§13. Expanding Maps.

Interesting examples of smooth dynamical systems often display a combination of expansion of distances in some directions, together with contraction of distances in complementary directions. Some expansion is a necessary prerequisite for chaotic behavior, and the interplay between expansion and contraction is a difficult and important topic. As an introduction to such studies, this section will concentrate on the much more restricted class of mappings which display only expansion. (We will also look at the somewhat more general class of covering maps, particularly in dimension one.) We first study "balanced" measures. These describe the distribution of random backward orbits. (See 13.5.) We then study absolutely continuous invariant measures, which describe the distribution of random forward orbits. (See 13.7.) Note that forward and backward orbits behave quite differently. For example, in a dynamical system with both attractors and repellors, forward orbits tend to converge towards the attractors while backward orbits tend to converge towards the repellors.

Suppose that $f : M \to M$ is a C^1-smooth map from a C^1-smooth manifold to itself. Choosing some Riemannian metric, we will denote the length of a tangent vector $v \in TM$ by $\|v\|$, and the derivative of the smooth map f, mapping each vector space T_xM linearly into $T_{f(x)}M$, by $Df : TM \to TM$.

Definition. The C^1-map $f : M \to M$ is called *expanding* if there exist constants $c > 0$ and $k > 1$ so that

$$\|Df^{\circ n}(v)\| \;\geq\; ck^n\|v\|$$

for all tangent vectors v and for all $n > 0$. An equivalent condition would be that for any smooth path $p : [0, 1] \to M$ the length $\mathbf{len}(p) = \int_0^1 \|p'(t)\|\, dt$ satisfies

$$\mathbf{len}(f^{\circ n} \circ p) \;\geq\; ck^n\, \mathbf{len}(p)\ .$$

(Compare §4E.) For other equivalent forms of the definition, see Problem 13-a.

Here the constant c depends on the particular choice of Riemannian metric. However the *expansion constant* k is independent of the metric. For if $\|v\|'$ is a different metric, then by compactness we can choose a constant $a > 1$ so that $\|v\|/a \leq \|v\|' \leq a\|v\|$ for all tangent vectors v. If $\|Df^{\circ n}(v)\| \;\geq\; ck^n\|v\|$, then it follows that $\|Df^{\circ n}(v)\|' \;\geq\; c'\,k^n\|v\|'$ where $c' = (c/a^2)$. (To obtain a differentiable conjugacy invariant, we could take the supremum of all such expansion constants.)

If f is an expanding map, then Df maps each tangent vector space T_xM bijectively onto the vector space $T_{f(x)}M$. Therefore f is locally a diffeomorphism. Since M is compact, it follows that f is a *covering map*. That is, each point of M has a connected neighborhood U so that each connected component of the preimage $f^{-1}(U)$ maps homeomorphically onto U. If M is connected, then denoting the number of such connected components by q, it follows easily that the number q is constant and finite, so that f is exactly q-to-one. Note that $f^{\circ n}$ is locally distance increasing, and hence locally volume increasing, for n sufficiently large. Integrating over M, it follows that f cannot be one-to-one, hence $q \geq 2$.

§13A. Balanced Measures. First consider an arbitrary q-fold covering map $f : X \to Y$ with $q \geq 2$, where X and Y are locally connected compact metric spaces. As in §10, let $\mathcal{M}(X)$ be the convex set consisting of all Borel probability measures on X.

Then any measure $\nu \in \mathcal{M}(Y)$ gives rise to a "pull-back" measure $\mu = f^*\nu \in \mathcal{M}(X)$. To illustrate the idea, if $\nu = \delta_y$ is the Dirac measure conventrated at a point of Y, then $f^*\delta_y$ will be the average of the Dirac measures concentrated at the q points of $f^{-1}(y)$. The definition can be given as follows.

Lemma 13.1. *Given a q-fold covering $f : X \to Y$, and given a probability measure $\nu \in \mathcal{M}(Y)$, there is one and only one probability measure $\mu \in \mathcal{M}(X)$ which satisfies any one of the following three equivalent conditions. Furthermore, the push forward $f_*(\mu)$ of this measure μ is equal to ν.*

- *If the connected measurable set $U \subset Y$ is evenly covered, then the components U_i of $f^{-1}(U)$ satisfy*

$$\mu(U_1) \;=\; \cdots \;=\; \mu(U_q) \;=\; \nu(U)/q\ . \tag{13:1}$$

- *If S is any measurable subset of X such that f restricted to S is one-to-one, then*

$$\mu(S) \;=\; \nu(f(S))/q\ . \tag{13:2}$$

- *For any continuous $\varphi : X \to \mathbb{R}$ we have*

$$\int_X \varphi\, d\mu \;=\; \int_Y (f_*\varphi)\, d\nu\ , \tag{13:3}$$

where the function $f_*\varphi : Y \to \mathbb{R}$ is defined as the average

$$f_*\varphi(y) \;=\; \frac{1}{q}\sum_{f(x)=y} \varphi(x) \tag{13:4}$$

of the values of φ at the pre-images of y.

Thus we have linear maps $\mathcal{M}(Y) \xrightarrow{f^*} \mathcal{M}(X) \xrightarrow{f_*} \mathcal{M}(Y)$ with composition equal to the identity map of $\mathcal{M}(Y)$.

Definition. In the special case of a self-covering map $f : X \to X$, a probability measure β on X is said to be *balanced* if $f^*\beta = \beta$. Note that a balanced measure is necessarily f-invariant, since $f_*\beta = f_*(f^*\beta) = \beta$.

Proof of Lemma 13.1. Using the Riesz Representation Theorem 10.8, we can take (13 : 3) as the definition of μ. Existence and uniqueness are then clear. Thus it only remains to show that the first two characterizations of μ are equivalent to (13 : 3).

Starting with (13 : 3), we can verify (13 : 2) as follows. Approximating the characteristic function $\mathbf{1}_S$ by continuous maps as in the proof of 10.4, it follows easily that

$$\mu(S) \;=\; \int_X \mathbf{1}_S\, d\mu \;=\; \int_Y (f_*\mathbf{1}_S)\, d\nu \;=\; \int_Y \mathbf{1}_{f(S)}\, d\nu/q \;=\; \nu(f(S))/q\ ,$$

which proves (13 : 2). Conversely, given (13 : 2) we can partition X into small sets on which φ is nearly constant, and approximate the continuous function φ by the associated step function, to prove (13 : 3).

The formula (13 : 2) clearly implies (13 : 1). Conversely, given (13 : 1), after partitioning S into small sets, we can approximate each one by a neighborhood N, and then apply (13 : 1) to each connected component of $f(N)$ to verify (13 : 2). $\square$

Theorem 13.2. *An expanding map from a compact connected manifold M to itself has a unique balanced probability measure β. This measure is ergodic and mixing, and satisfies $\beta(U) > 0$ for every non-empty open set U.*

The proof will be based on the following statement. Given a continuous function $\varphi : M \to \mathbb{R}$ let us iterate the averaging operator $\varphi \mapsto f_*\varphi$ of $(13:4)$.

Theorem 13.3. *If $f : M \to M$ is an expanding map on a connected manifold, then for any continuous $\varphi : M \to \mathbb{R}$ the averaged functions*

$$(f_*)^{\circ n}\varphi(x) \;=\; (f^{\circ n})_*\varphi$$

converge uniformly to a constant function as $n \to \infty$. Hence, starting with any probability measure $\mu \in \mathcal{M}(M)$, the pulled back measures $(f^*)^{\circ n}\mu = (f^{\circ n})^*\mu$ converge in the weak topology to a measure β which is independent of the choice of μ.*

This in turn will be proved using the following geometric construction.

Theorem 13.4. *If $f : M \to M$ is an expanding map of a connected manifold, then there exists a sequence of partitions $\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2, \ldots$ with the following properties. Each $\mathcal{P}_n$ is a collection of q^n disjoint measurable sets with union M. Furthermore, for $n > 0$, f maps each $P \in \mathcal{P}_n$ bijectively onto a set $f(P) \in \mathcal{P}_{n-1}$. In particular, f maps each $P \in \mathcal{P}_1$ bijectively onto the entire manifold M, which is the unique element of the partition $\mathcal{P}_0$. Finally, the* mesh *of $\mathcal{P}_n$, that is the maximum of the diameters of its elements, tends to zero as $n \to \infty$.*

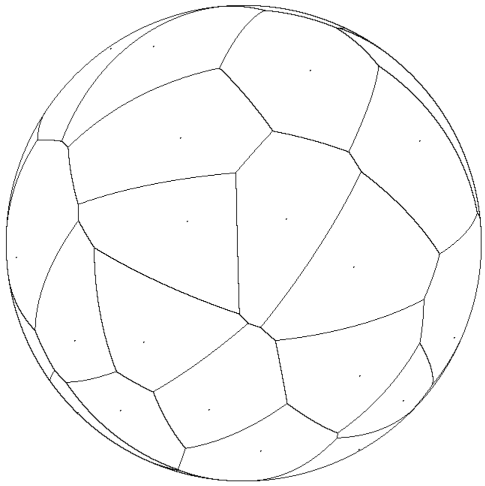

Figure 42. Voronoĭ cells in the 2-sphere

Proof of 13.4. We will work with some subdivision of M into closed topological cells. For example a smooth triangulation would certainly suffice. (Compare [Munkres 1966].) However, a subdivision into "Voronoĭ cells" (also called "Dirichlet cells") is much easier to construct and will suffice for our purpose. Let us assume that M is a C^3-differentiable manifold which has been provided with a C^2-differentiable Riemannian metric. Then there exists an $\epsilon > 0$ so that any two points with Riemannian distance $\mathrm{dist}(x, y) < \epsilon$ are joined by a unique minimal geodesic which depends smoothly on these endpoints. (See for example [Milnor 1963].) Choose finitely many points $x_1, x_2, \ldots, x_n \in M$ so that the open balls of radius $\epsilon/2$ centered at these points cover the manifold. By definition, the j-th associated *Voronoĭ cell* is the set $\overline{V}_j$ consisting of all $x \in M$ which satisfy

$$\mathrm{dist}(x, x_j) \le \mathrm{dist}(x, x_h) \qquad \text{for all} \quad h,$$

so that no other x_h is closer to x than x_j. Then it is easy to check that the $\overline{V}_j$ are non-overlapping closed topological cells, and that the union of the $\overline{V}_j$ is equal to M.

We will also construct a tree $T_0 \subset M$ with the points x_h as vertices. It will be convenient to inductively renumber the points x_h, as follows. For each $j > 1$, assuming inductively that the set $\{x_1, \ldots, x_{j-1}\}$ has already been renumbered, choose x_j so as to minimize the distance between x_j and $\{x_1, \ldots, x_{j-1}\}$. Furthermore, choosing $i(j) < j$ so as to minimize the distance between $x_{i(j)}$ and x_j, let E_j be the minimal geodesic line segment from $x_{i(j)}$ to x_j. This exists and is unique, since it is easy to check that $\mathrm{dist}(x_{i(j)}, x_j) < \epsilon$. It then follows easily that the midpoint of E_j belongs to both $\overline{V}_{i(j)}$ and $\overline{V}_j$. The union $E_2 \cup \cdots \cup E_N$ is the required tree T_0.

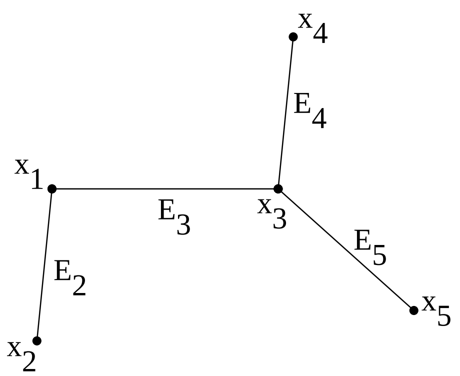

Figure 43. Contruction of the tree T_0.

In order to obtain a partition of M into disjoint sets we will remove some duplicated boundary points from the $\overline{V}_j$. To do this, set $V_1 = \overline{V}_1$, and set

$$V_j = (\overline{V}_1 \cup \overline{V}_2 \cup \cdots \overline{V}_j) \smallsetminus (\overline{V}_1 \cup \overline{V}_2 \cup \cdots \overline{V}_{j-1}) \qquad \text{for} \quad j > 1.$$

Evidently the V_j are disjoint sets with union M.

Since T_0 is simply connected, each of the q^n connected components of $f^{-n}(T_0)$ is again a tree which maps homeomorphically onto T_0. Number these connected components as T_n^α, where $1 \le \alpha \le q^n$. Siimilarly, each V_j is simply connected, so $f^{-1}(V_j)$ is the union of 2^q disjoints sets, each of which maps homeomorphically onto V_j. Let $V_{n,j}^\alpha$ be the unique component of $f^{-n}(V_j)$ which intersects T_n^α, and let P_n^α be the union

$$P_n^\alpha = V_{n,1}^\alpha \cup V_{n,2}^\alpha \cup \cdots \cup V_{n,N}^\alpha .$$

Thus P_n^α is a neighborhood of T_n^α. For each fixed n, it is not difficult to check that these sets P_n^α constitute a partition $\mathcal{P}_n$ of M into q^n disjoint connected measurable sets.

Since $f^{\circ n} : V_j^\alpha \xrightarrow{\cong} V_j$, and since any two points of V_j are connected by a geodesic of length less than ϵ, it follows that any two points of V_n^α are joined by a path of length at most ϵ/ck^n. Hence any two points of P_n^α are joined by a path of length at most $N\epsilon/ck^n$. This tends to zero as $n \to \infty$, which completes the proof of Theorem 13.4. $\square$

Proof of 13.3. We will need two basic properties of the transformation $\varphi \mapsto f_*\varphi$. First, for a composition $f \circ g$ of covering maps, note that

$$(f \circ g)_* \;=\; f_* \circ g_* \,. \tag{13 : 5}$$

Second, note that any upper and lower bounds for φ yield the same upper and lower bounds for $f_*\varphi$,

$$c_1 \le \varphi(x) \le c_2 \quad \text{for all } x \qquad \Longrightarrow \qquad c_1 \le f_*\varphi(y) \le c_2 \quad \text{for all } y \,. \tag{13 : 6}$$

This is clear, since f_* is an averaging operator.

Let $\varphi : M \to \mathbb{R}$ be a continuous function, and let y be an arbitrary point of M. Evidently the set of $x \in M$ with $f^{\circ n}(x) = y$ contains exactly one point in each partition element $P \in \mathcal{P}_n$. Since the mesh of $\mathcal{P}_n$ tends to zero, it follows by uniform continuity that the difference

$$\max_{x\in P} \varphi(x) \;-\; \min_{x\in P} \varphi(x) \,, \qquad \text{where} \qquad P \in \mathcal{P}_n \,,$$

tends uniformly to zero as $n \to \infty$. Hence the difference between the numbers

$$\frac{1}{q^n} \sum_{P\in\mathcal{P}_n} \min_{x\in P} \varphi(x) \;\;\le\;\; \frac{1}{q^n} \sum_{f^{\circ n}(x)=y} \varphi(x) \;\;\le\;\; \frac{1}{q^n} \sum_{P\in\mathcal{P}_n} \min_{x\in P} \varphi(x)$$

tends uniformly to zero as $n \to \infty$. Since the middle expression equals $f_*^{\circ n}\varphi$, this together with (13 : 6) proves that the $f_*^{\circ n}\varphi$ converge uniformly to some constant function $x \mapsto c(\varphi)$.

For any $\mu \in \mathcal{M}(M)$, it follows that the integrals

$$\int \varphi \, d(f^{*\circ n}\mu) \;=\; \int (f_*^{\circ n}\varphi) \, d\mu \;\approx\; \int c(\varphi) \, d\mu \;=\; c(\varphi)$$

converge to $c(\varphi)$ as $n \to \infty$. If we define a measure $\beta \in \mathcal{M}(M)$ by the equation

$$\int \varphi \, d\beta \;=\; c(\varphi)$$

for every continuous φ, then this argument proves that the sequence of measures $f^{*\circ n}\mu$ converges weakly to β; which completes the proof of 13.3. $\square$

Proof of Theorem 13.2. Clearly this measure $\beta = \lim_{n\to\infty}(f^*)^{\circ n}\mu$ satisfies $f^*\beta = \beta$, and hence is balanced. Now take μ to be an arbitrary balanced measure. Then $f^*\mu = \mu$ so $\mu = \lim (f^*)^{\circ n}\mu$, which proves that $\mu = \beta$.

Ergodicity follows from uniqueness. For if $T \subset M$ were a fully invariant Borel set with $0 < \beta(T) < 1$, then the measure

$$\mu(S) \;=\; \beta(S \cap T)/\beta(T)$$

would be balanced and different from β.

Every non-vacuous open set has measure $\beta(U) > 0$, since every such set contains some partition element $P \in \mathcal{P}_n$ with measure $\beta(P) = 1/q^n > 0$.

Recall from Problem 8-d that the transformation f is called *mixing* with respect to a probability measure β, if we have

$$\lim_{n\to\infty} \beta\Big(S \cap f^{-n}(T)\Big) \;=\; \beta(S)\,\beta(T) \qquad (13:7)$$

for every pair of measurable sets S and T. Using the partitions of 13.4, for any $\epsilon > 0$ we can choose a compact set $K = K_\epsilon \subset S$ and an open set $U = U_\epsilon \supset S$ so that $\beta(U \smallsetminus K) < \epsilon$. Let $K' = M \smallsetminus U$. If n is large enough so that the mesh of $\mathcal{P}_n$ is less than the distance between K and K', then no $P \in \mathcal{P}_n$ can intersect both K and K'. Let $N(n) \le q^n$ be the number of sets $P \in \mathcal{P}_n$ which intersect K. Then

$$\beta(K) \;\le\; N(n)/q^n \;\le\; \beta(K) + \epsilon\,.$$

Thus, as $\epsilon \to 0$ and $n \to \infty$, the ratio $N(n)/q^n$ must tend to $\beta(S)$. Now note that $f^{\circ n}$ carries each $P \in \mathcal{P}_n$ bijectively onto M, carrying $P \cap f^{-n}(T)$ onto T, and multiplying measures by q^n, so that

$$\beta(P \cap f^{-n}(T)) \;=\; \beta(T)/q^n\,.$$

Summing over the $N(n)$ sets P which intersect K, it follows that

$$\beta\Big(K \cap f^{-n}(T)\Big) \;\le\; N(n)\,\beta(T)/q^n \;\le\; \beta\Big(K \cap f^{-n}(T)\Big) + \epsilon\,,$$

where the expression in the middle converges towards $\beta(T)\,\beta(S)$. Passing to the limit as $\epsilon \to 0$, since

$$\beta\Big(K \cap f^{-n}(T)\Big) \;\le\; \beta\Big(S \cap f^{-n}(T)\Big) \;\le\; \beta\Big(K \cap f^{-n}(T)\Big) + \epsilon\,,$$

it follows that

$$\lim_{n\to\infty} \beta\Big(S \cap f^{-n}(T)\Big) \;=\; \beta(T)\,\beta(S)\,,$$

as required. This completes the proof of 13.2. $\square$

Random Backward Orbits. We will next describe a computationally effective procedure for approximating the balanced measure β. (Compare [Elton].) I am indebted to Jane Hawkins for suggesting this algorithm in the closely related case of rational maps of the Riemann sphere.) Starting with some arbitrary point $x_0 \in M$, construct a "random backward orbit"

$$x_0 \;\leftarrow\!\shortmid\; x_1 \;\leftarrow\!\shortmid\; x_2 \;\leftarrow\!\shortmid\; \cdots$$

inductively, so that for each n any one of the q preimages of x_n will be chosen as x_{n+1} with probability $1/q$. These choices should be independent, so that each of the q^n possibilities for x_n occurs with probability $1/q^n$. Let δ_{x_n} be the Dirac measure concentrated at the point x_n.

Lemma 13.5. *If f is an expanding map, and if the x_n are chosen as above, then with probability one the sequence $x_1, x_2, \ldots$ is evenly distributed with respect to the unique balanced probability measure β. In other words, with probability one, the sequence of averaged measures*

$$\mu_k \;=\; (\delta_{x_1} + \delta_{x_2} + \cdots + \delta_{x_k})/k$$

converges weakly to the measure β as $k \to \infty$.

Proof. Choose the sequence of partitions $\mathcal{P}_1, \mathcal{P}_2, \ldots$ as in 13.4. First choose some set $P \in \mathcal{P}_1$. Then each randomly chosen x_k with $k > 0$ has probability $1/q$ of belonging to

P , and these events are independent. Hence with probability one, by Borel's law of large numbers, the measure

$$\mu_k(P) \;=\; \big(\delta_{x_1}(P) + \cdots + \delta_{x_k}(P)\big)/k$$

converges to $\beta(P) = 1/q$ as $k \to \infty$. Applying this same argument to the map $f^{\circ n}$ and a partition element $P \in \mathcal{P}_n$ we see for $1 \le j \le n$ that each of the orbit elements

$$x_j \overset{f^{\circ n}}{\longleftarrow} x_{j+n} \overset{f^{\circ n}}{\longleftarrow} x_{j+2n} \overset{f^{\circ n}}{\longleftarrow} \cdots$$

independently has probability $1/q^n$ of being in P . Arguing as above and then averaging over j , it follows easily that with probability one

$$\lim_{k\to\infty} \mu_k(P) \;=\; \beta(P) \;=\; 1/q^n \qquad \text{for} \qquad P \in \mathcal{P}_n \,.$$

Finally, approximating an arbitrary continuous $\varphi : M \to \mathbb{R}$ by a step function which is constant on each $P \in \mathcal{P}_n$, and then letting $n \to \infty$, we see that

$$\lim_{k\to\infty} \int \varphi(x)\, d\mu_k(x) \;=\; \int \varphi(x)\, d\beta(x)$$

with probability one, as required. □

§**13B. Circle Covering Maps.** In the 1-dimensional case, most of these results are true without assuming that f is expanding or differentiable. (Compare §4E.) If $f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ is any continuous map of the circle, then there exists an essentially unique ***lift*** $F : \mathbb{R} \to \mathbb{R}$ so that the following diagram is commutative

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{F} & \mathbb{R} \\ \downarrow p & & \downarrow p \\ \mathbb{R}/\mathbb{Z} & \xrightarrow{f} & \mathbb{R}/\mathbb{Z}\,, \end{array}$$

where p denotes the natural projection map from $\mathbb{R}$ onto $\mathbb{R}/\mathbb{Z}$. More precisely, F is unique up to the addition of an integer constant. Since we must have

$$F(x+1) \;\equiv\; F(x) \pmod{\mathbb{Z}}$$

in order to make this diagram commutative, it follows that

$$F(x+1) \;=\; F(x) \,+\, d$$

where d is an integer constant called the ***degree*** of f . As an example, the map $x \mapsto xd$ has degree d . If F is a homeomorphism of $\mathbb{R}$, or equivalently if f is locally a homeomorphism, then evidently f is a q-fold covering map, where $q = |d| > 0$. Note that F is either monotone increasing or monotone decreasing according as $d > 0$ or $d < 0$.

Theorem 13.6. *If $f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ is a q-fold covering map with $q \ge 2$, then f has a unique balanced probability measure β , which is necessarily ergodic and mixing. Furthermore, as in* 13.5*, a random backward orbit is β-evenly distributed.*

Note however that we do *not* claim that every non-empty interval $I \subset \mathbb{R}/\mathbb{Z}$ has measure $\beta(I) > 0$. For example the map f may well have an attracting fixed point, and hence a non-trivial interval I with $I \supset f(I)$. Then $\beta(I) \ge \beta(f(I)) = q\,\beta(I)$ by (13 : 2), which implies that any such interval must have measure $\beta(I) = 0$.

Proof of 13.6. These assertions can all be proved by slight modifications of the arguments given above. Choosing some base point $x_0 \in \mathbb{R}/\mathbb{Z}$, we can form the partitions $\mathcal{P}_n$ of $\mathbb{R}/\mathbb{Z}$ where each $J \in \mathcal{P}_n$ is a half-open interval bounded by two of the points of $f^{-n}(x_0)$, and where $f^{\circ n}$ maps the interior of each each $J \in \mathcal{P}_n$ homeomorphically onto $\mathbb{R}/\mathbb{Z} \smallsetminus \{x_0\}$. Clearly most of these q^n intervals $J \in \mathcal{P}_n$ must be reasonably short. (The key difference which makes the one dimensional case easier is that every connected set with small Lebesgue measure must also have small diameter.)

As in 13.3, we will show, for any continuous $\varphi : \mathbb{R}/\mathbb{Z} \to \mathbb{R}$, that the sequence of averaged functions

$$f_*^{\circ n}(y) \;=\; \frac{1}{q^n} \sum_{f^{\circ n}(x)=y} \varphi(x)$$

converges uniformly to a constant function. Clearly the set $f^{-n}(y) \subset \mathbb{R}/\mathbb{Z}$ contains exactly one point x_J from each half-open interval $J \in \mathcal{P}_n$. Let us call such an interval J either *long* or *short* according as its length satisfies $\mathbf{len}(J) \geq 1/\sqrt{q^n}$ or $\mathbf{len}(J) < 1/\sqrt{q^n}$. Clearly there are at most $\sqrt{q^n}$ long intervals. Hence the sum of $\varphi(x_J)$ over the long intervals J is bounded by $C\sqrt{q^n}/q^n = C/\sqrt{q^n}$, where C is an upper bound for $|\varphi(x)|$. Since $C/\sqrt{q^n}$ tends to zero as $n \to \infty$, we can ignore all of these terms. On the other hand for short intervals, by uniform continuity, the difference

$$\max_{x\in J} \varphi(x) \;-\; \min_{x\in J} \varphi(x)$$

tends uniformly to zero as $n \to \infty$. It follows that the difference between the upper and lower Riemann sums

$$S_n^+ \;=\; \frac{1}{q^n} \sum_{J\in\mathcal{P}_n} \max_{x\in J} \varphi(x) \qquad \text{and} \qquad S_n^- \;=\; \frac{1}{q^n} \sum_{J\in\mathcal{P}_n} \min_{x\in J} \varphi(x)$$

tends to zero as $n \to \infty$. Since the values of $f_*^{\circ n}\varphi(y)$ are bounded between these two Riemann sums, it follows using $(13:6)$ that $f_*^{\circ n}\varphi$ converges uniformly to a constant function $y \mapsto c(\varphi)$. As in the proof of 13.2, it then follows that f has a unique balanced measure β, where $\int \varphi\, d\beta = c(\varphi)$; and it also follows that this measure is ergodic.

The proof that β is mixing goes through much as before. More precisely, the short intervals $J \in \mathcal{P}_n$ behave just as before, and we can simply ignore the long intervals. Similarly, the proof that a random backwards orbit is evenly distributed with probability one goes through much as before. $\square$

Alternatively, once β has been constructed, we can construct a monotone semiconjugacy from f to the linear map $t \mapsto t\,d$ as follows. (Compare 4.12.) First note that we can choose a fixed point $x_0 = f(x_0)$. In fact any circle map of degree $d \neq 1$ must have at least $|d-1| > 0$ fixed points, since the map $x \mapsto F(x) - x$ increases by $d-1$ over the interval $0 \leq x \leq 1$, and hence must take an integer value at least $|d-1|$ times.

Now define a monotone function $m : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ by setting

$$m(x) \;=\; \int_{x_0}^{x} d\beta \qquad \text{for} \qquad x_0 \leq x \leq x_0 + 1\,.$$

This yields a well defined degree one map from $\mathbb{R}/\mathbb{Z}$ to itself since $\int_{x_0}^{x_0+1} d\beta = 1$. Since f carries each short interval of measure $\beta(I)$ to an interval of length $q\,\beta(I) = |d|\,\beta(I)$, it follows easily that $m(f(x)) = m(y)\,d$. In other words, the following diagram is commutative:

$$\begin{array}{ccc} \mathbb{R}/\mathbb{Z} & \xrightarrow{f} & \mathbb{R}/\mathbb{Z} \\ \downarrow m & & \downarrow m \\ \mathbb{R}/\mathbb{Z} & \xrightarrow{d\cdot} & \mathbb{R}/\mathbb{Z}\,. \end{array} \qquad (13:8)$$

We say briefly that m *semiconjugates* f to multiplication by d . Now, for any interval $I \subset \mathbb{R}/\mathbb{Z}$, it follows easily that $\beta(I)$ is just the length (= Lebesgue measure) of the image $m(I) \subset \mathbb{R}/\mathbb{Z}$. In particular, the push forward $m_*\beta$ is equal to Lebesgue measure.

Since m is monotone of degree one, the preimage $m^{-1}(t)$ of any point is either a point or a closed interval I of measure $\beta(I) = 0$. There can be at most countably many such I . Thus the measure preserving map m from $(\mathbb{R}/\mathbb{Z}, \beta)$ to $(\mathbb{R}/\mathbb{Z}, \text{Lebesgue})$ is one-to-one except on a fully invariant set of measure zero. Now since ergodicity, mixing, and the even distribution of random backward orbits are verified for the expanding map $t \mapsto t\,d$ by 13.2 and 13.5, it follows easily that they are true for the map f also.

Remark. In the differentiable case, if $f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ is an expanding map (for example if $|f'(x)| > 1$ everywhere), then every non-trivial interval has measure $\beta(I) > 0$, and it follows that m is a homeomorphism from the circle to itself. In this case we say that m is a *topological conjusugacy* or that the map f is *topologically conjugate* to multiplication by the degree d . However, even if f is an infinitely differentiable expanding map, we should not expect m to be a diffeomorphism from the circle to itself. (Compare Problem 13-b below.)

§13C. Absolutely Continuous Invariant Measures. Any C^1-smooth manifold has a preferred measure class $[\lambda]$ which reduces to the Lebesgue measure class under any C^1-coordinate chart. To find a global measure in this class, it is convenient to choose a Riemannian metric, which in classical notation is written in terms of local coordinates $x^1, \ldots, x^r$ as $ds^2 = \sum g_{i\,j}(x^1, \ldots, x^r)\,dx^i\,dx^j$. If M is compact, using a partition of unity, the Riemann integrals $\int\cdots\int \varphi\sqrt{|\det(g_{i\,j})|}\,dx^1\cdots dx^r$ in local coordinates then piece together to yield an order preserving linear map $C(M,\mathbb{R}) \to \mathbb{R}$, and hence a finite measure λ on $(M, \mathcal{B}_M)$. We will always assume that λ has been normalized so that $\lambda(M) = 1$.

A finite measure μ on $(M, \mathcal{B}_M)$ is said to be *absolutely continuous* with respect to the Lebesgue measure class $[\lambda]$ if

$$\lambda(S) = 0 \qquad \Longrightarrow \qquad \mu(S) = 0$$

for all $S \in \mathcal{B}_M$. (Equivalently, by the Radon-Nikodým Theorem, μ is absolutely continuous with respect to $[\lambda]$ if and only if there exists a *density function* $\rho \ge 0$ in $L^1_\lambda(M,\mathbb{R})$ so that

$$\mu(S) = \int_S \rho(x)\,d\lambda(x) \qquad \text{for every} \qquad S \in \mathcal{B}_M\ .)$$

> **Theorem 13.7.** *If f is a C^2-smooth expanding map of the compact connected Riemannian manifold M , then the Lebesgue measure class $[\lambda]$ is ergodic, and contains a unique f-invariant probability measure μ .*

Evidently μ can also be described as the unique invariant probability measure which is absolutely continuous with respect to $[\lambda]$. Combining this result with the Birkhoff Ergodic Theorem, we obtain the following.

Corollary 13.8. *With* $f : M \to M$ *as in* 13.7*, Lebesgue almost every orbit of* f *is evenly distributed with respect to the invariant measure* μ *.*

The proof of 13.7 will show also that there exists a constant $K > 1$ so that

$$\lambda(S)/K \;\le\; \mu(S) \;\le\; \lambda(S)\,K \tag{13 : 9}$$

for every Borel set S . It will be convenient to use the abbreviated notation

$$\lambda(S) \;\asymp\; \mu(S) \;,$$

meaning that there exists some uniform constant K so that the ratio $\lambda(S)/\mu(S)$ is bounded between $1/K$ and K . (It is worth noting that one single constant K will be valid for all of the inequalities of this form which we consider.) The proof will depend on two lemmas.

If we choose an orthonormal basis for the tangent spaces T_xM and $T_{f(x)}M$, then the absolute value of the Jacobian determinant of $Df_x : T_xM \to T_{f(x)}M$ is a well defined real number, independent of the choice of bases. Denoting this number by $Jf(x) > 0$, note the identity

$$\lambda\big(f(S)\big) \;=\; \int_S Jf(x)\, d\lambda(x) \;, \tag{13 : 10}$$

which holds for any Borel set S which maps injectively under f .

Lemma 13.9. Bounded Distortion Lemma. *Choosing partitions* $\mathcal{P}_n$ *as in* 13.4*, there exists a constant* $K > 1$ *which is independent of* n *, so that the ratio of* $Jf^{\circ n}(x)$ *to* $Jf^{\circ n}(y)$ *satisfies*

$$\frac{1}{K} \;\le\; \frac{Jf^{\circ n}(x)}{Jf^{\circ n}(y)} \;\le\; K \tag{13 : 11}$$

whenever x *and* y *belong to the same set* $P \in \mathcal{P}_n$ *. We will write briefly*

$$Jf^{\circ n}(x) \;\asymp\; Jf^{\circ n}(y) \qquad \text{for} \quad x, y \in P \in \mathcal{P}_n \;.$$

It follows that

$$Jf^{\circ n}(x) \;\asymp\; \frac{1}{\lambda(P)} \;\asymp\; \frac{\lambda(S)}{\lambda\big(P \cap f^{-n}(S)\big)} \tag{13 : 12}$$

for any set S *of positive measure and any* $x \in P \in \mathcal{P}_n$ *.*

Proof. We will work with the quantity $\log Jf(x)$ which satisfies the chain rule

$$\log Jf^{\circ n}(x_0) \;=\; \sum_{i=0}^{n-1} \log Jf(x_i) \qquad \text{for any orbit} \qquad x_0 \mapsto x_1 \mapsto \cdots \mapsto x_n \;.$$

Since f is assumed to be C^2-differentiable, and since the Riemannian metric can be chosen to be C^1-smooth, it follows that the function $x \mapsto \log Jf(x)$ is C^1-smooth. Since M is compact, we can choose some constant $\kappa > 0$ so that

$$|\log Jf(x) \;-\; \log Jf(y)| \;\le\; \kappa\, \mathrm{dist}(x, y) \;,$$

where $\mathrm{dist}(x, y)$ is the Riemannian distance. Recall from the proof of 13.4 that there are constants $c' = \epsilon N/c > 0$ and $k > 1$ so that each set $P \in \mathcal{P}_n$ has diameter at most c'/k^n . Thus if x_0 and y_0 belong to the same set of the partition $\mathcal{P}_n$, then x_i and y_i belong to the same set of $\mathcal{P}_{n-i}$, and it follows that

$$\Big|\log Jf^{\circ n}(x) \;-\; \log Jf^{\circ n}(y)\Big| \;\;\le\;\; \sum_{i=0}^{n-1}\Big|\log Jf(x_i) - \log Jf(y_i)\Big| \;\;\le\;\; \sum_{i=0}^{n-1}\frac{\kappa\, c'}{k^{n-i}} \;<\; \sum_{j=1}^{\infty}\frac{\kappa\, c'}{k^j}\;,$$

where $j = n - i$. Setting the right side equal to $\log K$, the inequality (13 : 11) follows. Now choosing any base point $x_0 \in P$, and integrating the inequality $Jf^{\circ n}(x_0) \asymp Jf^{\circ n}(x)$ over P , it follows from (6) that

$$Jf^{\circ n}(x_0)\,\lambda(P) \;\;\asymp\;\; \int_P Jf^{\circ n}(x)\;d\lambda(x) \;\;=\;\; \lambda(M) \;\;=\;\; 1\;,$$

since $f^{\circ n}$ maps P bijectively onto M . In other words,

$$Jf^{\circ n}(x) \;\;\asymp\;\; 1/\lambda(P) \qquad \text{for} \quad x \in P\;.$$

More generally, given any Borel set S , if we integrate over the set $P\cap f^{-n}(S)$ which maps bijectively onto S , then the same argument will yield the identity

$$Jf^{\circ n}(x_0)\;\lambda(P\cap f^{-n}(S)) \;\;\asymp\;\; \lambda(S)\;.$$

The required inequality (13 : 12) now follows. □

Proof of 13.7: Existence. Let $\lambda_n = f_*^{\circ n}(\lambda)$ be the n-th push-forward of the Riemannian measure $\lambda = \lambda_0$. We will first prove that $\lambda_n(S) \asymp \lambda(S)$ or more precisely

$$\lambda(S)/K \;\le\; \lambda_n(S) \;\le\; \lambda(S)\,K \tag{13 : 13}$$

for any Borel set S , where the constant K , as in (13 : 11), is independent of n . For each $P^\alpha \in \mathcal{P}_n$, it will be convenient to set $S^\alpha = P^\alpha \cap f^{-n}(S)$, so that $f^{-n}(S)$ is the disjoint union of the S^α . Then

$$\lambda(S^\alpha) \;\;\asymp\;\; \lambda(P^\alpha)\,\lambda(S)$$

by (13 : 12). Summing over the elements $P^\alpha \in \mathcal{P}_n$, it follows that

$$\lambda_n(S) \;\;=\;\; \lambda\Big(f^{-n}(S)\Big) \;\;=\;\; \sum\lambda(S^\alpha) \;\;\asymp\;\; \lambda(S)\;,$$

as asserted. Now set $\mu_n = \Big(\lambda_1 + \cdots + \lambda_n\Big)/n$, and choose a subsequence $\{n_j\}$ so that the sequence of measures $\{\mu_{n_j}\}$ converges weakly to some limit μ , Then it follows as in 13.2 that μ is an invariant measure, $f_*(\mu) = \mu$. The inequality $\lambda(S)/K \le \mu_n(S) \le \lambda(S)\,K$ follows from (13 : 13), and it is not difficult to pass to the weak limit, so as to obtain the corresponding inequality (13 : 9). □

It will be convenient to use the notation $P_n(x) \in \mathcal{P}_n$ for the unique set in the partition $\mathcal{P}_n$ which contains some given point x . We will say that x is a *point of density* for the Borel set S , the measure λ , and the sequence $\{P_n(x)\}$ if the limit

$$\lim_{n\to\infty}\frac{\lambda\Big(S\cap P_n(x)\Big)}{\lambda\Big(P_n(x)\Big)}$$

exists and is equal to one. We will need the following version of the Lebesgue Density Theorem.

13.10. Density Lemma. *Given a finite measure ν on $(X, \mathcal{B}_X)$, a sequence of partitions $\mathcal{P}_n$ with mesh tending to zero, and a Borel set S of positive measure,* *there exists a subsequence* $\{n_j\}$ *so that* $[\nu]$*-almost every point* $x \in S$ *is a point of density for the sequence of neighborhoods* $\{P_{n_j}(x)\}$ *.*

Proof. Fixing some number $0 < \xi < 1$ and some $n > 0$, let us define a partition element $P \in \mathcal{P}_n$ to be "good" if $\nu(P \cap S) \geq \xi\, \nu(P)$, but "bad" if $0 < \nu(P \cap S) < \xi\, \nu(P)$. For any $\epsilon > 0$, we can choose $K \subset S \subset U$ with $\nu(U \smallsetminus K) < \epsilon$ as in the proof of 3.2, and choose n large enough so that any $P \in \mathcal{P}_n$ which intersects K is contained in U . Let B_n be the sum of the measures $\nu(P)$ of those bad P which intersect K , and let G_n be the sum of the measures of the good P which intersect K . It then follows that

$$\nu(K) \;\leq\; \xi\, B_n + G_n \;\leq\; B_n + G_n \;\leq\; \nu(K) + \epsilon\ ,$$

hence $B_n + G_n \leq \xi\, B_n + G_n + \epsilon$, which implies that

$$B_n \;\leq\; \frac{\epsilon}{1-\xi}\ .$$

For fixed ξ , this tends to zero as $\epsilon \to 0$ and hence $n \to \infty$. Similarly, if B'_n is the measure of the set of points $x \in S$ such that the set $P_n(x) \in \mathcal{P}_n$ which contains it is bad, then $B'_n \leq B_n + \epsilon$ also tends to zero as $n \to \infty$. Thus we can choose a sequence of numbers $n(k)$ so that the series $\sum_k B'_{n(k)}$ is convergent. It then follows easily that, for $[\nu]$-almost every point $x \in S$, we have

$$\frac{\nu\Big(P_{n(k)}(x) \cap S\Big)}{\nu\Big(P_{n(k)}(x)\Big)} \;\geq\; \xi$$

for all sufficiently large k . Choosing a sequence of ξ_i converging to one, and choosing corresponding subsequences of $\{n(k)\}$. and finally using a diagonal construction, we can find a sequence of n so that the ratios $\nu(P_n(x) \cap S)/\nu(P_n(x))$ will converge to one for almost every $x \in S$. □

Proof of 13.7: Ergodicity. Suppose that S is a fully invariant set with $\lambda(S) < 1$. Applying 13.10 to the complement $M \smallsetminus S$, we can find a point $x \notin S$ and a sequence of sets $P_{n_j}(x) \in \mathcal{P}_{n_j}$ so that the ratio

$$\frac{\lambda\Big(S \cap P_{n_j}(x)\Big)}{\lambda\Big(P_{n_j}(x)\Big)}$$

tends to zero as $j \to \infty$. Since $f^{\circ n}$ maps $S \cap P_n(x)$ bijectively onto S , it follows from (13 : 12) that

$$\lambda(S) \;\leq\; K\, \frac{\lambda\Big(S \cap P_{n_j}(x)\Big)}{\lambda\Big(P_{n_j}(x)\Big)}\ .$$

Here K is fixed, so the right hand side can be made arbitrarily small, which proves that $\lambda(S) = 0$. Thus the measure class $[\lambda]$ is ergodic. Since there can be at most one invariant probability measure in a given ergodic measure class, this also proves that the measure μ is unique. □

Problem 13-a. If M is compact and $f : M \to M$ is C^1-smooth, show that the following three conditions are equivalent:

- The map f is expanding.
- For any Riemannian metric $\|v\|$, there exist an integer $n > 0$ so that

$$\|Df^{\circ n}(v)\| \;>\; \|v\|$$

for all non-zero tangent vectors v.

- There exists a Riemannian metric, which we can take to be of the form

$$\|v\|_0^2 \;=\; \|v\|^2 + \|Df(v)\|^2 + \ldots + \|Df^{\circ n-1}(v)\|^2 ,$$

so that $\|Df(v)\|_0 > \|v\|_0$ for all non-zero tangent vectors v. (These are called "adapted metrics".)

Problem 13-b. Consider the family of circle maps

$$f_c(x) \;\equiv\; xd + c\,\sin(2\pi x)/2\pi \quad (\mathrm{mod}\ \mathbb{Z}) ,$$

with derivative $f_c'(x) = d + c\,\cos(2\pi x)$. Show that f_c is an $|d|$-fold covering map whenever $|c| \le |d|$, and is expanding whenever $|c| < |d| - 1$. Note that $0 = f_c(0)$ is a fixed point with multiplier $f'(0) = d + c$. If $c \ne 0$, show that f_c cannot be differentiably conjugate to a linear map. If $|d + c| < 1$, show that f_c has an attracting fixed point. Thus, if $|c| \le |d|$ so that the monotone semconjugacy m exists, show that it is is no longer a homeomorphism but rather has an interval of constancy around $x = 0$.

Problem 13-c. Show that a diffeomorphism $F(x) = y$ from $\mathbb{R}$ to $\mathbb{R}$ is the lift of a degree two expanding map having Lebesgue measure as its absolutely continuous invariant measure if and only if the inverse map can be written as

$$2\,F^{-1}(y) \;=\; y + \varphi(y) + c ,$$

where $\varphi(y+1) = -\varphi(y)$, and where $|\varphi'(y)| < 1$. (The case $\varphi(y) = \cos(3\pi y)/10$ is illustrated in Figure 44.)

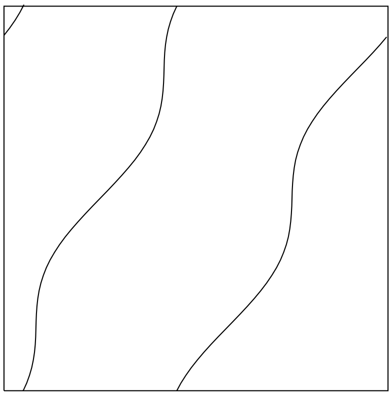

Figure 44. Graph of a degree two expanding map having Lebesgue measure as its absolutely continuous invariant measure. Here the opposite sides of the unit square are to be identified with each other. The reciprocals of the slopes, at two points on the same horizontal line $y = \text{constant}$, *always add up to* $+1$.

Some References.

I. P. Cornfeld, S. V. Fomin, and Ya. G. Sinai, "Ergodic Theory", Springer 1982.

A. Katok and B. Hasselblatt, "Introduction to the Modern Theory of Dynamical Systems", Cambridge U. Press 1995.

K. Krzyzewski and W. Szlenk, "On invariant measures of expanding differentiable mappings, Studia Math. **33** (1969) 83-92.

R. Mañé, "Ergodic Theory and Differentiable Dynamics", Springer 1987.

M. Shub, *Expanding maps*, pp. 273-276 of "Global Analysis", Proc. Symp. Pure Math. **14**, AMS 1970.

Ya. Sinai, "Topics in Ergodic Theory", Princeton U. Press 1994.