tag{6}$$

where the number on the right is expressed in base 3 and the second "1" occurs i places after the decimal point. It now follows from (6) and base 3 addition that $h(F(x)) = h(x) + \mathbf{1}$ and therefore $F(I_\infty) = I_\infty$.

(c) Since S is the completion of the nonnegative integers under the identification (4), the set

$$\{(x_0, x_1, x_2, \ldots, x_n, 0, 0, 0, \ldots) : n = 0, 1, 2, \ldots, x_i = 0 \text{ or } 1\}$$

which equals $\{n\mathbf{1} : n = 0, 1, 2, \ldots\}$ is dense in S. It is easy to establish that the map which takes x to $x + z$ is a homeomorphism from S to S for any fixed $z \in S$. Thus

$$\{n\mathbf{1} : n = 0, 1, 2, \ldots\} + z$$

is dense in S for any z. But $h^{-1}(\{n\mathbf{1} : n = 0, 1, 2, \ldots\} + z)$ is precisely the F-orbit of $h^{-1}(z)$. Therefore the F-orbit of any $y = h^{-1}(z) \in I_\infty$ is dense in I_∞. □

II. ERGODIC MEASURES FOR F

A measure μ on a set X is called a probability measure if $\mu(X) = 1$; the pair (X, μ) is then called a probability space. Given a measurable transformation $T: X \to X$ on a probability space (X, μ), μ is T-invariant if $\mu = \mu \circ T^{-1}$, i.e., for any measurable set $B \subset X$, $\mu(B) = \mu(T^{-1}(B))$. The probability measure μ is ergodic if $T^{-1}(A) = A$ implies that $\mu(A)$ is 0 or 1.

One way to study the attracting, often fractal, sets of a dynamical system (for example the ternary Cantor set for F) is to study the invariant probability measures which are supported on the attractors. (The support of a measure on [0, 1] is the intersection of all closed subsets of [0, 1] whose complements have zero measure.) This is an especially fruitful approach when the attracting set lies in a high dimensional space so that a purely geometric description is unfeasible (cf. [9]). Particularly important, though generally difficult to find, are the invariant ergodic probability measures for the dynamical system. For Devaney's transformation $F: [0, 1] \to [0, 1]$, we will find all F-invariant ergodic probability measures.

To construct a probability measure on I_∞, we borrow some ideas from the theory of iterated function systems as developed by Barnsley [10]. Let $P[0, 1]$ denote the set of all Borel probability measures on [0, 1], i.e., probability measures on the σ-algebra generated by all open sets on [0, 1]. For any $\mu, \nu \in P[0, 1]$, define

$$d(\mu, \nu) \equiv \sup\{ \,|\int f d\mu - \int f d\nu \,| : f : [0, 1] \to [0, 1] \text{ and } |f(x) - f(y)| \leqslant |x - y|, \ \forall x, y \in [0, 1]\} .$$

Then d is a complete metric on $P[0, 1]$. Let $w_1, w_2: [0, 1] \to [0, 1]$ by $w_1(x) = \frac{1}{3}x$ and $w_2(x) = \frac{1}{3}x + \frac{2}{3}$. Let $M: P[0, 1] \to P[0, 1]$ by $M(\mu) = \frac{1}{2}\mu \circ w_1^{-1} + \frac{1}{2}\mu \circ w_2^{-1}$. M is called the Markov operator for an iterated function system defined by w_1, w_2. It follows as a special case of a more general theorem from ref. 10 that

$$d(M(\mu), M(\nu)) \leqslant \tfrac{1}{3} d(\mu, \nu) , \tag{7}$$

so that by the well-known contraction mapping theorem, M has a unique fixed point ν_∞, i.e., $M(\nu_\infty) = \nu_\infty$.

Notice that the intervals I_n^k defined above are in one-to-one correspondence with iterations of the form $w_{i_1} \circ w_{i_2} \circ \cdots \circ w_{i_n}([0, 1])$, where $i_k = 1$ or 2. For example, $w_1 \circ w_2([0, 1]) = \left[\frac{2}{9}, \frac{1}{3}\right] \equiv I_2^2$.

LEMMA 2.1. *For all* $n \geqslant 0$ *and* $k \leqslant 2^n, \nu_\infty(I_n^k) = 2^{-n}$.

Proof. The proof follows by induction. For $n = 0$, $I_0^1 = [0, 1]$ and $\nu_\infty([0, 1]) = 1 = 2^0$ because ν_∞ is a probability measure. Assume $M(\nu_\infty) = \nu_\infty$ and the induction hypothesis, $\nu_\infty(I_n^k) = \nu_\infty(w_{i_1} \circ w_{i_2} \circ \cdots \circ w_{i_n}[0, 1]) = 2^{-n}$

for any sequence $i_1, \ldots, i_n$ and any $k \leqslant 2^n$. For any $j \leqslant 2^{n+1}$, there is a sequence $i_1, \ldots, i_{n+1}$ such that

$$\begin{aligned} \nu_\infty(I^j_{n+1}) &= \nu_\infty(w_{i_1} \circ w_{i_2} \circ \cdots \circ w_{i_n} \circ w_{i_{n+1}}[0, 1]) \\ &= \tfrac{1}{2}\nu_\infty \circ w_1^{-1}(w_{i_1} \circ w_{i_2} \circ \cdots \circ w_{i_n} \circ w_{i_{n+1}}[0, 1]) \\ &\quad + \tfrac{1}{2}\nu_\infty \circ w_2^{-1}(w_{i_1} \circ w_{i_2} \circ \cdots \circ w_{i_n} \circ w_{i_{n+1}}[0, 1])\,. \end{aligned} \tag{8}$$

Notice that $\nu_\infty \circ w_j^{-1}(w_{i_1} \circ w_{i_2} \circ \cdots \circ w_{i_n} \circ w_{i_{n+1}}[0, 1]) = 2^{-n}$ or 0 depending on whether or not $j = i_1$. Combining this observation with (8) shows $\nu_\infty(I^j_{n+1}) = \frac{1}{2}2^{-n} = 2^{-n-1}$. $\square$

PROPOSITION 2.1. *The measure* ν_∞ *is supported on the Cantor set* I_∞.

Proof. By Lemma 2.1 $\nu_\infty(I_n) = \sum_{k=1}^{2^n} \nu_\infty(I^k_n) = 2^n 2^{-n} = 1$ for all n. Since $I_0 \supset I_1 \supset I_2 \supset \cdots$ is a decreasing sequence, $\nu_\infty(I_\infty) = \nu_\infty\left(\bigcap_{n=1}^{\infty} I_n\right) = \lim_{n\to\infty} \nu_\infty(I_n) = 1$. If C is a proper closed subset of I_∞, there is an open interval U in the complement of C whose intersection with I_∞ is nonempty. Let $x \in U \cap I_\infty$. For some positive integers n and k, $x \in I^k_n \subset U$. Then since I^k_n is in the complement of C, $\nu_\infty(C) \leqslant 1 - 2^{-n} < 1$. $\square$

The next step is to show that ν_∞ is an invariant measure for the dynamical system $F: [0, 1] \to [0, 1]$. This means that $\nu_\infty = \nu_\infty \circ F^{-1}$, i.e., for any Borel set $B \subset [0, 1]$, $\nu_\infty(B) = \nu_\infty(F^{-1}(B))$.

LEMMA 2.2. *If* μ *is any probability measure on* $[0, 1]$ *and* $\mu(I^k_n) = 2^{-n}$ *for all* n *and* k, *then* $\mu = \nu_\infty$.

Proof. The condition $\mu(I^k_n) = 2^{-n}$ implies that $\mu\{x\} = 0$ for every $x \in [0, 1]$. This is clearly true if $x \notin I_\infty$ since $\mu(I_\infty) = 1$, as follows from the proof of Proposition 2.1. If $x \in I_\infty$, then for every n there exists a k such that $x \in I^k_n$. Thus $\mu\{x\} \leqslant 2^{-n}$ for every n and hence $\mu\{x\} = 0$.

Any interval of the form $[a, b] \subset [0, 1]$ where a and b have finite ternary expansions (finite "decimal" expansions in base 3) is a disjoint union of sets of the following form:

1) I^k_n for some n and k
2) intervals in the complement of I_n for some n
3) $\{a\}, \{b\}$.

Since μ and ν_∞ agree on each of the sets in 1, 2, and 3, $\mu[a,b] = \nu_\infty[a,b]$. It is a standard result in measure theory that two probability measures which are equal on a collection of measurable sets, closed under finite intersections, are equal on the σ-algebra generated by that collection. In this case the σ-algebra generated by sets of the form $[a,b] \subset [0,1]$ where a and b have finite ternary expansions is just the Borel σ-algebra on $[0,1]$. □

It is well-known [11] that the "adding machine" is uniquely ergodic. This fact may also be deduced from Proposition 2.2 below.

PROPOSITION 2.2. *ν_∞ is invariant under F. If μ is a probability measure invariant under F and $\mu(I_\infty) = 1$, then $\mu = \nu_\infty$.*

Proof. To show that ν_∞ is invariant under F, it suffices by Lemma 2.2 to show that

$$\nu_\infty(F^{-1}(I_n^k)) = 2^{-n}$$

for all k and n. By Theorem 1.1 (a), given k there exists a unique integer j such that

$$I_n^j \subset F^{-1}(I_n^k) \subset I_n^j \cup [0,1]\backslash I_\infty \tag{9}$$

(because F is a permutation on the intervals in I_n). Thus,

$$2^{-n} = \nu_\infty(I_n^j) \leqslant \nu_\infty(F^{-1}(I_n^k)) \leqslant \nu_\infty(I_n^j \cup [0,1]\backslash I_\infty) = 2^{-n}.$$

Hence ν_∞ is invariant under F. Suppose μ is a probability measure invariant under F and $\mu(I_\infty) = 1$. Then by (9) and F-invariance,

$$\mu(I_n^j) \leqslant \mu(F^{-1}(I_n^k)) = \mu(I_n^k) \leqslant \mu(I_n^j \cup [0,1]\backslash I_\infty) = \mu(I_n^j). \tag{10}$$

Since F is a cyclic permutation on the intervals in I_n, for any n and any positive integers $j, k \leqslant 2^n$, $\mu(I_n^j) = \mu(I_n^k)$. As there are 2^n intervals of the form I_n^k in I_n, it follows that $\mu(I_n^k) = 2^{-n}$ for all n and k. Therefore, $\mu = \nu_\infty$, by Lemma 2.2. □

Before investigating the ergodicity of ν_∞, we introduce some other invariant measures for F. For a fixed $x \in [0,1]$, let δ_x be the probability measure on $[0,1]$ which assigns 1 to any Borel measurable set containing x and assigns 0 to all other measurable sets. The probability measure δ_x is sometimes called the "point mass at x" or the "Dirac delta function at x."

For each nonnegative integer n, let y_n be the smallest number in $[0, 1]$ which lies in the unique orbit with prime period 2^n for the function F and define

$$\nu_n = \frac{1}{2^n} \sum_{k=0}^{2^n-1} \delta_{F^k(y_n)}. \tag{11}$$

The measure ν_n assigns mass 2^{-n} to each point in the unique orbit with period 2^n of F, i.e., $\nu_n(B) = k2^{-n}$ if B contains exactly k points from $P_n \equiv \mathrm{Per}_{2^n}(F)$. It is not difficult to check that for each $n = 0, 1, 2, \ldots, \nu_n$ is invariant with respect to F, and ν_n is the only F-invariant probability measure on the set of points P_n.

Let $C[0, 1]$ be the Banach space consisting of all continuous functions with the maximum norm $\| \ \|$ given by $\|f\| \equiv \max\{|f(x)| : x \in [0, 1]\}$. Let $M_F[0, 1] \equiv \{\mu \in P[0, 1] : \mu = \mu \circ F^{-1}\}$ be the set of invariant probability measures on $[0, 1]$. $M_F[0, 1]$ may be identified in a natural way with a metrizable, compact, convex subset of the dual space of $C[0, 1]$. The compact, metrizable topology on $M_F[0, 1]$ is the weakest topology which makes the map $\mu \to \int f(x)\mu(dx)$ continuous for each $f \in C[0, 1]$; it is called the vague or weak-* topology. An extreme point of the convex set $M_F[0, 1]$ is a measure μ which is not a convex combination of any other two points in $M_F[0, 1]$, i.e., μ is extreme if whenever $\mu = \alpha\mu_1 + (1 - \alpha)\mu_2$, $0 < \alpha < 1$, and $\mu_1, \mu_2 \in M_F[0, 1]$, then $\mu = \mu_1 = \mu_2$. It is a consequence of the Krein-Milman Theorem that the set of extreme points of $M_F[0, 1]$ is non-empty. The following theorem is a specialization of a well-known result in ergodic theory (see for example [12]).

THEOREM 2.1. *The F-invariant measure μ is an extreme point of $M_F[0, 1]$ if and only if μ is ergodic with respect to F on $[0, 1]$.*

As a consequence of Theorem 2.1, we have the following proposition.

PROPOSITION 2.3. *For each $n = 0, 1, 2, \ldots, \infty$, ν_n is an extreme point of $M_F[0, 1]$ and is therefore ergodic with respect to F.*

Proof. Consider the case $n = \infty$. Suppose there exist F-invariant probability measures μ_1 and μ_2 and $\alpha \in (0, 1)$ such that $\nu_\infty = \alpha\mu_1 + (1 - \alpha)\mu_2$. Then since $\nu_\infty(I_\infty) = 1$, $\mu_1(I_\infty) = \mu_2(I_\infty) = 1$. Then by Proposition 2.2, $\mu_1 = \mu_2 = \nu_\infty$. Thus ν_∞ is an extreme point of $M_F[0, 1]$ and is ergodic by Theorem 2.1. The cases, $n = 0, 1, 2, \ldots$ are handled in the same manner. □

PROPOSITION 2.4. *The measures* $\nu_0, \nu_1, \nu_2, \ldots, \nu_\infty$ *are the only probability measures on* $[0, 1]$ *ergodic with respect to* F.

Proof. Let μ be an F-invariant probability measure with support A_μ. Let $\mathscr{P} \equiv \{x \in [0, 1] : F^k(x) \in P_n \text{ for some } k, n\}$ be the set of periodic or eventually periodic points. By Theorem 1.1 $A_\mu \setminus \mathscr{P} \subset \bigcup_{k=0}^{\infty} F^{-k}(I_n)$ for each n. Thus

$$\mu(A_\mu \setminus \mathscr{P}) \leqslant \mu\left(\bigcup_{k=0}^{\infty} F^{-k}(I_n)\right).$$

Since $I_n \subset F^{-1}(I_n) \subset F^{-2}(I_n) \subset \cdots$ is an increasing sequence of sets,

$$\mu\left(\bigcup_{k=0}^{\infty} F^{-k}(I_n)\right) = \lim_{k \to \infty} \mu(F^{-k}(I_n)) = \mu(I_n) \geqslant \mu(A_\mu \setminus \mathscr{P})$$

for all n. Since $I_\infty = \bigcap_{n=1}^{\infty} I_n$, is a decreasing sequence of sets,

$$\mu(I_\infty) \geqslant \mu(A_\mu \setminus \mathscr{P}).$$

Similarly,

$$\mu\left(\bigcup_{k=0}^{\infty} F^{-k}(P_n)\right) = \lim_{k \to \infty} \mu(F^{-k}(P_n)) = \mu(P_n) \geqslant \mu\left(A_\mu \cap \bigcup_{k=0}^{\infty} F^{-k}(P_n)\right)$$

Thus,

$$\mu\left(\bigcup_{n=0}^{\infty} P_n\right) = \sum_{n=0}^{\infty} \mu(P_n) \geqslant \mu(A_\mu \cap \mathscr{P}).$$

Since $A_\mu = (A_\mu \setminus \mathscr{P}) \cup (A_\mu \cap \mathscr{P})$, it follows that

$$\begin{aligned} 1 \geqslant \mu\left(I_\infty \cup \bigcup_{n=0}^{\infty} P_n\right) &= \mu(I_\infty) + \mu\left(\bigcup_{n=0}^{\infty} P_n\right) \\ &\geqslant \mu(A_\mu \setminus \mathscr{P}) + \mu(A_\mu \cap \mathscr{P}) = \mu(A_\mu) = 1. \end{aligned}$$

Hence $\mu\left(I_\infty \cup \bigcup_{n=0}^{\infty} P_n\right) = 1$.

From Theorem 1.1, $\bigcup_{n=0}^{\infty} A_n \setminus \bigcup_{n=0}^{\infty} P_n$ is a union of open intervals. Therefore

$$\left(I_\infty \cup \left(\bigcup_{n=0}^{\infty} P_n\right)\right)^c = \bigcup_{n=0}^{\infty} A_n \setminus \bigcup_{n=0}^{\infty} P_n$$

is open. Hence $I_\infty \cup \left(\bigcup_{n=0}^{\infty} P_n\right)$ is a closed subset of $[0, 1]$. Since $\mu\left(I_\infty \cup \bigcup_{n=0}^{\infty} P_n\right) = 1$ the definition of the support of a measure implies that

$A_\mu \subset I_\infty \cup \bigcup_{n=0}^{\infty} P_n$.

Since F is a continuous function and A_μ is closed, $F^{-1}(A_\mu)$ is also closed. By F-invariance, $\mu(F^{-1}(A_\mu)) = 1$ and therefore $A_\mu \subset F^{-1}(A_\mu)$. Hence $F(A_\mu) \subset A_\mu$. Suppose $A_\mu \cap I_\infty \neq \varnothing$ and let $x \in A_\mu \cap I_\infty$. Then $F^k(x) \in A_\mu \cap I_\infty$ for all k. Since by Theorem 1.2 the orbit of x is dense in I_∞ and $A_\mu \cap I_\infty$ is closed, it follows that $A_\mu \cap I_\infty = I_\infty$, i.e., $I_\infty \subset A_\mu$. A similar argument shows that if $A_\mu \cap P_n \neq \varnothing$, then $P_n \subset A_\mu$.

Thus A_μ is a union of one or more of $I_\infty, P_0, P_1, P_2, \ldots$ Furthermore, if $\mu(I_\infty) = 1$ then by Proposition 2.2, $\mu = \nu_\infty$. Similarly if $\mu(P_n) = 1$ for some n, then $\mu = \nu_n$. If $0 < \mu(I_\infty) < 1$, then

$$F^{-1}\left(\bigcup_{k=0}^{\infty} F^{-k}(I_\infty)\right) = \bigcup_{k=0}^{\infty} F^{-k}(I_\infty)$$

is an invariant set for F and

$$\mu\left(\bigcup_{k=0}^{\infty} F^{-k}(I_\infty)\right) = \lim_{k\to\infty} \mu(F^{-k}(I_\infty)) = \mu(I_\infty)\,.$$

Therefore μ is not ergodic. Similarly if $0 < \mu(P_n) < 1$, then μ is not ergodic. $\square$

Using the above results and some general theorems from ergodic theory, it is possible to give alternative descriptions of the measure ν_∞. The following theorem may be found in [12].

THEOREM 2.2. *Let X be a compact metric space, $T: X \to X$ a continuous map, and assume ν is the unique probability measure on X which is invariant with respect to T. Then for any continuous real valued function f on X.*

$$\frac{1}{N}\sum_{k=0}^{N-1} f(T^k(x)) \to \int f\,d\nu \tag{12}$$

uniformly for all $x \in X$.

Note that by the Birkhoff Ergodic Theorem the convergence in (12) holds for any integrable function f pointwise for almost all $x \in X$. Before applying Theorem 2.2 to F, we cite the following special case of a theorem of Choquet [13].

THEOREM 2.3. *For any* $\mu \in M_F[0, 1]$ *there exists a Borel probability measure* m_μ *on the set of extreme points of* $M_F[0, 1]$ *such that* $\mu = \int \nu m_\mu(d\nu)$.

PROPOSITION 2.5. *For any continuous function* f *on* $[0, 1]$

(a) $\int f(x) \nu_\infty(dx) = \lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} f(F^k(x_0))$ *for all* x_0 *not eventually periodic, and uniformly for all* x_0 *in* I_∞.

(b) $\int f(x) \nu_\infty(dx) = \lim_{n \to \infty} \int f(x) \nu_n(dx) = \lim_{n \to \infty} \frac{1}{2^n} \sum_{k=0}^{2^n - 1} f(F^k(y_n))$ *where* y_n *is the smallest number in* P_n.

Proof. Part (a) follows from Theorem 2.2 with $X = I_\infty$ and Theorem 1.1. To prove part (b) consider the sequence $\{\nu_n\}$ in $M_F[0, 1]$. Since $M_F[0, 1]$ is compact and metrizable in the vague topology any subsequence of $\{\nu_n\}$ has a convergent subsequence. Let $\{\nu_{n_k}\}$ be a subsequence of $\{\nu_n\}$ converging to $\mu \in M_F[0, 1]$. By Proposition 2.4 and Theorem 2.3,

$$\mu = \alpha_0 \nu_0 + \alpha_1 \nu_1 + \alpha_2 \nu_2 + \cdots + \alpha_\infty \nu_\infty$$

for some sequence of nonnegative real numbers $\{\alpha_n\}$ such that $\sum_{n=0}^{\infty} \alpha_n + \alpha_\infty = 1$. For ease of notation, denote by $\nu_n(f)$ the integral $\int f(x) \nu_n(dx)$. Then for any continuous function f on $[0, 1]$,

$$\nu_{n_k}(f) \to \alpha_0 \nu_0(f) + \alpha_1 \nu_1(f) + \alpha_2 \nu_2(f) + \cdots + \alpha_\infty \nu_\infty(f) \tag{13}$$

as $k \to \infty$. Let f be continuous on $[0, 1]$, $f \equiv 0$ outside of the open interval A_0 and $f(y_0) = 1$ (where as before $F(y_0) = y_0$ is the unique fixed point and smallest number in the period 1 orbit). By Theorem 1.1 $\nu_n(f) = 0$ when $n \geqslant 1$ and $\nu_0(f) = 1$. It follows that $\alpha_0 = 0$. Choosing a continuous function f such that $f(y_n) = 1$ and $f \equiv 0$ outside of the open set A_{n-1} and using a similar argument shows that $\alpha_n = 0$ for each integer n. Since μ is a probability measure, it follows that $\alpha_\infty = 1$. Thus $\mu = \nu_\infty$. Since every subsequence of $\{\nu_n\}$ has a subsequence converging to ν_∞, it follows that

$$\nu_n \to \nu_\infty \tag{14}$$

in the vague topology of $M_F[0, 1]$, which is equivalent to part (b) of Proposition 2.5. □

Part (a) of Proposition 2.5 may be understood in an intuitive way. Consider a system with initial "state" $x_0 \in [0, 1]$, whose state at integer time n is given recursively by $x_n = F(x_{n-1})$. Let f be an observable, i.e., a continuous function from $[0, 1]$ to $\mathbf{R}$. Proposition 2.5 (a) then says that the time average $\lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} f(F^k(x_0))$ of f is equal to the "phase space average" $\int f(x) \nu_\infty(dx)$ of f for any initial state x_0 which is not eventually periodic. The identification of time averages with phase space averages is the theoretical foundation of the statistical mechanical derivation of thermodynamics.

One way to measure chaos is to calculate Liapunov exponents. These exponents measure the rate of separation of nearby points under iterations of the map defining a dynamical system. When x and x_0 are close,

$$F^n(x) - F^n(x_0) \approx D_{x_0} F^n \cdot (x - x_0) .$$

If we also require that $F^n(x) - F^n(x_0) \approx \exp(n\lambda(x_0))$ asymptotically as n increases (for x_0 "infinitesimally close" to x), then a natural definition for $\lambda(x_0)$ is

$$\begin{aligned} \lambda(x_0) &\equiv \lim_{n \to \infty} \frac{1}{n} \log \left| D_{x_0} F^n \cdot (x - x_0) \right| \\ &= \lim_{n \to \infty} \frac{1}{n} \log \left| F'(F^{n-1}(x_0)) \cdot F'(F^{n-2}(x_0)) \cdots F'(x_0)(x - x_0) \right| \\ &= \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \log \left| F'(F^k(x_0)) \right| \end{aligned} \tag{15}$$

assuming, of course, that F is differentiable at $F^k(x_0)$ for all $k \geqslant 0$. In our case, F is differentiable at all but countably many points. For a point x in $[0, 1]$ where F fails to be differentiable, let us make the convention that $D_x F \equiv 1$, the smaller in magnitude of the one-sided derivatives. If x_0 is in I_∞, (15) can be calculated directly or via Proposition 2.5 (a) with $f(x) \equiv \log |F'(x)|$ so that

$$\lambda(x_0) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \log \left| F'(F^k(x_0)) \right| = \int \log |F'(x)| \nu_\infty(dx) . \tag{16}$$

In either case, $\lambda(x_0) = 0$ for all x_0 in I_∞ because $F'(x) \equiv 1$ on I_∞. Similar reasoning shows that if x is eventually in the period 2^n orbit of F, i.e., $x \in \bigcup_{k=0}^{\infty} F^{-k}(P_n)$ then

$$\lambda(x) = \frac{1}{2^n} \log \left(\frac{7}{3}\right). \tag{17}$$

There is no generally accepted mathematical definition of a chaotic map. One widely accepted definition [9] requires that the Liapunov exponent be positive. It follows that $F: I_\infty \to I_\infty$ is not chaotic in this sense. It is not difficult to verify that F is also not chaotic according to the definition given in Devaney [7]. The qualitative behavior that nearby points in I_∞ do not separate with increasing iterations of F is also manifested by the fact, established in Theorem 1.1 (a), that F is a permutation on the intervals I_n^k for any fixed n.

A measurable partition of a probability space (X, ν) is a collection

$$\xi = \{B_1, B_2, \ldots, B_n\}$$

of measurable subsets of X whose union is X and which are pairwise disjoint. The entropy $H(\xi)$ is given by

$$H(\xi) = -\sum_{i=1}^{n} \nu(B_i) \log[\nu(B_i)] \tag{18}$$

with the convention that $0 \log 0 = 0$. If $T: X \to X$ is a fixed measurable transformation, define ξ^m to be the measurable partition of X consisting of all sets of the form $B_{i_1} \cap T^{-1}(B_{i_2}) \cap \cdots \cap T^{-m-1}(B_{i_m})$ where $B_{i_k} \in \xi$. For a T-invariant probability measure ν, the entropy $h(\nu, \xi)$ of ν relative to ξ is defined by

$$h(\nu, \xi) = \lim_{m \to \infty} \frac{1}{m} H(\xi^m), \tag{19}$$

where it can be shown that the limit in (19) exists. The entropy $h(\nu)$ is then defined by

$$h(\nu) = \sup h(\nu, \xi) \tag{20}$$

where the supremum is over all measurable partitions of X. The entropy $h(\nu)$ is sometimes called the Kolmogorov-Sinai invariant and it measures the asymptotic rate of creation of information by iterating T. It is invariant under measure preserving isomorphisms. If X is a compact metric space (with the Borel σ-algebra) and T is continuous, then [14]

$$h(\nu) = \lim_{\operatorname{diam}\xi \to 0} h(\nu, \xi)\,, \tag{21}$$

where $\operatorname{diam}\xi = \max_i\{\text{diameter of } B_i \in \xi\}$.

We apply (21) to $F: I_\infty \to I_\infty$ with the invariant probability measure ν_∞. For any positive integer n, let $\xi(n) = \{I_n^1 \cap I_\infty, I_n^2 \cap I_\infty, \ldots, I_n^{2^n} \cap I_\infty\}$. As $n \to \infty$, $\operatorname{diam}\xi(n) \to 0$. By Theorem 1.1, $(\xi(n))^m = \xi(n)$ for every m and

$$\begin{aligned} H(\xi(n)) &= -\sum_{k=1}^{2^n} \nu_\infty(I_n^k \cap I_\infty)\log[\nu_\infty(I_n^k \cap I_\infty)] \\ &= -\sum_{k=1}^{2^n} 2^{-n}\log 2^{-n} \\ &= n\log 2. \end{aligned} \tag{22}$$

Therefore $h(\nu_\infty, \xi(n)) = \lim_{m \to \infty} \frac{1}{m} n \log 2 = 0$ for all n. Thus

$$h(\nu_\infty) = \lim_{n \to 0} h(\nu_\infty, \xi(n)) = 0\,.$$

Essentially the same argument shows that $h(\nu_\infty) = 0$ when we regard $F: [0, 1] \to [0, 1]$. The only modification needed is to add terms to the partition $\xi(n)$ which partition the complement of I_∞ in such a way that the diameters of these zero measure pieces decrease to zero as $n \to \infty$.

A similar calculation shows that $h(\nu_n) = 0$ for $F: [0, 1] \to [0, 1]$ for any nonnegative integer n. Because $[0, 1]$ is a compact metric space, the topological entropy $h_t(f)$ for our map $F: [0, 1] \to [0, 1]$ may be defined [4] as

$$h_t(F) = \sup\{h(\nu) : \nu \text{ is an ergodic } F\text{-invariant probability measure}\}\,.$$

From the above analysis, the topological entropy is clearly zero.

To what extent is the behavior of the function F generic? Let $T: [0, 1] \to [0, 1]$ be a continuous map with zero topological entropy, and let ν be an ergodic T-invariant probability measure on $[0, 1]$ which is not supported on any periodic orbit of T. Misiurewicz [2] pointed out that all periodic orbits of T have periods which are powers of 2. He proved that the dynamical system determined by T and ν on $[0, 1]$ is isomorphic to the adding machine on 2-adic integers explained in Theorem 1.2 together with the measure $\nu_\infty \circ h^{-1}$ (where h is the homeomorphism defined in Theorem 1.2). It follows that the dynamical system determined by T and ν on $[0, 1]$ is therefore isomorphic to the dynamical system determined by F and ν_∞. An example of such a map with zero topological entropy is the function $x^2 - c_0$ discussed in the introduction.