IN Variant MEASURES FOR
SET-VALUED DYNAMICAL SYSTEMS

WALTER MILLER AND ETHAN AKIN

Abstract. A continuous map on a compact metric space, regarded as a dynamical system by iteration, admits invariant measures. For a closed relation on such a space, or, equivalently, an upper semicontinuous set-valued map, there are several concepts which extend this idea of invariance for a measure. We prove that four such are equivalent. In particular, such relation invariant measures arise as projections from shift invariant measures on the space of sample paths. There is a similarly close relationship between the ideas of chain recurrence for the set-valued system and for the shift on the sample path space.

1. Introduction

All our spaces will be nonempty, compact, metric spaces. For such a space $X$ let $P(X)$ denote the space of Borel probability measures on $X$ with $\delta : X \to P(X)$ the embedding associating to $x \in X$ the point measure $\delta_x$. The support $|\mu|$ of a measure $\mu$ in $P(X)$ is the smallest closed subset of measure 1. If $f : X_1 \to X_2$ is Borel measurable then the induced map $f_* : P(X_1) \to P(X_2)$ associates to $\mu$ the measure $f_*(\mu)$ defined by

\[
  f_*(\mu)(B) = \mu(f^{-1}(B))
\]

for all $B$ Borel in $X_2$.

We regard a continuous map $f$ on $X$ as a dynamical system by iterating. A measure $\mu \in P(X)$ is called an invariant measure when it is a fixed point for the map $f_* : P(X) \to P(X)$, i.e. $f_*(\mu) = \mu$.

A relation $F : X_1 \to X_2$ is a subset of $X_1 \times X_2$. We associate to $x \in X_1$ the subset of $X_2$ given by $F(x) = \{y : (x, y) \in F\}$. The relation is a function when each $F(x)$ is a singleton. For $A \subset X_1$, the image $F(A) = \bigcup \{F(x) : x \in A\} = \{y : y \in F(x) \text{ for some } x \in A\}$. The inverse relation $F^{-1} : X_2 \to X_1$ is $\{(y, x) : (x, y) \in F\}$. So for $B \subset X_2$, $F^{-1}(B) = \{x : F(x) \cap B \neq \emptyset\}$. In particular, the domain of $F$, denoted $\text{Dom}(F)$, is given by $F^{-1}(X_2) = \{x \in X_1 : F(x) \neq \emptyset\}$. $F$ is a closed (or Borel) relation when it is a closed (resp. Borel) subset of $X_1 \times X_2$.

A function is a closed relation if it is a continuous function. If $F : X_1 \to X_2$ and $G : X_2 \to X_3$ are relations then the composed relation $G \circ F : X_1 \to X_3$ is $\{(x, z) : z \in G(y) \text{ for some } y \in F(x)\}$, i.e. the projection to $X_1 \times X_3$ of the subset $(F \times X_3) \cap (X_1 \times G) \subset X_1 \times X_2 \times X_3$. The composition of closed relations is closed.

Received by the editors June 14, 1996.

1991 Mathematics Subject Classification. Primary 54H20, 58F10, 34C35.

Key words and phrases. Set-valued dynamical system, dynamics of a relation, sample path spaces, invariant measure, basic set, chain recurrence.

©1999 American Mathematical Society

1203
For the elementary facts about relations, and the notation we will use, see Akin (1993), Chapter 1.

Both for its own sake and for application to the usual function case, the dynamics of a relation \( F \) on \( X \), i.e. \( F : X \to X \), have recently received some attention, e.g. Aubin and Frankowska (1990), McGehee (1992) and Akin (1993). As in the function case, the dynamics are obtained by iterating. By convention, \( F^0 \) is the identity map \( 1_X \), \( F^1 = F \), \( F^2 = F \circ F \) and \( F^n \) is the \( n \)-fold composition for any positive integer \( n \); and we write \( F^{-n} \) for \((F^{-1})^n\).

For a relation there are two concepts of invariant measure in the literature which extend the continuous function notion. We introduce two more and will prove in Section 3 that for a closed relation \( F \) on \( X \) all four definitions are equivalent.

(1) Aubin, Frankowska and Lasota (1991) call a measure \( \mu \in P(X) \) invariant for \( F \) if

\[
\mu(B) \leq \mu(F^{-1}(B))
\]

for all Borel subsets \( B \) of \( X \). We will then call \( \mu \) an invariant measure. If \( F = f \) is a function then \( f^{-1}(A) \) and \( f^{-1}(X \setminus A) \) are disjoint, and so, applying (1.2) to \( B = A \) and \( B = X \setminus A \), we obtain equality. By (1.1) this condition says \( \mu = f_* \mu \), and so \( \mu \) is an invariant measure for the map.

(2) A Markov kernel \( \kappa \) from \( X_1 \) to \( X_2 \) is a function from \( X_1 \) to \( P(X_2) \), associating to each \( x \in X_1 \) a measure \( \kappa_x \) on \( X_2 \) such that for each Borel set \( A \), \( \kappa_x(A) \) is a Borel measurable real valued function on \( X_1 \). For example, if \( f : X_1 \to X_2 \) is a Borel measurable function, then \( \delta_f = \delta \circ f : X_1 \to P(X_2) \) is the Markov kernel associated with \( f \), associating to \( x \) the point measure \( \delta_{f(x)} \). For a Markov kernel \( \kappa : X_1 \to P(X_2) \) there is an induced mapping on measures \( \kappa_* : P(X_1) \to P(X_2) \) defined by

\[
\kappa_*(\mu)(B) = \int_{X_1} \kappa_x(B) \mu(dx)
\]

for \( B \) a Borel subset of \( X_2 \). Notice that \( \kappa_*(\mu) \) depends only on the measures \( \kappa_x \) for \( x \in \mu \). Clearly, for \( \kappa = \delta_f, \kappa_* = f_* \). If \( F : X_1 \to X_2 \) we say that the Markov kernel \( \kappa : X_1 \to P(X_2) \) is supported by \( F \) when

\[
|\kappa_x| \subset F(x) \text{ for all } x \in X_1.
\]

For example, let \( f : X_1 \to X_2 \) be a Borel measurable selection function for \( F \), a function satisfying \( f(x) \in F(x) \) for all \( x \in X_1 \). Then \( \delta_f : X_1 \to P(X_2) \) is a Markov kernel supported by \( F \).

1.1. Lemma. If \( F : X_1 \to X_2 \) is a closed relation such that \( \text{Dom}(F) = X_1 \), i.e. \( F^{-1}(X_2) = X_1 \), then there exists a Borel measurable selection function for \( F \).

Proof. Let \( X_3 \subset [0,1] \) denote the Cantor set. By Hocking and Young (1961), Theorem 3.28, there exists a continuous map \( h \) from \( X_3 \) onto the compact metric space \( X_2 \). \( h^{-1} \circ F \) is a closed relation from \( X_1 \) to \( X_3 \). Let \( \hat{f}(x) = \sup h^{-1}(F(x)) \). Since each \( h^{-1}(F(x)) \) is nonempty and closed, \( \hat{f} \) is a selection function for \( h^{-1} \circ F \).

\[
\hat{f} : X_1 \to \mathbb{R} \text{ is upper semicontinuous and so is Borel measurable. } f = h \circ \hat{f} : X_1 \to X_2 \text{ is therefore a Borel measurable selection map for } F. \]
For a closed relation $F$ on $X$, Miller (1995) calls $\mu$ invariant for $F$ if there exists a Markov kernel $\kappa : X \to P(X)$ such that
\begin{equation}
|\kappa_x| \subset F(x) \text{ for all } x \in |\mu| \quad \text{and} \quad \mu = \kappa_\ast(\mu).
\end{equation}
We will then call $\mu$ an invariant\textsubscript{2} measure.

If $F = f$ is a continuous function on $X$, then $\delta_f$ is the unique Markov kernel supported by $f$ and $(\delta_f)_\ast(\mu) = f_\ast(\mu)$. So condition (1.5) specializes to the usual notion of invariant measure when $f$ is a map.

(3) If $\mu_{12}$ is a measure on the product space $X_1 \times X_2$, then the coordinate projections $\pi_i : X_1 \times X_2 \to X_i$ ($i = 1, 2$) induce the marginal measures $\mu_i = \pi_i\ast(\mu_{12})$ ($i = 1, 2$).

For a closed relation $F$ on $X \times X$ we will call $\mu$ an invariant\textsubscript{3} measure if there exists $\mu_{12} \in P(X \times X)$ such that
\begin{equation}
|\mu_{12}| \subset F, \quad \pi_1\ast(\mu_{12}) = \mu = \pi_2\ast(\mu_{12}).
\end{equation}
If $F = f$ is a continuous function then the restriction of $\pi_1$ to the subset $f$ is a homeomorphism with inverse given by $\text{grph}_f : X \to X \times X : \text{grph}_f(x) = (x, f(x))$. The unique measure $\mu_{12}$ on $X \times X$ with support in $f$ and satisfying $\pi_1\ast(\mu_{12}) = \mu$ is $(\text{grph}_f)_\ast(\mu)$. Since $\pi_2 \circ \text{grph}_f = f$, $\pi_2\ast(\text{grph}_f)_\ast(\mu) = f_\ast(\mu)$. So again (1.6) specializes to the proper notion.

(4) Let $X^Z$ denote the set of bi-infinite sequences in $X$ regarded as functions of $Z$ to $X$. With the product topology it is compact. Let $\pi_0 : X^Z \to X$ be the projection $\pi_0(\xi) = \xi_0$. The shift homeomorphism $s : X^Z \to X^Z$ is defined by
\begin{equation}
s(\xi)_i = \xi_{i+1}, \quad i \in Z.
\end{equation}

For a closed relation $F$ on $X$ we denote by $X_F$ the sample path space for $F$, \{ $\xi : \xi_{i+1} \in F(\xi_i)$ for all $i \in Z$ \}, a closed, $s$ invariant subset of $X^Z$. So $s$ restricts to a homeomorphism on $X_F$, which we denote $s_F$. The measures on $X_F$ are called sample path measures. So $\nu \in P(X^Z)$ is a sample path measure when $|\nu| \subset X_F$.

Such a measure is called an invariant sample path measure when it is $s$ invariant or, equivalently, when it is $s_F$ invariant regarded as a measure on $X_F$.

We will call $\mu \in P(X)$ an invariant\textsubscript{4} measure for $F$ if there exists an invariant sample path measure which projects to $\mu$ via $\pi_0 : X^Z \to X$, $\pi_0(\xi) = \xi_0$. That is, there exists $\nu \in P(X^Z)$ such that
\begin{equation}
|\nu| \subset X_F, \quad s_\ast(\nu) = \nu
\end{equation}
and
\begin{equation}
\pi_0\ast(\nu) = \mu.
\end{equation}
If $F = f$ is a homeomorphism on $X$ then $\pi_0$ restricts to a homeomorphism of $X_f$ to $X$ inducing a conjugacy of $s_f$ with $f$. So $\mu$ is $f$ invariant on $X$ exactly when it is the image of an $s_f$ invariant measure under the conjugacy $\pi_0$.

The image of the map $\pi_0 : X_F \to X$ need not be all of $X$. We pause to describe this subset.

Call a relation $F$ on $X$ surjective if $F(X) = X = F^{-1}(X)$ or, equivalently, if $\text{Dom}(F^{-1}) = X = \text{Dom}(F)$. For any subset $B$ of $X$ the restriction of $F$ to $B$, denoted $F_B$, is defined by $F_B = F \cap (B \times B)$. A subset $B$ will be called a surjective subset when $F_B$ is a surjective relation on $B$. Thus, $B$ is a surjective subset exactly when
\begin{equation}
B \subset F(B) \cap F^{-1}(B).
\end{equation}
1.2. Lemma. For a closed relation $F$ on $X$ let $\pi_0 : X_F \to X$ be the projection from the sample path space,

\begin{equation}
\pi_0(X_F) = \bigcap_{n \in \mathbb{Z}} F^n(X).
\end{equation}

We call this subset the dynamic domain of $F$. $\pi_0(X_F)$ is a surjective subset of $X$, and if $B$ is any surjective subset of $X$ then $B \subset \pi_0(X_F)$. In particular, $F$ is a surjective relation iff $X = \pi_0(X_F)$ and so iff the projection $\pi_0 : X_F \to X$ is surjective.

Proof. Since $s_F$ is bijective on $X_F$ and the coordinate $n$ projection $\pi_n$ is $\pi_0(s_F)^n$ for every integer $n$, we clearly have

\begin{equation}
\pi_n(X_F) = \pi_0(X_F) \quad \text{for all } n \in \mathbb{Z}.
\end{equation}

This implies $\pi_0(X_F)$ satisfies (1.9) and so is a surjective subset. By induction on $n = 1, 2, \ldots$ we see that $B \subset F^n(X) \cap F^{-n}(X)$ for any surjective subset $B$, and so any surjective subset is contained in $\bigcap_{n \in \mathbb{Z}} F^n(X)$. Finally, if $x$ is in this intersection, then for each positive integer $m$, \{ $\xi : \xi_0 = x$ and $\xi_{i+1} \in F(\xi_i)$ for $|i| \leq m$ \} is a nonempty compact subset of $X^Z$. The intersection of this decreasing sequence is $\pi_0^{-1}(x) \cap X_F$, which is therefore nonempty. Thus, equation (1.10) holds and $\pi_0(X_F)$ is the maximum surjective subset. In particular, the entire space $X$ is a surjective subset, i.e. $F$ is surjective, iff $X = \pi_0(X_F)$. \hfill $\square$

In Section 2 we will consider the case where $X$ is finite. The general equivalence uses the finite space results, which also have a nice geometrical representation using directed graphs. Furthermore, the finite “pixel” approximations to the space leads us to consider relations even when the original $F = f$ was a homeomorphism. In Section 3 we complete the general proof of the equivalence.

Since map results project from $s_F$ on $X_F$ to $F$ on $X$, it is useful to compare these as dynamical systems. In Section 4 we consider the chain relations. The set of maximal chain transitive subsets of $X_F$, the basic sets for $s_F$, is mapped bijectively by $\pi_0$ onto the set of basic sets for $F$. That is, each $s_F$ basic set maps onto an $F$ basic set via $\pi_0$, and each $F$ basic set is the image of a unique $s_F$ basic set.

Finally, in Section 5 we apply these results to show that every invariant measure supported on a basic set can be $\epsilon$ approximated by the finite measures associated with periodic $\epsilon$ chains, for arbitrarily small $\epsilon$, and can also be expressed as the limit measures of asymptotic chains. The proof uses the embedding of $X_F$ in $X^Z$ even in the case when $F = f$ is a homeomorphism.

We conclude by introducing the particular metrics on $P(X)$ and $X^Z$ which we will use. On $P(X)$ the Hutchinson metric is given by

\begin{equation}
d(\mu_1, \mu_2) = \sup \{ \int u(x) \mu_1(dx) - \int u(x) \mu_2(dx) \},
\end{equation}

where the sup is taken over all real-valued functions $u$ with Lipschitz constant $\leq 1$. Then $\delta$ is an isometric embedding of $X$ into $P(X)$, and if $f : X_1 \to X_2$ is Lipschitz then $f_* : P(X_1) \to P(X_2)$ is Lipschitz with the same Lipschitz constant (see Hutchinson (1981) and Akin (1993), Exercise 8.16).

On the product $X^Z$ we define

\begin{equation}
d(\xi, \eta) = \sup \{ \min(\hat{d}(\xi, \eta_i), 1/|i|) : i \in \mathbb{Z} \},
\end{equation}
where \( \min(d, 1/0) = d \) by convention. It is not hard to show that \( d \) is a metric on \( X^Z \) yielding the product topology, and, for \( \epsilon > 0 \),

\[
(1.14) \quad d(\xi, \eta) \leq \epsilon \Leftrightarrow d(\xi, \eta_i) \leq \epsilon \text{ for all } i \text{ s.t. } |i| < 1/\epsilon.
\]

In particular, the projection map \( \pi_0 : X^Z \to X \) is distance nonincreasing, i.e. it is Lipschitz with Lipschitz constant = 1. Thus, the same is true for \( \pi_0 : P(X^Z) \to P(X) \).

2. Finite Spaces

When \( X \) is a finite set we can represent a relation \( F \subset X \times X \) as a directed graph with vertices the points of \( X \) and an edge from \( x \) to \( y \) exactly when \( (x, y) \in F \). Then the characteristic function \( \chi_F \) is a 0 - 1 square matrix.

A measure \( \mu_{12} \) on \( F \) can be regarded as a real valued function on \( X \times X \) satisfying

\[
(2.1) \quad 0 \leq \mu_{12} \leq \chi_F, \quad \sum_{(x,y) \in X \times X} \mu_{12}(x, y) = 1,
\]

i.e. a probability distribution on the edges of the graph. The marginal measures \( \mu_1 \) and \( \mu_2 \) induced by \( \mu_{12} \) are the distributions on the set of vertices \( X \) given by

\[
(2.2) \quad \mu_1(x) = \sum_{y \in X} \mu_{12}(x, y) = \sum_{y \in F(x)} \mu_{12}(x, y), \quad \mu_2(y) = \sum_{x \in X} \mu_{12}(x, y) = \sum_{x \in F^{-1}(y)} \mu_{12}(x, y),
\]

where the empty sum is 0 by convention. Thus, \( \mu_1(x) \) is the probability that an edge begins at \( x \) and \( \mu_2(y) \) is the probability that an edge ends at \( y \). The associated Markov kernel is the conditional distribution on the terminal vertices assuming the initial one:

\[
(2.3) \quad \kappa_x(y) = \frac{\mu_{12}(x, y)}{\mu_1(x)},
\]

a distribution well defined when \( \mu_1(x) > 0 \), which is then supported on \( F(x) \). For \( x \not\in \mu_1 \), i.e. \( \mu_1(x) = 0 \), we choose \( \kappa_x \) to be an arbitrary distribution on \( X \). Notice that if, for some \( x \), \( F(x) = \emptyset \), i.e. no edge begins at \( x \), then \( \kappa_x \) cannot be chosen supported by \( F \). But we do have

\[
(2.4) \quad |\kappa_x| \subset F(x) \quad \text{for} \quad x \in |\mu_1|, \quad \mu_2(y) = \sum_x \kappa_x(y) \mu_1(x).
\]

Conversely, if \( \mu_1 \) and \( \mu_2 \) are distributions on \( X \) and \( \kappa \) is a Markov kernel, then we obtain the associated distribution on \( X \times X \):

\[
(2.5) \quad \mu_{12}(x, y) = \kappa_x(y) \mu_1(x).
\]

If \( \kappa \) satisfies (2.4), then the marginals of \( \mu_{12} \) are \( \mu_1 \) and \( \mu_2 \) and the support of \( \mu_{12} \) is contained in \( F \).

In particular, a distribution \( \mu \) on \( X \) is an invariant measure for \( F \), i.e. there exists \( \mu_{12} \) satisfying (2.1) with \( \mu_1 = \mu = \mu_2 \), iff there is a Markov kernel \( \kappa \) satisfying (2.4) with \( \mu_1 = \mu = \mu_2 \) and so iff \( \mu \) is invariant for \( F \).

From \( \mu_{12} \) on \( X \times X \), supported by \( F \) and with common marginals \( \mu \), we can define a measure \( \nu \) on \( X^Z \) with support \( X_F \) and such that \( s_\nu(\nu) = \nu \). For integers
\(a, \ b\) with \(a < b\) and a sequence \(\{x_i : a \leq i \leq b\}\) in \(X\) we define \(\nu\) on the associated cylinder set by

\[
(2.6) \quad \nu(\{\xi \in X^\mathbb{Z} : \xi_i = x_i \text{ for } a \leq i \leq b\}) = \frac{\prod_{i=a}^{b-1} \mu_{12}(x_i, x_{i+1})}{\prod_{i=a+1}^{b-1} \mu(x_i)},
\]

where the empty product is 1. Clearly, \((\pi_{n,n+1})_\ast(\nu) = \mu_{12}\) for any integer \(n\), where \(\pi_{n,n+1}\) is the projection from \(X^\mathbb{Z}\) to \(X \times X\) via the \(n\) and \(n + 1\) coordinates. Hence \(\pi_n(\nu) = \mu\), and because \(\mu_{12}(x, y) = 0\) unless \((x, y) \in F\) it follows that \(\nu\) is supported by \(X_F\). Obviously \(s_\ast(\nu) = \nu\), and so \(\nu\) is an invariant sample path measure.

Conversely, if \(\nu\) is an invariant sample path measure such that \(\pi_0(\nu) = \mu\) and so \(\pi_n(\nu) = \mu\), then \((\pi_{n,n+1})_\ast(\nu) = \mu_{12}\) satisfies (2.1) and (2.2) with \(\mu_1 = \mu = \mu_2\). Thus, \(\mu\) is invariant for \(F\) iff it is invariant for \(F\).

Finally, the equivalence of the invariant\(_1\) and invariant\(_3\) conditions is a special case of the following theorem of Brualdi (1968).

**2.1. Theorem.** Let \(X_1\) and \(X_2\) be nonempty finite sets and let \(F\) be a subset of the product \(X_1 \times X_2\). Let \(\mu_i\) be a distribution on \(X_i\) for \(i = 1, 2\). There exists a distribution \(\mu_{12}\) on \(X_1 \times X_2\) supported by \(F\) and with marginals \(\mu_1\) and \(\mu_2\) iff for all subsets \(B\) of \(X_2\), \(\mu_2(B) \leq \mu_1(F^{-1}(B))\), that is, if \(A_1 \supset F^{-1}(A_2)\) for subsets \(A_i\) of \(X_i\) \((i = 1, 2)\) then

\[
(2.7) \quad \sum_{y \in A_2} \mu_2(y) \leq \sum_{x \in A_1} \mu_1(x).
\]

**Proof.** Necessity of (2.7) is easy (see the general proof of (3) \(\Rightarrow\) (1) in Theorem 3.1 below). We prove sufficiency using Theorem 2.6 of Gale (1960), a separation theorem which says that for an \(m \times n\) matrix \(A\) and \(1 \times n\) row vector \(b\) either the equation \(cA = b\) has a nonnegative \(1 \times m\) row vector solution \(c\), or for some \(n \times 1\) column vector \(d\), \(Ad\) is nonnegative while \(bd < 0\).

Using \(F\) to index the rows and disjoint copies of \(X_1\) and \(X_2\) to index the columns we define

\[
(2.8) \quad A_{(x,y)z} = \begin{cases} 1 & \text{if } z = x \ or \ z = y, \\ 0 & \text{otherwise}, \end{cases}
\]

\[
b_z = \begin{cases} \mu_1(z) & \text{for } z \in X_1, \\ \mu_2(z) & \text{for } z \in X_2, \end{cases}
\]

The required distribution \(\mu_{12}\) is exactly a nonnegative solution \(c\) of the equation \(cA = b\). We prove it exists by showing that \(Ad\) nonnegative implies \(bd \geq 0\).

A vector \(d\) such that \(Ad\) is nonnegative satisfies \(d_x + d_y \geq 0\) whenever \((x, y) \in F\). Choose a positive constant \(k\) large enough that for all \(x \in X_1, y \in X_2\):

\[
(2.9) \quad u_1(x) = d_x + k > 0, \quad u_2(y) = -d_y + k > 0.
\]

Thus, \(u_i\) is a positive real valued function on \(X_i\) for \(i = 1, 2\), and \(Ad\) nonnegative implies

\[
(2.10) \quad u_1(x) \geq u_2(y) \text{ whenever } (x, y) \in F.
\]

For any \(\epsilon \geq 0\), (2.10) implies

\[
(2.11) \quad \{u_1 \geq \epsilon\} \supset F^{-1}\{u_2 \geq \epsilon\}
\]
where \( \{u_i \geq \epsilon\} \equiv \{z \in X_i : u_i(z) \geq \epsilon\} \) \( (i = 1,2) \). Thus, condition (2.7) implies

\[
\mu_1(\{u_1 \geq \epsilon\}) \geq \mu_2(\{u_2 \geq \epsilon\}).
\]

Now let \( 0 \leq \epsilon_0 < \epsilon_1 < \ldots < \epsilon_N \) list all the values of \( u_1 \) and \( u_2 \) in ascending order. Define \( \delta_0 = \epsilon_0 \) and \( \delta_i = \epsilon_i - \epsilon_{i-1} \) for \( i = 1, \ldots, N \). Then

\[
\sum_x u_1(x)\mu_1(x) = \sum_{i=0}^{N} \epsilon_i \mu_1(\{u_1 = \epsilon_i\}) = \sum_{j=0}^{N} \sum_{i=j}^{N} \delta_j \mu_1(\{u_1 = \epsilon_i\})
\]

\[
= \sum_{j=0}^{N} \sum_{i=j}^{N} \delta_j \mu_1(\{u_1 = \epsilon_i\}) \geq \sum_{j=0}^{N} \delta_j \mu_2(\{u_2 \geq \epsilon_j\}) = \sum_{y} u_2(y)\mu_2(y).
\]

Since \( \mu_i \) is a distribution it follows that \( \sum_z k\mu_i(z) = k \) \( (i = 1,2) \), and so (2.13) and (2.9) imply \( bd \geq 0 \). \( \Box \)

The finite case provides a useful tool for studying dynamical systems on arbitrary spaces. Consider for a compact, metric space \( X \) a partition into finitely many sets, or equivalently, let \( E \subset X \times X \) be an equivalence relation with a finite number of distinct equivalence classes. We will call such a relation a \textit{finite equivalence relation}. Let \( X/E = \{E(x) : x \in X\} \) be the finite set of equivalence classes, with \( \pi : X \rightarrow X/E \) the projection map associating to \( x \in X \) its class \( E(x) \).

We will call \( E \) a \textit{Borel equivalence relation} when \( E \) is a Borel subset of \( X \times X \). Letting \( i_x : X \rightarrow X \times X \) be the inclusion given by \( i_x(y) = (x,y) \), we see that \( E(x) = (i_x)^{-1}(E) \) is Borel when \( E \) is Borel. Conversely, a finite equivalence relation \( E = \bigcup\{E(x) \times E(x) : x \in X\} \) is Borel when each equivalence class is Borel. Thus, \( E \) is Borel exactly when \( \pi_E \) is Borel measurable.

For a finite equivalence relation \( E \) the \textit{mesh of \( E \)} is the maximum of the diameters of the equivalence classes of \( E \).

\section*{2.2. Lemma.} \textit{Let \( E \) be a finite, Borel equivalence relation on \( X \) with mesh \( \leq \epsilon \). If \( \mu_1, \mu_2 \in P(X) \) satisfy \( (\pi_E)_*(\mu_1) = (\pi_E)_*(\mu_2) \) in \( P(X/E) \), then the Hutchinson distance \( d(\mu_1, \mu_2) \) is at most \( 2\epsilon \).

Proof.} Let \( \mu_E = (\pi_E)_*(\mu_1) = (\pi_E)_*(\mu_2) \) and let \( u \) be a real valued map on \( X \) with Lipschitz constant at most 1. Since the diameter of \( E(x) \) is at most \( \epsilon \), we have for \( y \in E(x) \)

\[
u(x) - \epsilon \leq u(y) \leq u(x) + \epsilon.
\]

If \( \mu_E(E(x)) = \mu_i(E(x)) \) \( (i = 1,2) \) is positive, then we can define the conditional expectations \( (i = 1,2) \):

\[
E_i(u|E(x)) = \frac{1}{\mu_E(E(x))} \int_{E(x)} u(y)\mu_i(dy)
\]

and get from (2.14)

\[
\mu_i(E(x)) = E_1(u|E(x)) \leq u(x) + \epsilon,
\]

and so for every class \( E(x) \) of positive measure we have

\[
E_1(u|E(x)) - E_2(u|E(x)) \leq 2\epsilon.
\]
Multiply by $\mu_E(E(x))$ and sum over $X/E$ to get
\begin{equation}
\int u(x)\mu_1(dx) - \int u(x)\mu_2(dx) \leq 2\epsilon.
\end{equation}

By the definition (1.12) of the Hutchinson metric we see that $d(\mu_1, \mu_2) \leq 2\epsilon$. □

A section is a choice function for the equivalence classes, i.e. a map $\omega : X/E \to X$ such that $\pi_E \circ \omega = 1_{X/E}$. So if $\mu_E$ is any measure on $X/E$, then
\begin{equation}
(\pi_E)_* \omega_* (\mu_E) = \mu_E,
\end{equation}
where
\begin{equation}
\omega_* (\mu_E) = \sum_{A \in X/E} \mu_E(A) \delta_{\omega(A)}.
\end{equation}
That is, $\omega_* (\mu_E)$ is the weighted average of the point measures $\delta_{\omega(A)}$ with weights $\mu_E(A)$.

Now let $F$ be a relation from $X_1$ to $X_2$ and $E_i$ be a finite equivalence relation on $X_i (i = 1, 2)$. Let $E_1 \times E_2$ denote the equivalence relation on $X_1 \times X_2$ such that $(E_1 \times E_2)(x, y) = E_1(x) \times E_2(y)$ and so that the projection $\pi_{E_1 \times E_2}$ is identified with $\pi_{E_1} \times \pi_{E_2}$. Define $F_E = (\pi_{E_1} \times \pi_{E_2})(F)$, a relation from $X_1/E_1$ to $X_2/E_2$, so that
\begin{equation}
(E_1(x), E_2(y)) \in F_E
\iff (E_1(x) \times E_2(y)) \cap F \neq \emptyset
\iff F(E_1(x)) \cap E_2(y) \neq \emptyset
\iff E_1(x) \cap F^{-1}(E_2(y)) \neq \emptyset.
\end{equation}

If $E$ is a closed relation and $\mu_{12}$ satisfies $|\mu_{12}| \subset F$ then $\mu_{12}E \equiv (\pi_{E_1} \times \pi_{E_2})_* (\mu_{12})$ is a measure on $(X_1/E_1) \times (X_2/E_2)$ with $|\mu_{12}E| \subset F_E$.

We call a section $\omega : X_1/E_1 \times X_2/E_2 \to X_1 \times X_2$ adapted to $F$ if $\omega$ maps $F_E$ to $F$, i.e.
\begin{equation}
(E_1(x), E_2(x)) \in F_E \Rightarrow (\omega(E_1(x)), E_2(x)) \in F.
\end{equation}

By (2.21) such sections always exist. If $\mu_{12}E$ is a measure on $(X_1/E_1) \times (X_2/E_2)$ which is supported by $F_E$ and if $\omega$ is a section adapted to $F$, then $\omega_* (\mu_{12}E)$ is a measure on $X_1 \times X_2$ supported by $F$.

3. Invariant Measures

3.1. Theorem. For compact metric spaces $X_1, X_2$ let $F \subset X_1 \times X_2$ be a closed relation. For a pair of measures $(\mu_1, \mu_2) \in P(X_1) \times P(X_2)$ the following conditions are equivalent,

(1) For every Borel set $A_2 \subset X_2$:
\begin{equation}
\mu_2(A_2) \leq \mu_1(F^{-1}(A_2)).
\end{equation}

(2) There exists a function $\kappa : X_1 \to P(X_2)$ such that for every Borel set $A_2 \subset X_2$ the function $\kappa_x(A_2)$ is a Borel measurable function of $x \in X_1$,
\begin{equation}
\mu_2(A_2) = \int_{X_1} \kappa_x(A_2) \mu_1(dx),
\end{equation}
and for each $x$ in the support of $\mu_1$, the measure $\kappa_x$ is supported by $F(x)$, i.e.
\begin{equation}
|\kappa_x| \subset F(x) \text{ for all } x \in |\mu_1|.
\end{equation}
(3) There exists a measure $\mu_{12} \in P(X_1 \times X_2)$ which projects to $\mu_i$ via $\pi_i : X_1 \times X_2 \to X_i$ ($i = 1, 2$), i.e. for every Borel subset $A_1 \subset X_1$ and $A_2 \subset X_2$,
\[ \mu_1(A_1) = \mu_{12}(A_1 \times X_2), \]
\[ \mu_2(A_2) = \mu_{12}(X_1 \times A_2), \]
and the measure $\mu_{12}$ is supported by $F$, i.e.
\[ |\mu_{12}| \subset F. \]

Let $F_*$ denote the set of pairs $(\mu_1, \mu_2)$ which satisfy these conditions. $F_*$ is a closed relation, i.e. a closed subset of $P(X_1) \times P(X_2)$ whose domain satisfies
\[ \text{Dom}(F_*) = \{ \mu_1 \in P(X_1) : |\mu_1| \subset \text{Dom}(F) \}. \]

Furthermore, for the inverse relation $F^{-1} \subset X_2 \times X_1$,
\[ (F^{-1})_* = (F_*)^{-1}. \]

If $F : X_1 \to X_2$ is a continuous map, then $F_*$ is the induced continuous map from $P(X_1)$ to $P(X_2)$.

Proof. (2) $\Rightarrow$ (3): Given $\mu_1 \in P(X_1)$ and a Markov kernel $\kappa : X_1 \to P(X_2)$, define $\mu_{12} \in P(X_1 \times X_2)$ on product sets by
\[ \mu_{12}(A_1 \times A_2) = \int_{A_1} \kappa_x(A_2) \mu_1(dx). \]

For $u$ a real, bounded, measurable function on $X_1 \times X_2$
\[ \int_{X_1 \times X_2} u(x, y) \mu_{12}(d(x, y)) = \int_{X_1} (\int_{X_2} u(x, y) \kappa_x(dy)) \mu_1(dx). \]

Clearly, $\mu_{12}(A_1 \times X_2) = \mu_1(A_1)$ in any case, and (3.2) implies $\mu_{12}(X_1 \times A_2) = \mu_2(A_2)$. To prove $|\mu_{12}| \subset F$ it suffices to show that when a continuous function $u$ vanishes on $F$ its integral, given by (3.9), is zero. For $x$ fixed in $|\mu_1|$, $u(x, y) = 0$ for $y \in F(x)$. Because $|\kappa_x| \subset F(x)$ we have, for $x \in |\mu_1|$,\[ \int_{X_2} u(x, y) \kappa_x(dy) = \int_{F(x)} u(x, y) \kappa_x(dy) = 0. \]

But the outer integral in (3.9) is unchanged if we restrict the range of integration from $X_1$ to $|\mu_1|$. So by (3.10) the integral in (3.9) is zero.

(3) $\Rightarrow$ (2): Because $\mu_{12}$ is a measure on a separable metric space there exists a regular conditional distribution for the $\sigma$ algebra of subsets $B_1 = \{ \pi^{-1}_1(A_1) = A_1 \times X_2 : A_1 \subset X_1 \text{ Borel} \}$ (see, e.g., Gihman and Skorohod (1974), Theorem I.3.3). This is a $B_1$ measurable Markov kernel $\bar{\kappa} : X_1 \times X_2 \to P(X_1 \times X_2)$ such that for $A_1 \subset X_1$ Borel and $u : X_1 \times X_2 \to \mathbb{R}$ bounded and Borel measurable we have
\[ \int_{A_1 \times X_2} u(x, y) \mu_{12}(d(x, y)) = \int_{A_1 \times X_2} \int_{X_1 \times X_2} u(\bar{x}, \bar{y}) \bar{\kappa}_{(x, y)}(d(\bar{x}, \bar{y})) |\mu_{12}(d(x, y)). \]

We can fix $y = y_0 \in X_2$ to define $\kappa : X_1 \to P(X_2)$ by $\kappa_x = (\pi_2)_* \bar{\kappa}_{(x, y_0)}$. Because $\bar{\kappa}$ is $B_1$ measurable and $\pi_1 \cdot (\mu_{12}) = \mu_1$, (3.8) follows from (3.11) with $u$ the characteristic function of $X_1 \times A_2$. With $A_1 = X_1$ (3.2) follows because $\pi_2 \cdot (\mu_{12}) = \mu_2$. If $U_1 \times U_2$ is open in $X_1 \times X_2$ and disjoint from $F$, then $\mu_{12}(U_1 \times U_2) = 0$, and so from (3.8) $\kappa_x(U_2) = 0$ for $\mu_1$ almost every $x \in U_1$. By letting $U_1$ and $U_2$ vary over countable...
bases for the topologies we can obtain a Borel set $N$ of $\mu_1$ measure zero such that for $x \notin N$, $|\kappa_x| \subset F(x)$. Because $|\mu_1| \subset F$ it follows that $|\mu_1| = |\pi_1 \cdot (\mu_1)| \subset \pi_1(F) = \text{Dom}(F)$. Use Lemma 1.1 to construct $f : \text{Dom}(F) \to X_2$, a Borel selection function for $F$, and replace $\kappa_x$ by $\delta_{f(x)}$ for all $x \in N \cap |\mu_1|$. Condition (3.3) then holds for the adjusted function $\kappa$.

(3) $\Rightarrow$ (1): If $A_2 \subset X_2$ is Borel, then $F \cap (X_1 \times A_2)$ is a Borel subset of $X_1 \times X_2$ and its image under $\pi_1$ is $F^{-1}(A_2)$. Thus, the latter set, while not necessarily Borel, is analytic and so is measurable with respect to any Borel measure on $X_1$. Since $\mu_1(F^{-1}(A_2))$ is always defined for $A_2 \subset X_2$ Borel.

If $\mu_{12} \in P(X_1 \times X_2)$ satisfies the conditions of (3), then, because $|\mu_{12}| \subset F$,

$$\mu_2(A_2) = \mu_{12}(X_1 \times A_2) = \mu_{12}((X_1 \times A_2) \cap F) \leq \mu_{12}(F^{-1}(A_2) \times X_2) = \mu_1(F^{-1}(A_2)),$$

proving (3.1).

(1) $\Rightarrow$ (3): Fix $\epsilon > 0$ and for $i = 1, 2$ choose $E_i$ a finite equivalence relation on $X_i$ with mesh at most $\epsilon$. As in Section 2 we let $F_E$ denote $(\pi_{E_1} \times \pi_{E_2})(F) \subset (X_1/E_1) \times (X_2/E_2)$. Let $\nu_i = (\pi_{E_i})_*(\mu_i)$ ($i = 1, 2$). Choose a section $\omega : (X_1/E_1) \times (X_2/E_2) \to X_1 \times X_2$ which is adapted to $F$.

For $B_2 \subset X_2/E_2$,

$$\mu_1(F^{-1}(A_2) \times X_2) = \mu_1(F^{-1}(B_2)).$$

Consequently, (3.1) implies

$$\nu_2(B_2) = \mu_2((\pi_{E_2})^{-1}(B_2)) \leq \mu_1(F^{-1}((\pi_{E_2})^{-1}(B_2)))$$

$$\leq \mu_1((\pi_{E_1})^{-1}(F^{-1}(B_2))) = \nu_1(F^{-1}(B_2)).$$

Thus, condition (2.7) holds for the pair $(\nu_1, \nu_2) \in P(X_1/E_1) \times P(X_2/E_2)$.

Now we apply Brualdi’s Theorem, Theorem 2.1, to $F_E$ and obtain a measure $\mu_{12}$ on $(X_1/E_1) \times (X_2/E_2)$ which satisfies

$$|\nu_{12}| \subset F_E,$$

$$\pi_{i*}(\nu_{12}) = \nu_i \quad (i = 1, 2).$$

Use the section $\omega$ to define the finite measure $\mu_E = (\omega)_*(\nu_{12})$ on $X_1 \times X_2$. Because $\omega$ is adapted to $F$ and $\mu_{12}$ is supported by $F_E$, it follows that $\mu_E$ is supported by $F$. Also, $\pi_i \circ (\pi_{E_1} \times \pi_{E_2}) = \pi_{E_i} \circ \pi_i : X_1 \times X_2 \to X_i/E_i$ ($i = 1, 2$), and so we have

$$|\mu_E| \subset F,$$

$$\mu_i(\pi_{E_i})*(\mu_E) = \nu_i = (\pi_{E_i})_*(\mu_i) \quad (i = 1, 2).$$

Apply Lemma 2.2. Because the mesh of $E_i$ is at most $\epsilon$,

$$d(\mu_i, (\pi_{E_i})_*(\mu_E)) \leq 2\epsilon \quad (i = 1, 2).$$

Now vary $\epsilon$ and let $\mu_{12}$ be a limit point in $P(X_1 \times X_2)$ of the $\mu_E$’s as $\epsilon \to 0$. From (3.16) the limit measure is supported by $F$, and (3.17) implies $(\pi_i)_*(\mu_{12}) = \mu_i$ ($i = 1, 2$). Thus, (3.4) and (3.5) of (3) hold.

This completes the proof of the equivalence.

The set of measures on $X_1 \times X_2$ supported by the closed set $F$ is closed in $P(X_1 \times X_2)$ (e.g. by Akin (1993), Proposition 8.1, the relation $\mu \to |\mu|$ is lower semicontinuous). So by compactness and (3) the image of this set under $\pi_{1*} \times \pi_{2*}$ is
closed in \(P(X_1) \times P(X_2)\). Clearly, \(|\mu_{12}| \subset F \) implies \(|\mu_1| = |\pi_1 \ast (\mu_{12})| \subset \text{Dom}(F) = \pi_1(F)\). Conversely, if \(|\mu_1| \subset \text{Dom}(F)\), let \(f : \text{Dom}(F) \to X_2\) be a measurable selection function for \(F\); \(\kappa_x = \delta_{f(x)}\) satisfies (3.3), and so, with \(\mu_2\) defined by (3.2), \((\mu_1, \mu_2)\) is in \(F_*\) by (2).

Equation (3.7) is clear from (3), and the case when \(F\) is a continuous map was analyzed in Section 1.

**3.2. Theorem.** Let \(F \subset X \times X\) be a closed relation on a compact metric space \(X\). A measure \(\mu \in P(X)\) is called an invariant measure for \(F\) when it satisfies the following equivalent conditions.

1. For every Borel set \(A \subset X\)
   \[\mu(A) \leq \mu(F^{-1}(A)).\]
   (3.18)

2. There exists a Markov kernel \(\kappa : X \to P(X)\) satisfying
   \[x \in |\mu| \Rightarrow |\kappa_x| \subset F(x)\]
   and
   \[\mu = \kappa_\ast(\mu).\]
   (3.19)

3. There exists a measure \(\tilde{\mu} \in P(X \times X)\) satisfying
   \[|\tilde{\mu}| \subset F\]
   and
   \[\mu = (\pi_0)_\ast(\tilde{\mu}) = (\pi_2)_\ast(\tilde{\mu}).\]
   (3.20)

4. There exists a measure \(\nu \in P(X^Z)\) which is invariant with respect to the shift homeomorphism \(s\) on \(X^Z\) and satisfying, with \(X_F\) the sample path space for \(F\),
   \[|\nu| \subset X_F\]
   and
   \[\mu = (\pi_0)_\ast(\nu).\]
   (3.21)

The set \(P_F(X)\) of \(F\) invariant measures is compact and convex in \(P(X)\). It is nonempty if and only if \(X_F\) is nonempty and so if and only if the dynamic domain \(\pi_0(X_F)\) is nonempty. In general, if \(\mu \in P_F(X)\) then

\[|\mu| \subset \pi_0(X_F).\]

(3.25)

**Proof.** By Theorem 3.1 each of the first three conditions is equivalent to \((\mu, \mu) \in F_*\).

(4) \(\Rightarrow\) (3): Define \(\tilde{\mu} \in P(X \times X)\) by

\[\tilde{\mu} = (\pi_{0,1})_\ast(\nu).\]

(3.26)

Condition (3.22) then follows from (3.24) and shift invariance. Since the set \((\pi_{0,1})^{-1}(X \times X \setminus F)\) is open and disjoint from \(X_F\), (3.23) implies this set has \(\nu\) measure 0, and so \(\tilde{\mu}\) is supported by \(F\).

(2) \(\Rightarrow\) (4): Let \(\mu^1 = \mu \in P(X^1)\) and inductively use \(\mu^n \in P(X^n)\) and \(X^n \xrightarrow{\pi_n} X \xrightarrow{\pi} P(X)\) to define \(\mu^{n+1} \in P(X^{n+1})\) via (3.8). Applying the Kolmogorov Product Theorem to this sequence of measures, we obtain from the Markov kernel \(\kappa\) and the stationary measure \(\mu\) the associated sample path measure \(\nu\). The resulting measure is clearly shift invariant and satisfies (3.24). By the proof of (2) \(\Rightarrow\) (3) in
Theorem 3.1, \( \mu^2(X \times X \setminus F) = 0 \). Since \( X^Z \setminus X_F = \bigcup_{i \in Z} \{(\pi_{i,i+1})^{-1}(X \times X \setminus F)\} \) we see that the complement of \( X_F \) has \( \nu \) measure 0, i.e. (3.23) holds.

The set of \( s_F \) invariant measures in \( P(X_F) \) is compact and convex and nonempty—by the Krylov-Bogolyubov Theorem—when \( X_F \) is nonempty. So its image under \( \pi_\omega \), satisfies the same properties. The inclusion (3.25) follows from (3.23) and (3.24).

For \( \epsilon \geq 0 \) call \( \mu \in P(X) \) an \( \epsilon \) invariant measure for the closed relation \( F \) on \( X \) if there exists \( (\mu_1, \mu_2) \in F_\epsilon \) such that \( d(\mu, \mu) \leq \epsilon \) for \( i = 1, 2 \). So an invariant measure is exactly a 0 invariant measure. If \( (\mu_1, \mu_2) \in F_\epsilon \) with \( d(\mu_1, \mu_2) \leq \epsilon \), then \( \mu_1 \) and \( \mu_2 \) are both \( \epsilon \) invariant measures. If \( \mu' \) is an invariant measure and \( d(\mu', \mu) \leq \epsilon \), then \( \mu \) is an \( \epsilon \) invariant measure. On the other hand,

### 3.3. Lemma

Let \( F \) be a closed relation on \( X \). For every \( \epsilon > 0 \) there exists \( \epsilon_1 > 0 \) so that if \( \mu \) is an \( \epsilon_1 \) invariant measure then there exists an invariant measure \( \mu' \) such that \( d(\mu', \mu) < \epsilon \).

**Proof.** The set of \( \epsilon_1 \) invariant measures is a closed and hence compact subset of \( P(X) \). As \( \epsilon_1 \) decreases to 0 the sets decrease toward the set of invariant measures, and so are eventually contained in its \( \epsilon \) neighborhood.

From the proof of Theorem 3.1 we extract some useful information about the pixel approximation.

### 3.4. Theorem

Let \( E \) be a finite Borel equivalence relation on \( X \) with mesh \( \leq \epsilon/2 \). Let \( \pi_E : X \rightarrow X/E \) be the projection to the set of equivalence classes. Assume \( \omega : X/E \rightarrow X \) is a section, i.e. \( \pi_E \circ \omega = 1_{X/E} \). For a closed relation \( F \) on \( X \), let \( F_E = (\pi_E \times \pi_E)(F) \) be the induced relation on the finite set \( X/E \).

(a) For \( \mu \) a measure on \( X \), let \( \mu_E = (\pi_E)_*(\mu) \). If \( \mu_\omega = \omega_*(\mu_E) \), then \( (\pi_E)_*(\mu_\omega) = \mu_E \) as well, and

\[
(3.27) \quad d(\mu, \mu_\omega) \leq \epsilon.
\]

If \( \mu \) is an \( F \) invariant measure, then \( \mu_E \) is an \( F_E \) invariant measure.

(b) For \( \mu_E \) a measure on \( X/E \) let \( \mu_\omega = \omega_*(\mu_E) \). If \( \mu_E \) is an \( F_E \) invariant measure, then \( \mu_\omega \) is an \( \epsilon \) invariant measure for \( F \) with \( (\pi_E)_*(\mu_\omega) = \mu_E \).

**Proof.** \( \pi_E \circ \omega = 1_{X/E} \) implies \( \pi_E*(\mu_\omega) = \mu_E \). So Lemma 2.2 implies (3.27).

If \( \tilde{\mu} \in P(X \times X) \) is supported by \( F \) with \( (\pi_i)_*(\tilde{\mu}) = \mu \) for \( i = 1, 2 \), then \( \tilde{\mu}_E = (\pi_E \times \pi_E)_*(\tilde{\mu}) \) is supported by \( F_E \) with \( (\pi_i)_*(\tilde{\mu}_E) = \mu_E \) for \( i = 1, 2 \). So \( \mu \) invariant implies \( \mu_E = (\pi_E)_*(\mu) \) is \( F_E \) invariant, proving (a).

Let \( \omega' : (X/E) \times (X/E) \rightarrow X \times X \) be a section of \( \pi_E \times \pi_E \) which is adapted to \( F \). If \( \tilde{\mu}_E \in P((X/E) \times (X/E)) \) with \( (\pi_i)_*(\tilde{\mu}_E) = \mu_E \) for \( i = 1, 2 \), then \( \tilde{\mu}_\omega = (\omega')_*(\tilde{\mu}_E) \) is supported by \( F \). So if \( \mu_{i,\omega} = (\pi_i)_*(\tilde{\mu}_\omega) \) \( (i = 1, 2) \), then \( (\mu_{1,\omega}, \mu_{2,\omega}) \in F_\epsilon \). No two of \( \{\mu_{i,\omega}, \mu_{2,\omega}, \mu_{3,\omega}\} \) need be equal, but we do have, for \( i = 1, 2 \),

\[
(3.28) \quad (\pi_i)_*(\mu_\omega) = (\pi_i)_*(\tilde{\mu}_E) = \mu_E = (\pi_E)_*(\mu_\omega).
\]

By Lemma 2.2 we have \( d(\mu_{i,\omega}, \mu_{2,\omega}) \leq \epsilon \) \( (i = 1, 2) \), and so \( \mu_\omega \) is \( \epsilon \) invariant. \( \square \)

### 4. Chain Recurrence

Since \( F \) invariant measures on \( X \) are the image under \( \pi_0 \) of the \( s_F \) invariant sample path measures on \( X_F \), it will be useful to relate the dynamics of \( s_F \) on \( X_F \) with \( F \) on \( X \).
For a closed relation $F$ on $X$ and $\epsilon \geq 0$, an $\epsilon$ chain is a sequence $\{x_i\}$ indexed by some subinterval $I$ of $\mathbb{Z}$ such that

$$d(x_{i+1}, F(x_i)) \leq \epsilon \quad \text{for } i, i + 1 \in I. \tag{4.1}$$

Here $I$ can be finite or infinite, but has length at least one, so that at least one pair $i, i + 1$ lies in $I$. By compactness (4.1) says there exists $z_{i+1} \in F(x_i)$ such that $d(x_{i+1}, z_{i+1}) \leq \epsilon$. In particular, for a chain ($= 0$ chain) we have $x_{i+1} \in F(x_i)$ for $i, i + 1 \in I$. A sequence $\{x_i : i \in I\}$ is called an asymptotic chain if $I$ is unbounded on the right and

$$\lim_{i \to \infty} d(x_{i+1}, F(x_i)) = 0. \tag{4.2}$$

An $\epsilon$ asymptotic chain is an asymptotic chain which is also an $\epsilon$ chain. When $I = \mathbb{Z}$ we obtain points $\xi$ of $X^\mathbb{Z}$, and so we speak of the points $\xi$ of $X^\mathbb{Z}$ which are $\epsilon$ chains, asymptotic chains or $\epsilon$ asymptotic chains. In particular, $X_F$ is the set of 0 chains in $X^\mathbb{Z}$.

If $F$ is a surjective relation on $X$, i.e.

$$F(X) = X = F^{-1}(X), \tag{4.3}$$

then for any $\epsilon$ chain $\{x_i : i \in I\}$ there exists points $\xi \in X^\mathbb{Z}$ such that

$$\xi_i = x_i, \quad i \in I, \tag{4.4}$$

$$\xi_{i+1} \in F(\xi_i) \quad \text{unless } i, i + 1 \in I.$$ 

We will call such a point $\xi$ an extension of $\{x_i : i \in I\}$.

The chain relation $CF \subset X \times X$ is defined by $(x, y) \in CF$ when for every $\epsilon > 0$ there exists an $\epsilon$ chain $\{x_0, \ldots, x_n\}$ (0 $< n < \infty$) such that $x_0 = x$ and $x_n = y$, i.e. there is an $\epsilon$ chain from $x$ to $y$. $CF$ is a closed, transitive relation containing $F$. A point $x$ is called chain recurrent if $(x, x) \in CF$. It is easy to see that $x$ is chain recurrent iff for every $\epsilon > 0$ there exists $\xi \in X^\mathbb{Z}$, an $\epsilon$ chain with $\xi_0 = x$ which is periodic, i.e. $\xi_{i+n} = \xi_i$ for some $n > 0$ and all $i \in \mathbb{Z}$. The set of chain recurrent points is denoted $[CF]$. It is a closed subset of $X$ on which $(CF) \cap (CF^{-1})$ is a closed equivalence relation. The equivalence classes are called the basic sets for $F$ (see Akin (1993), Chapters 1-3).

$F$ is called chain transitive if $CF = X \times X$, i.e. every point is chain related to every other. A closed subset $A$ of $X$ is called a chain transitive subset if the closed relation $F_A = F \cap (A \times A)$ is a chain transitive relation on $A$. For example, by Akin (1993), Theorem 4.5, any basic $A$ for $F$ is a chain transitive subset.

### 4.1. Lemma

For a closed relation $F$ on $X$

$$|CF| \subset \bigcap_{n \in \mathbb{Z}} (CF)^n(X) = \pi_0(X_F). \tag{4.5}$$

Any basic set for $F$ is a surjective subset. In fact,

$$B = F(B) \cap F^{-1}(B) = (CF(B)) \cap (CF^{-1}(B)). \tag{4.6}$$

**Proof.** It is easy to check that $CF(X) = F(X)$ (see Akin (1993), Exercise 2.3e) and so, for all $n \in \mathbb{Z}$, $(CF)^n(X) = F^n(X)$. The equality in (4.5) follows from (1.10). Clearly, $x \in CF(x)$ implies $x \in (CF)^n(X)$ for all $n$.

A basic set $B$ is a chain transitive subset, and so $B = |CF(B)|$, where $F_B = F \cap (B \times B)$ is the closed relation obtained by restricting to $B$. By (4.5) applied to $F_B$ we see that $F_B$ is a surjective relation, and so $B$ is a surjective subset. Since $F \subset CF$ we have $B \subset F(B) \cap F^{-1}(B) \subset (CF(B)) \cap (CF^{-1}(B))$. On the other hand,
if \((x_1, x), (x, x_2) \in CF\) for \(x_1, x_2 \in B\), then, because \(B\) is a basic set, \((x_2, x_1) \in CF\), and so the three points \(x_1, x_2, x\) are all \(CF\) equivalent. That is, \(x \in B\). So (4.6) follows.

4.2. Lemma. (a) Assume \(\epsilon > 0\) and \(n\) is a positive integer. There exists \(\epsilon_1 > 0\) so that if \(\{y_i : i \in [-k, k]\}\) is an \(\epsilon_1\) chain with \(k \leq n\), then there exists \(\xi \in X_F\) such that \(d(y_i, \xi) < \epsilon\) for all \(i \in [-k, k]\).

(b) Assume \(\{x_i : i \in [-l, l]\}\) is a \(0\) chain, \(\epsilon > 0\) and \(n\) is a positive integer. Then there exists \(\epsilon_1 > 0\) so that if \(\{y_i : i \in [-k, k]\}\) is an \(\epsilon_1\) chain with \(l \leq k \leq n\) such that \(y_i = x_i\) for all \(i \in [-l, l]\), then there exists \(\xi \in X_F\) such that \(\xi = x_i\) for all \(i \in [-l, l]\) and \(d(y_i, \xi) < \epsilon\) for all \(i \in [-k, k]\).

Proof. (a) Let \(X_{F, \delta}\) denote the set of \(\delta\) chains in \(X^Z\). As \(\delta\) decreases to 0 this family of compacta decreases to \(X_F\). So there exists \(\epsilon_1 > 0\) so that \(X_{F, \epsilon_1}\) is contained in the open \(\tilde{\epsilon} = \min(\epsilon, 1/n)\) neighborhood of \(X_F\). Given an \(\epsilon_1\) chain, \(\{y_i : i \in [-k, k]\}\) let \(\eta\) be its extension in \(X_{F, \epsilon_1}\). There exists \(\xi \in X_F\) such that \(d(\eta, \xi) < \tilde{\epsilon}\), and as in (1.14) \(d(\eta, \xi) < \epsilon\) for all \(i \in [-n, n]\).

For (b) repeat the proof, intersecting \(X_{F, \delta}, X_F\) and \(X_{F, \epsilon_1}\) with the closed set \(\{\xi \in X^Z : \xi = x_i\) for all \(i \in [-l, l]\}\).

4.3. Theorem. Let \(F\) be a closed surjective relation on \(X\), i.e. \(F(X) = X = F^{-1}(X)\). Let \(s_F\) be the shift homeomorphism on the sample path space \(X_F\) and \(\pi_0\) be the time zero projection of \(X_F\) onto \(X\).

(a) For a point \(\xi \in X_F\) the following conditions are equivalent:

1. \(\xi\) is \(s_F\) chain recurrent, i.e. \(\xi \in |CS_F|\).
2. \(\xi_i \in CF(\xi_{i+1})\) for all \(i \in \mathbb{Z}\).
3. There exists a basic set \(B\) for \(F\) such that \(\xi_i \in B\) for all \(i \in \mathbb{Z}\).

(b) For points \(\xi, \eta\) in \(|CS_F|\) we have \(\xi \in CS_F(\eta)\) if and only if \(\pi_0(\xi) \in CF(\pi_0(\eta))\).

In general, \(\pi_0 \times \pi_0\) maps \(CS_F\) onto \(CF\), i.e.

\[
(\pi_0 \times \pi_0)(CS_F) = CF.
\]

(c) If \(B\) is a basic set for \(F\) then \(X_{F, 0} = X_F \cap B^Z\) is a basic set for \(s_F\) mapping onto \(B\) via \(\pi_0\). Every basic set for \(s_F\) is constructed this way. Thus, \(\pi_0\) maps distinct \(s_F\) basic sets onto distinct \(F\) basic sets and

\[
\pi_0(|CS_F|) = |CF|.
\]

(d) \(F\) is a chain transitive relation on \(X\) if and only if \(s_F\) is a chain transitive homeomorphism on \(X_F\).

Proof. If \(\xi \in X_F\) then \(\xi_1 \in F(\xi_0)\) and so \(\pi_0(s_F(\xi)) \in F(\pi_0(\xi))\). With the metric on \(X^Z\) given by (1.13) the map \(\pi_0\) is distance nonincreasing (by (1.14)), so the image under \(\pi_0\) of an \(\epsilon\) chain for \(s_F\) in \(X_F\) is an \(\epsilon\) chain for \(F\) in \(X\). In particular, we have

\[
(\pi_0 \times \pi_0)(CS_F) \subset CF.
\]

So if two points in \(X_F\) are \(CS_F\) equivalent then their images in \(X\) are \(CF\) equivalent. It follows that for any \(s_F\) basic set \(B \subset X_F\) there is a unique \(F\) basic set \(B \subset X\) such that \(\pi_0(B) \subset B\).

Now we describe a construction we will call the lifting trick. Given \(\epsilon > 0\), choose an integer \(n > (1/\epsilon) + 1\). Given fixed 0 chains \(\{x_i : i \in [-l_1, l_1]\}\) and \(\{y_j : j \in [-l_2, l_2]\}\) with \(l_1, l_2 \leq n\), choose \(\epsilon_1 > 0\) small enough to satisfy the
conditions of Lemma 4.2a for $\epsilon/2$ and $n$ and Lemma 4.2b for $\epsilon/2$, $n$ and each of the
two 0 chains.

Now suppose $\xi$ is an $\epsilon_1$ chain in $X^2$ such that for some integer $N > l_1 + l_2$,

$$
\xi_i = x_i, \quad i \in [-l_1, l_1],
\xi_{i+N} = y_i, \quad i \in [-l_2, l_2].
$$

(4.10)

By Lemma 4.2 we can choose a sequence $\eta^0, \ldots, \eta^N$ in $X_F$ so that $d(\eta^i, s^k(\xi)_i) < 
\epsilon/2$ for $i \in [-n, n]$, $\eta^0 = \xi_i$ for $i \in [-l_1, l_1]$ and $\eta^N = \xi_{i+N}$ for $i \in [-l_2, l_2]$.

Observe that for $0 \leq k < N$ and $i \in [-N, n - 1]$

$$
d(s_F(\eta^k)_i, \eta^{k+1}_i) \leq d(\eta^{k+1}_i, \xi_{i+k+1}) + d(\eta^k_i, \xi_{i+k+1}).
$$

(4.11)

So, since $n - 1 > 1/\epsilon$, we have $d(s_F(\eta^k), \eta^{k+1}) \leq \epsilon$ by (1.14). This says that

$$
\eta^0 = x_i, \quad i \in [-l_1, l_1],
\eta^N = y_i, \quad i \in [-l_2, l_2].
$$

(4.12)

Now we prove the theorem.

(d) If $s_F$ is chain transitive, i.e. $C s_F = X_F \times X_F$, then (4.9) implies that

$CF = X \times X$ because $F$ is surjective.

Conversely, assume $F$ is chain transitive, $\zeta^1, \zeta^2 \in X_F$ and $\epsilon > 0$. We choose

$n > (1/\epsilon) + 1$, and let $l_1 = l_2 = n$, $x_i = \zeta^1_i$ and $y_i = \zeta^2_i$ for $i \in [-n, n]$. Since $\zeta^1_n$ and

$\zeta^2_n \in X$ and $F$ is chain transitive, we can build a sequence $\xi$ so that $\xi_i = \zeta^1_i$ for

$i \in [-n, n]$, and $\{\xi_n, \xi_{n+1}, \ldots, \xi_{n+M}\}$ is an $\epsilon_1$ chain for $F$ in $X$ with $\xi_n = \zeta^1_n$ and

$\xi_{n+M} = \zeta^2_n$. Let $\xi_{n+m+j} = \zeta^j_{n}$ for $j = 0, \ldots, 2n$, and extend to get $\xi \in X^Z$.

Then use the lifting trick construction, choosing $\eta^0 = \zeta^1$ and $\eta^N = \zeta^2$, where

$N = 2n + M$. We have constructed an $\epsilon$ chain for $s_F$ from $\zeta^1$ to $\zeta^2$. As $\epsilon$ was

arbitrary it follows that $\zeta^2 \in C s_F(\zeta^1)$, and so $s_F$ is chain transitive.

(a) (1) $\Rightarrow$ (3): If $\hat{B}$ is an $s_F$ basic set then by (4.9) there is a basic set $B$ for $F$
such that $\pi_0(B) \subset B$. For a map (though not in general for a closed relation) any

basic set is invariant (see Akin (1993), Corollary 4.11). So $\xi \in \hat{B}$ implies $s_F^k(\xi) \in \hat{B}$

for all $k \in \mathbb{Z}$. Hence $\hat{\xi}_k = \pi_0(s_F^k(\xi)) \in \hat{B}$ for all $k \in \mathbb{Z}$.

(3) $\Rightarrow$ (1): On the basic set $\hat{B}$ the closed relation $F_{\hat{B}}$ is chain transitive. By

(4.6), part (d) proved above applies to $F_{\hat{B}}$, and so $s_F$ restricted to $X_{F_{\hat{B}}}$ is chain

transitive. Furthermore, $\pi_0(X_{F_{\hat{B}}}) = B$. Since $(\xi_i, \xi_{i+1}) \in F \cap (B \times B) = F_B$

for all $i \in \mathbb{Z}$, $\xi \in X_{F_{\hat{B}}}$. Every point of $X_{F_{\hat{B}}}$ is chain recurrent because $X_{F_{\hat{B}}}$ is a chain

transitive subset.

(3) $\Rightarrow$ (2): Obvious.

(2) $\Rightarrow$ (3): $\xi_{i+1} \in F(\xi_i)$ and $F \subset CF$. So, by assumption of (2), $\xi_i$ and $\xi_{i+1}$

are $CF$ equivalent for all $i$. As $CF$ is transitive, this means that all the $\xi_i$’s are $CF$

equivalent to one another, and so lie in some common $F$ basic set.

(c) If $B$ is an $F$ basic set then, as above, $X_{F_B} = X_F \cap B^Z$ is a chain transitive

subset mapping onto $B$. So it is entirely contained in some $s_F$ basic set $\hat{B}$. But $\pi_0$

maps $\hat{B}$ into some $F$ basic set. As distinct basic sets are disjoint, it must be that

$\pi_0(\hat{B}) \subset B$. Then $\hat{B} \subset X_F \cap B^Z$, and so we have

$$
\hat{B} = X_F \cap B^Z = X_{F_B}
$$

(4.13)

is a basic set mapping onto $B$. Any other basic set mapping into $B$ is contained in

$X_F \cap B^Z$ by part (a), and so equals $\hat{B}$ because distinct basic sets are disjoint.
Thus, $\pi_0$ maps distinct basic sets of $s_F$ to distinct basic sets of $F$, and (4.8) holds.

(b) If $\xi \in \mathcal{C}F(\eta)$, then $\pi_0(\xi) \in \mathcal{C}F(\pi_0(\eta))$ by (4.9). If $\xi, \eta \in |\mathcal{C}F|$, then by (a) all the $\xi_i$’s lie in some common basic set $B_1$ and all the $\eta_j$’s lie in some common basic set $B_2$. If $\xi_0 \in \mathcal{C}F(\eta_0)$ then for all $i, j \in \mathbb{Z}$ we have $\xi_i, \eta_j \in \mathcal{C}F(\eta_0)$ by transitivity of the relation $\mathcal{C}F$. So for any $\epsilon_1 > 0$ and positive integer $n$ we can connect $\eta_0$ to $\xi_{-n}$ via an $\epsilon_1$ chain. Then, as in the proof of (d), use the lifting trick to connect $\eta$ to $\xi$ via an $\epsilon$ chain for $s_F$. In general, if $y \in \mathcal{C}F(x)$ then we can connect $x$ to $y$ by an $\epsilon_1$ chain and use the lifting trick again, this time with $l_1 = l_2 = 0$, to get an $\epsilon$ chain for $\{\eta^0, \ldots, \eta^N\}$ for $s_F$ such that $\eta^0 = x$ and $\eta^N = y$. This proves (4.7). $\square$

For any continuous map $f$ on $X$ the omega limit point set of $x \in X$, $\omega_f(x)$, is the set of limit points of the orbit sequence $\{f^k(x) : k = 0, 1, \ldots\}$.

**4.4. Corollary.** Let $\xi \in X^Z$ be an asymptotic chain for a closed surjective relation $F$ on $X$,

\[(4.14) \quad \text{Lim}_{k \to -\infty} d(s^k(\xi), X_F) = 0.\]

The limit point set $\omega_s(\xi)$ is an $s_F$-chain transitive subset of $X_F$ and so is contained in some $s_F$-basic set. The image $\pi_0(\omega_s(\xi))$ is the set of limit points of the sequence $\{\xi_k : k = 0, 1, \ldots\}$ in $X$. It is an $F$ chain transitive subset of $X$ and so is contained in some $F$ basic set.

**Proof.** Given $\epsilon > 0$, choose $n > 1/\epsilon$ and then $\epsilon_1 > 0$ to satisfy the condition of Lemma 4.2a. Because $\xi$ is an asymptotic chain, there exists $N \in \mathbb{Z}$ such that $d((\xi_{i+1}, F(\xi_i))) < \epsilon_1$ for $i \geq N$. Hence, for $k > N + n$, $\{s^k(\xi)_i : i \in [-n, n]\}$ is an $\epsilon_1$ chain in $X$ for $F$. By Lemma 4.2a there is $\eta^k \in X_F$ such that $d(s^k(\xi)_i, \eta^k_i) < \epsilon$ for $i \in [-n, n]$, and so $d(s^k(\xi), X_F) < \epsilon$ for all $k > N$ by (1.14). This proves (4.14), which implies $\omega_s(\xi) \subseteq X_F$. By Akin (1993), Proposition 4.14, $\omega_s(\xi)$ is a chain transitive subset of $X^Z$. Since it is in fact a subset of $X_F$ on which $s = s_F$, it follows that $\omega_s(\xi)$ is contained in some basic set for $s_F$ in $X_F$. By compactness the image under $\pi_0$ of the limit point set of the sequence $\{s^k(\xi)\}$ is the limit point set of the sequence $\{\xi_k = \pi_0(s^k(\xi))\}$. By (4.9) $\pi_0(\omega_s(\xi))$ is an $F$ chain transitive subset of $X$ and so is contained in an $F$ basic set. $\square$

To apply all these results when the closed relation $F$ is not surjective we restrict to the dynamic domain $\pi_0(X_F)$. By Lemma 1.2, $\pi_0(X_F) = \bigcap_{n \in \mathbb{Z}} F^n(X)$ and the restriction $F_{\pi_0}(X_F)$ is a surjective relation on $\pi_0(X_F)$. Furthermore, by (4.5) and Lemma 1.2, applied to $\mathcal{C}F$, $|\mathcal{C}F| \subset \pi_0(X_F)$ and

\[(4.15) \quad \pi_0(X_F) = \mathcal{C}F(\pi_0(X_F)) \cap \mathcal{C}F^{-1}(\pi_0(X_F)).\]

From Theorem 4.5 of Akin (1993) it follows that

\[(4.16) \quad \mathcal{C}(F_{\pi_0}(X_F)) = (\mathcal{C}F)_{\pi_0}(X_F) = (\mathcal{C}F) \cap (\pi_0(X_F) \times \pi_0(X_F)).\]

We conclude this section by observing that, while chain transitivity for $s_F$ and for $F$ are equivalent, the same is not true for other related ideas.

Let $f$ be a topologically transitive homeomorphism on a space $X$. Suppose that $f$ has a unique proper minimal subset, a fixed point $e \in X$. For example, the *stopped torus map* described in Chapter 9 of Akin (1993) is such a map. Let $F = f \cup \{e\} \times X$. The only closed subsets $A$ of $X$ such that $F(A) \subset A$ are $\emptyset$ and $X$. So $F$ is a “minimal” relation. The sample path space $X_F$ consists of $X_f$ together
with sequences ξ ∈ X^Z such that for some n ∈ Z we have ξ_i = e for all i < n and ξ_{i+1} = f(ξ_i) for all i > n. For any ξ ∈ X_F the limit sets ω_F(ξ) and ω(s_F(ξ)) are contained in X_f, and so the orbit of ξ is not dense in X_F. Thus, s_F is not even topologically transitive.

5. Approximating Invariant Measures

We relate chains to invariant measures, first defining for any finite sequence \{x_0, \ldots, x_{n-1}\} in X the finite measure

\[\sigma(x_0, \ldots, x_{n-1}) = \frac{1}{n} \sum_{i=0}^{n-1} \delta_{x_i}.\]

Recall that \(\mu \in P(X)\) is called an \(\epsilon\) invariant measure for \(F\) if there exists \((\mu_1, \mu_2) \in F_*\) such that \(d(\mu_1, \mu), d(\mu_2, \mu) \leq \epsilon\).

5.1. Lemma. If \(\{x_0, \ldots, x_n\}\) is an \(\epsilon\) chain for \(F\), then \(\sigma(x_0, \ldots, x_{n-1})\) is an \((\epsilon + \frac{1}{n}d(x_0, x_n))\) invariant measure for \(F\). In particular, if \(x_0 = x_n\), then \(\sigma(x_0, \ldots, x_{n-1})\) is an \(\epsilon\) invariant measure for \(F\).

Proof. There exists \(\{z_1, \ldots, z_n\}\) such that \(z_{i+1} \in F(x_i)\) and \(d(z_{i+1}, x_{i+1}) \leq \epsilon\) for \(i = 0, \ldots, n-1\). So the pairs \((\delta_{x_i}, \delta_{z_{i+1}}) \in F_*\) and \(d(\delta_{z_{i+1}}, \delta_{x_{i+1}}) \leq \epsilon\) for \(i = 0, \ldots, n-1\). By convexity of \(F_*\) and convexity properties of the Hutchinson metric we have

\[(\sigma(x_0, \ldots, x_{n-1}), \sigma(z_1, \ldots, z_n)) \in F_*,\]

and

\[d(\sigma(z_1, \ldots, z_n), \sigma(x_1, \ldots, x_n)) \leq \epsilon.\]

On the other hand, by definition of the Hutchinson metric

\[d(\sigma(x_0, \ldots, x_{n-1}), \sigma(x_1, \ldots, x_n)) = \frac{1}{n} d(\delta_{x_0}, \delta_{x_n}) = \frac{1}{n} d(x_0, x_n).\]

So the result follows from (5.3) and (5.4) by the triangle inequality.

If \(f\) is a continuous map on \(X\), we define

\[\sigma_n(x, f) = \sigma(x, f(x), \ldots, f^{n-1}(x)) = \frac{1}{n} \sum_{i=0}^{n-1} \delta_{f^i(x)}.\]

By Lemma 5.1, \(\sigma_n(x, f)\) is a \(\frac{1}{n}d(x, f^n(x))\) invariant measure for \(f\). The set of limit points of the sequence \(\{\sigma_n(x, f) : n = 1, 2, \ldots\}\) in \(P(X)\) is denoted \(M(x, f)\). It is a nonempty, compact connected set of invariant measures for \(f\), all of whose supports lie in the omega set \(\omega(f)\) (see, e.g., Akin (1993), Propositions 8.3 and 8.8). We call \(x\) a convergence point for \(f\) when \(M(x, f)\) consists of a single measure, then denoted \(\mu_x\). So \(x\) is a convergence point exactly when the sequence \(\{\sigma_n(x, f)\}\) converges in \(P(X)\). The limit is then \(\mu_x\). For example, if \(x\) is a periodic point then it is a convergence point. In fact,

\[f^n(x) = x \Rightarrow \mu_x = \sigma_n(x, f).\]
5.2. Lemma. Let \( \{x_0, \ldots, x_n-1\} \) and \( \{y_0, \ldots, y_{N-1}\} \) be sequences in \( X \) and let \( \epsilon > 0 \). Assume there is an injection \( g : \{0, \ldots, n-1\} \to \{0, \ldots, N-1\} \) such that
\[
(5.7) \quad d(x_i, y_{g(i)}) \leq \frac{\epsilon}{2}, \quad i = 0, \ldots, n-1,
\]
and that with \( D \) the diameter of the space \( P(X) \), with respect to the Hutchinson metric,
\[
(5.8) \quad \frac{N-n}{N} \cdot D < \frac{\epsilon}{2}.
\]
Then we have
\[
(5.9) \quad d(\sigma(x_0, \ldots, x_{n-1}), \sigma(y_0, \ldots, y_{N-1})) < \epsilon.
\]

Proof. Let \( \tilde{\sigma} = \sigma(y_{g(0)}, \ldots, y_{g(n-1)}) \in P(X) \). From (5.7) we have
\[
d(\sigma(x_0, \ldots, x_{n-1}), \tilde{\sigma}) \leq \frac{\epsilon}{2}.
\]
Now call the image set of \( g \) in \( \{0, \ldots, N-1\} \) the set of good values and the rest of \( \{0, \ldots, N-1\} \) the set of bad values. Clearly,
\[
(5.10) \quad \sigma(y_0, \ldots, y_{N-1}) = \frac{1}{N} (\sum_{j \text{ bad}} \delta_{y_j} + n\tilde{\sigma}).
\]
If we replace each \( \delta_{y_j} \) (j bad) by \( \tilde{\sigma} \) we change the right side to \( \tilde{\sigma} \). Hence,
\[
(5.11) \quad d(\tilde{\sigma}, \sigma(y_0, \ldots, y_{N-1})) \leq \frac{1}{N} \sum_{j \text{ bad}} d(\tilde{\sigma}, \delta_{y_j}),
\]
which is less than \( \frac{\epsilon}{2} \) by (5.8). From the triangle inequality, (5.9) follows. \( \square \)

5.3. Theorem. Assume that the closed relation \( F \) on \( X \) is chain transitive, i.e. \( CF = X \times X \). Let \( x \in X \). Let \( X_F \) be the sample path space with \( s_F \) the restriction to \( X_F \) of the shift homeomorphism \( s \) on \( X \).

(a) For a measure \( \nu \in P(X^Z) \) the following conditions are equivalent:

1. \( \nu \) is an invariant sample path measure, i.e. \( |\nu| \subset X_F \) and \( s_F^*(\nu) = \nu \).
2. There exists \( \xi \in X^Z \), an asymptotic chain for \( F \), such that \( \xi \) is a convergence point for the shift \( s \) and \( \mu_\xi = \nu \).
3. For every \( \epsilon > 0 \) there exists \( \xi \in X^Z \), a periodic \( \epsilon \) chain for \( F \), such that \( \xi_0 = x \) and \( d(\mu_\xi, \nu) \leq \epsilon \).
4. For every \( \epsilon > 0 \) and positive integer \( N \) there exist \( \xi \in X^Z \), an \( \epsilon \) chain for \( F \), and \( n \geq N \) such that \( d(\sigma_n(\xi, s), \nu) \leq \epsilon \).

(b) For a measure \( \mu \in P(X) \) the following conditions are equivalent:

1. \( \mu \) is an invariant measure for \( F \), i.e. \( \mu \in P_F(X) \).
2. There exists \( \xi \in X^Z \), an asymptotic chain for \( F \), such that
\[
\mu = \lim_{n \to \infty} \sigma(\xi_0, \ldots, \xi_{n-1}).
\]
3. For every \( \epsilon > 0 \), there exists an \( \epsilon \) chain \( \{x_0, \ldots, x_n\} \) for \( F \) with \( x_0 = x_n = x \) and such that \( d(\sigma(x_0, \ldots, x_{n-1}), \mu) \leq \epsilon \).
4. For every \( \epsilon > 0 \) and positive integer \( N \) there exists an \( \epsilon \) chain \( \{x_0, \ldots, x_n\} \) with \( n \geq N \) and such that \( d(\sigma(x_0, \ldots, x_{n-1}), \mu) \leq \epsilon \).
Proof. Notice first that since all of \( X \) consists of a single \( F \) basic set, \( F(X) = X = F^{-1}(X) \) by Lemma 4.1.

(a) We prove first that conditions (2), (3) and (4) are equivalent, and then that the set of measures satisfying these conditions is exactly the \( s_F \) invariant measures on \( X_F \).

(2) \( \Rightarrow \) (4): Let \( \xi \) be an asymptotic chain which is a convergence point for \( s \) with \( \mu_\xi = \nu \). Given \( \epsilon > 0 \), there exists \( N_0 \) such that \( d(\xi_{i+1}, F(\xi_i)) < \epsilon \) for \( i \geq N_0 \). Thus, \( \{\xi_i : i \geq N_0\} \) is an \( \epsilon \) chain for \( F \). Let \( \eta \in X^\mathbb{Z} \) be an extension of this sequence, i.e. \( \eta_i = \xi_i \) for \( i \geq N_0 \) and \( \eta_i \in F(\eta_{i-1}) \) for \( i \leq N_0 \). Then \( \eta \) is an \( \epsilon \) asymptotic chain for \( F \), and, since \( \eta_i = \xi_i \) for \( i \geq N_0 \),

\[
\lim_{k \to \infty} d(s^k(\eta), s^k(\xi)) = 0,
\]

i.e. the orbit sequences for \( \xi \) and \( \eta \) are asymptotic in \( X^\mathbb{Z} \). It easily follows that

\[
\lim_{k \to \infty} d(s_k(\eta), s_k(\xi)) = 0,
\]

and so \( \eta \) is a convergence point for \( s \) with \( \mu_\eta = \mu_\xi = \nu \). Since the sequence \( \{s_n(\eta, s)\} \) converges to \( \nu \), condition (4) is clear.

(4) \( \Rightarrow \) (3): Given \( \epsilon > 0 \), we can use compactness and \( F(X) = X \) to choose a finite subset \( K \) such that \( K \) and \( F(K) \) are \( \epsilon \) dense subsets of \( X \), i.e. for every \( y \in X \) there exist \( z_0, z_1 \in K \) such that \( d(z_0, y) < \epsilon \) and \( d(z_1, y) < \epsilon \) for some \( z_1 \in F(z_1) \). For each point \( z \) of \( K \) choose an \( \epsilon \) chain from \( z \) to \( x \) and an \( \epsilon \) chain from \( x \) to \( z \), and let \( L \) be the maximum of the lengths of these chains. Let \( D \) be the diameter of \( P(X^\mathbb{Z}) \). Choose positive integers \( N_0 \) and \( N_1 \) so that

\[
\frac{1}{N_0} < \frac{\epsilon}{4}, \quad N_1 > N_0,
\]

\[
\frac{2(N_0 + L + 1)}{N_1} D < \frac{\epsilon}{4}.
\]

By (4) there exist an \( \epsilon/2 \) chain \( \eta \in X^\mathbb{Z} \) for \( F \) and \( n \geq N_1 \) such that \( d(s_n(\eta, s), \nu) < \epsilon/2 \). By definition of \( K \) we can choose \( z^- \), \( z^+ \in K \) such that \( d(F(z^-), \nu) < \epsilon \) and \( d(F(z^+), \nu) < \epsilon \). Let \( \{x_0, \ldots, x_{-L-1}\} \) and \( \{x_0^+, \ldots, x_{L}^+\} \) be \( \epsilon \) chains with \( l_- \leq L \), \( x_0 = x_0^+ = x \), \( x_0^+ = z^+ \) and \( x_0^- = z^- \). Now we build a periodic chain \( \tilde{\eta} \) as follows: Let \( \tilde{\eta}_i = \eta_{i} \) for \( i \in [-N_0, n + N_0 - 1] \). For \( i \in [-N_0 - L, -1, -N_0 - 1] \) let \( \tilde{\eta}_i \) be a shift of the \( x^- \) chain from \( x \) to \( x^- \). For \( i \in [n + N_0, n + N_0 + L] \) let \( \tilde{\eta}_i \) be a shift of the \( x^+ \) chain from \( z_+ \) to \( x \). In particular,

\[
\tilde{\eta}_{-N_0-L-1} = \eta_{n+N_0+L},
\]

and so we can extend \( \tilde{\eta} \) to define an \( N_3 \) periodic \( \epsilon \) chain for \( F \), where \( N_3 = n + 2N_0 + L + L_+ + 1 \). Because \( N_0^{-1} < \epsilon/4 \) we have from (1.14)

\[
d(s^i(\eta), s^i(\tilde{\eta})) < \epsilon/4 \text{ for } i = 0, \ldots, n_1 - 1.
\]

Now we use Lemma 5.2 to compare the sequences \( \{s^i(\eta) : i \in [0, n_1 - 1]\} \) and \( \{s^i(\tilde{\eta}) : [0, N_3 - 1]\} \) with \( y(i) = i \) for \( i \in [0, n_1 - 1] \). Since \( N_3 - n_1 \leq 2(N_0 + L + 1), \) (5.16) and (5.14) imply (5.7) and (5.8) respectively. By periodicity of \( \tilde{\eta} \), the lemma implies \( d(s_n(\eta, s), \mu_{\tilde{\eta}}) < \epsilon/2 \). So \( d(\nu, \mu_{\tilde{\eta}}) < \epsilon \). To obtain \( \xi \) required by (3) shift \( \tilde{\eta} \) to move \( x \) to position 0.

(3) \( \Rightarrow \) (2): For \( k = 1, 2, \ldots \) choose a periodic \( 2^{-k} \) chain \( \xi_k \) in \( X^\mathbb{Z} \) with \( \xi_k = x \) and \( d(\mu_k, \nu) \leq 2^{-k} \), where \( \mu_k = \mu_{\xi_k} \). So if \( n_k \) is a period for \( \xi_k \), i.e. a positive integer such that \( \xi_{i+n_k} = \xi_i \) for all \( i \in \mathbb{Z} \), then \( \mu_k = s_{n_k}(\xi_k) \). For \( N_k \) any increasing
sequence of positive integers we can define \( \eta \) to be the extension to an element of \( X^Z \) of the sequence \( \{ \eta_i : i = 0, 1, \ldots \} \) defined by first going around the length \( n_1 \) initial loop of \( \xi^1 \) \( N_1 \) times, then the length \( n_2 \) initial loop of \( \xi^2 \) \( N_2 \) times, and so on. Clearly the resulting \( \eta \) is an asymptotic chain. We show that when \( \{ N_k \} \) is increasing fast enough, \( \eta \) is a convergence point for \( s \) with \( \mu_\eta = \nu \).

Since any multiple of a period is a period, we can assume that \( n_k > 2^{k+1} \) for \( k = 1, 2, \ldots \). Now let \( N_1 \geq 4(2n_1 + 2n_2) \) and inductively choose \( N_k \) \((k>1)\) so that

\[ \frac{(\sum_{i=1}^{k-1} n_i N_i) + 2n_k + 2n_{k+1}}{N_k} \leq \frac{1}{2} \cdot 2^{-k}. \]  

(5.17)

Now suppose \( N \) is between \( \sum_{i=1}^{k} n_i N_i \) and \( \sum_{i=1}^{k+1} n_i N_i \). We regard the “bad values” of \( \{ 0, \ldots , N - 1 \} \) as all those up to \( \sum_{i=1}^{k} n_i N_i \), the first and last loops of \( \xi^k \), the \( \xi^{k+1} \) loop in which \( N \) occurs and, if this one is not initial, the initial \( \xi^{k+1} \) loop as well. For each remaining, “good”, value \( j \), \( s^j(\eta) \) on the interval \([-n_k, n_k)\) looks like one of the \( N_k - 2 \) interior loops of \( \xi^k \) or else like a piece of some loops of \( \xi^{k+1} \).

So by Lemma 5.2, \( \sigma_N(\eta, s) \) is at most \( 2^{-k} \) away from some convex combination of \( \mu_k \) and \( \mu_{k+1} \). Each of these is at most \( 2^{-k} \) away from \( \nu \), and so \( \sigma_N(\eta, s) \) is at most \( 2^{-(k-1)} \) away from \( \nu \). Thus the sequence \( \{ \sigma_N(\eta, s) \} \) converges to \( \nu \) as \( N \to \infty \).

This completes the proof that conditions (2), (3) and (4) on \( \nu \in P(X^Z) \) are equivalent. If \( \nu = \mu_\xi \) with \( \xi \) a convergence point for \( s \), then \( \nu \) is \( s \) invariant and \( |\nu| \subset \omega s(\xi) \). So if \( \xi \) is an asymptotic chain, then \( |\nu| \subset X_\xi \) by Corollary 4.4. Thus, such measures are \( s_\xi \) invariant measures, i.e. (1) holds for such \( \nu \)'s. If \( \nu \) is an ergodic invariant measure for \( s_\xi \) then by the Birkhoff Ergodic Theorem there exists \( \xi \in X_\xi \), a convergence point for \( s_\xi \), with \( \mu_\xi = \nu \). In fact, the set of such \( \xi \)'s has \( \nu \) measure 1. So an ergodic \( s_\xi \) invariant measure certainly satisfies condition (2). Every invariant measure is a weighted average of the ergodic measures. That is, the set of \( s_\xi \) invariant measures is the smallest closed, convex subset of \( P(X_\xi) \) which contains the ergodic measures. Clearly, the set of measures in \( P(X_\xi) \) which satisfy (3) or (4) is a closed set. We complete the proof by showing that the set of measures satisfying (3) is convex. By closure it is enough to show that if \( \nu^1, \nu^2 \) satisfy (3) then the midpoint \( \frac{1}{2} \nu^1 + \frac{1}{2} \nu^2 \) satisfies (3).

Given \( \epsilon > 0 \), choose periodic \( \epsilon/2 \) chains \( \xi^1, \xi^2 \) for \( F \) in \( X^Z \) with \( \xi_0^1 = \xi_0^2 = x \) and \( d(\mu_{\xi^1}, \nu^i) \leq \frac{\epsilon}{4} \) for \( i = 1, 2 \). By multiplying a period of \( \xi^1 \) times a period of \( \xi^2 \) and then multiplying by some large positive integer we can choose a positive integer \( N_0 \), a common period for \( \xi^1 \) and \( \xi^2 \), which satisfies \( N_0^{-1} < \frac{\epsilon}{2} \). With \( D \) the diameter of \( P(X^Z) \) choose \( N_1 \) a positive integer such that

\[ \frac{2}{N_1} D = \frac{4N_0}{2N_1 \cdot N_0} D < \frac{\epsilon}{4}. \]  

(5.18)

Now let \( \xi \) be the sequence in \( X^Z \) of period \( 2N_1 \cdot N_0 \) consisting of \( N_1 \) copies of the initial loop of \( \xi^1 \) of length \( N_0 \) followed by \( N_1 \) copies of the initial loop of \( \xi^2 \) of length \( N_0 \). As usual we apply Lemma 5.2; the “bad values” are the first and last loops of \( \xi^1 \) and of \( \xi^2 \), a total of \( 4N_0 \) out of \( 2N_1 N_0 \). For each of the remaining values \( j \), \( s^j(\xi) \) looks on the interval \([-N_0, N_0)\) like one of the \( \xi^1 \) loops \((N_1 - 1)N_0 \) values of \( j \) or like one of the \( \xi^2 \) loops (same number of values of \( j \)). So by Lemma 5.2 we have \( d(\mu_\xi, \frac{1}{2} \mu_{\xi^1} + \frac{1}{2} \mu_{\xi^2}) < \frac{\epsilon}{2} \), and by convexity properties of the Hutchinson metric we have \( d(\mu_\xi, \frac{1}{2} \nu^1 + \frac{1}{2} \nu^2) < \epsilon. \)
(b) (1) ⇒ (2), (3), (4): If \( \mu \) is an \( F \) invariant measure then by Theorem 3.2 there is an \( s_F \) invariant measure \( \nu \) on \( X_F \) such that \( \mu = \pi_0(\nu) \). Apply (2), (3) and (4) from (a) to \( \nu \) and project by \( \pi_{0*} \) to get (2), (3) and (4) respectively for \( \mu \).
(2) ⇒ (4) and (3) ⇒ (4): Obvious.
(4) ⇒ (1): By using (4) and extending to sequences in \( X^Z \) we can choose for \( k = 1, 2, \ldots \) a \( 1/k \) chain \( \xi^k \in X^Z \) and \( n_k > k \) such that \( d(\sigma(\xi^k_0, \ldots, \xi^k_{n_k-1}), \mu) \leq \frac{1}{k} \).
By going to a subsequence we can assume that \( \{\sigma_{n_k}(\xi^k, s): k = 1, 2, \ldots\} \) converges to some measure \( \nu \in P(X^Z) \). As \( \nu \) satisfies (4) of part (a), it is an \( s_F \) invariant measure on \( X_F \). Since \( \pi_{0*}(\sigma_{n_k}(\xi^k, s)) = \sigma(\xi^k_0, \ldots, \xi^k_{n_k-1}) \), we have \( \pi_{0*}(\nu) = \mu \).
Thus, \( \mu \) satisfies (1).

5.4. Corollary. For a closed relation \( F \) on \( X \), let \( X_F \) be the sample path space for \( F \) with \( s_F \) the restriction to \( X_F \) of the shift homeomorphism \( s \) on \( X^Z \) and with \( \pi_0: X^Z \to X \) the time zero projection.
(a) For a measure \( \nu \in P(X^Z) \) the following conditions are equivalent:

1. There exists a basic set \( B \subset X_F \) for \( s_F \) such that \( \nu \) is an invariant path measure supported by \( B \), i.e. \( \nu(B) = 1 \) and \( (s_F)_*(\nu) = \nu \).
2. There exists \( \xi \in X^Z \), an asymptotic chain for \( F \), such that \( \nu \) is among the limit measures associated with \( \xi \), i.e. \( \nu \) is a limit point for the sequence \( \{\sigma_n(\xi, s)\} \) in \( P(X^Z) \).
3. For every \( \epsilon > 0 \) there exists \( \xi \in X^Z \), a periodic \( \epsilon \) chain for \( F \), such that \( d(\mu_\xi, \nu) \leq \epsilon \).

Every ergodic measure for \( s_F \) : \( X_F \to X_F \) satisfies these conditions. Every invariant sample path measure can be decomposed as an average of such measures.
(b) For a measure \( \mu \in P(X) \) the following conditions are equivalent:
1. There exists a basic set \( B \subset X \) for \( F \) such that \( \mu \) is an invariant measure for \( F \) supported by \( B \), i.e. \( |\mu| \subset B \) and \( (\mu, \mu) \in F_* \).
2. There exist a basic set \( B \subset X_F \) for \( s_F \) and an invariant sample path measure \( \nu \) supported by \( B \) such that \( (\pi_{0*})_*(\nu) = \mu \).
3. There exists an asymptotic chain \( \{x_0, x_1, \ldots\} \) for \( F \) such that \( \mu \) is a limit point for the sequence \( \{\sigma(\{x_0, \ldots, x_n-1\})\} \) in \( P(X) \).
4. For every \( \epsilon > 0 \) there exists an \( \epsilon \) chain \( \{x_0, \ldots, x_n\} \) for \( F \) such that \( x_0 = x_n \), and \( d(\sigma(x_0, \ldots, x_n), \mu) \leq \epsilon \).

Every extreme point of the compact convex set of invariant measures for \( F \) is the image under \( \pi_{00} \) of some ergodic measure for \( s_F \), and so satisfies these conditions. Every invariant measure for \( F \) can be decomposed as the average of such measures.

Proof. (a) Applying Theorem 4.3 to the restriction of \( F \) to \( \pi_0(X_F) \), we see that a basic set \( B \) for \( s_F \) is just \( X_F \cap B^Z \), where \( B \) is the \( F \) basic set \( \pi_0(B) \). So if \( \nu \) satisfies (1) then we can apply Theorem 5.3a to \( F_B \) because the basic set \( B \) is a chain transitive subset (see Akin (1993), Theorem 4.5). We get asymptotic chains in \( B \) with limit measure \( \nu \) and periodic chains in \( B \) whose measures approximate \( \nu \). Thus, (1) ⇒ (2) and (1) ⇒ (3) are clear.
(2) ⇒ (1): Each limit measure is \( s \) invariant, and each is supported by \( \omega(s(\xi)) \), which is contained in an \( s_F \) basic set by Corollary 4.4.
(3) ⇒ (1): If \( \{\xi^k\} \) is a sequence of \( 1/k \) chains in \( X^Z \) with period \( n_k \) such that \( \{\mu_\xi^k\} \) converges to \( \nu \), then, as the limit of \( s \) invariant measures, \( \nu \) is \( s \) invariant. By going to a subsequence if necessary we can assume that the sequence of supports of \( \mu_\xi^k \), the periodic orbits, converges to a set \( A \) in the space of closed subsets of \( X^Z \).
equipped with the Hausdorff metric. By Akin (1993), Proposition 8.1, the support of the limit \( \nu \) is contained in the limit of the supports, \( A \). By Akin (1993), Exercise 7.37d \( A \), is a chain transitive subset for \( s \). Finally, it follows from Lemma 4.2a that \( A \subset X_F \). So \( A \) is contained in some basic set for \( s_F \) by chain transitivity.

The support of any ergodic measure for \( s_F \) is contained in some topologically transitive subset and so is certainly contained in some basic set (see, e.g. Akin (1993), Corollary 8.9). For any invariant measure \( \nu \) the ergodic decomposition expresses \( \nu \) as an average (i.e. an integral of a measure on the set) of ergodic measures.

(b) (1) \( \Rightarrow \) (2): If \( \mu \in F_s \) and \( |\mu| \subset B \), then there exists a measure \( \tilde{\mu} \) on \( X \times X \) with \( |\tilde{\mu}| \subset F \) and \( \pi_1.(\tilde{\mu}) = \mu = \pi_2.(\tilde{\mu}) \). Hence, \( \tilde{\mu}(B \times X) = \tilde{\mu}(X \times B) = 1 \), and so \( |\tilde{\mu}| \subset F \cap (B \times B) \). Thus, \( \mu \) is an invariant measure for \( F_B \). By Theorem 3.2 there exists an invariant measure \( \nu \) on \( X_{F_B} \) such that \( (\pi_0)_*.(\nu) = \mu \). By Theorem 4.3d applied to \( F_B \), \( X_{F_B} \) is a chain transitive subset and so is contained in (and in fact equals) a basic set \( \tilde{B} \) for \( s_F \).

(2) \( \Rightarrow \) (3),(4): \( \nu \) satisfies (1) of part (a). Applying (2) and (3) of part (a) and projecting by \( \pi_{0\nu} \), we get (3) and (4) respectively.

(3) \( \Rightarrow \) (1): By Lemma 5.1 any limit measure of the sequence is the limit of measures \( \epsilon \) invariant for \( F \) for any \( \epsilon > 0 \) and so is an \( F \) invariant measure. (To be precise you apply Lemma 5.2 to ignore the bad values, after which the sequence is an \( \epsilon/2 \) chain.) The support of \( \mu \) is contained in the set of limit points of the sequence \( \{x_0,x_1,\ldots\} \), which is, by the proof of Corollary 4.4, a chain transitive subset contained in some \( F \) basic set.

(4) \( \Rightarrow \) (2): There is a sequence \( \{\xi^k\} \) of period \( 1/k \) chains in \( X^\mathbb{Z} \) with period \( n_k \) such that \( \pi_{0\nu}(|\mu_{\xi^k}|) \) converges to \( \mu \). By going to a subsequence we can assume that \( \{\mu_{\xi^k}\} \) converges to \( \nu \in P(X^\mathbb{Z}) \), which satisfies (3) and hence (1) of (a). As \( \mu = \pi_{0\nu}.(\nu) \), \( \mu \) satisfies (2).

(2) \( \Rightarrow \) (1): Obvious, since \( \pi_{0\nu}(\tilde{B}) \) is an \( F \) basic set.

If \( \mu \) is an extreme point for the compact convex set \( P_F(X) \) then by Theorem 3.2 the set

\[
\text{C}_\mu = \{\nu : \pi_{0\nu}.(\nu) = \mu \text{ and } \nu \text{ is an } s_F \text{ invariant measure on } X_F\}
\]

is nonempty. It is clearly compact and convex, and so contains extrema. Such an extremum \( \nu \) is an extremum for the entire set of invariant measures. (A segment containing \( \nu \) maps to a segment containing \( \mu \). As \( \mu \) is extreme, the image segment is constantly \( \mu \) and so the original segment lies in \( C_\mu \). As \( \nu \) is an extremum of \( C_\mu \), the segment is constantly \( \nu \).) Thus, \( \nu \) is ergodic.

Any element of a compact convex set is an average of the extrema. \( \square \)

If \( F = f \) is a continuous map and \( \mu \) is an extreme point of \( P_f(X) \), then \( \mu \) is an ergodic measure satisfying Birkhoff’s Theorem: \( \mu \) a.e. \( x \) of \( X \) is a convergence point with \( \mu_x = \mu \). The analogue for a closed relation \( F \) on \( X \) uses the set \( C_\mu \subset P_{s_F}(X_F) \) associated to \( \mu \in P_F(X) \) by (5.19). If \( \mu \) is an extreme point of \( P_F(X) \), then, for any \( \nu \in C_\mu \), \( \nu \) a.e. \( \xi \in X_F \) is a convergence point for \( s_F \) for which \( (\pi_0)_*\mu_\xi = \mu \).

5.5. Theorem. For a closed relation \( F \) on \( X \) let \( \mu \) be an extreme point of the compact convex set \( P_F(X) \) of \( F \) invariant measures. Let \( C_\mu \) be the compact convex subset of \( P(X_F) \) consisting of \( s_F \) invariant measures \( \nu \) such that \( (\pi_0)_*\nu = \mu \). The
set of convergence points $\xi$ for $s_F$ such that $\mu_\xi \in C_\mu$ has $\nu$ measure 1 for every $\nu \in C_\mu$.

For every $\nu \in C_\mu$ and $g \in L^1(X, \mu)$,
\begin{equation}
\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} g(\xi_k) = \int_X g \, d\mu,
\end{equation}
for $\nu$ a.e. $\xi \in X_F$. In particular, the asymptotic fraction of times a $\nu \in C_\mu$ typical sample path 'visits' a Borel set $A \subset X$ is given by $\mu(A)$.

**Proof.** Integrating over the set of $s_F$ convergence points, we see from the Ergodic Decomposition Theorem that for any $\nu \in C_\mu$
\begin{equation}
\nu = \int \mu_\xi \nu(d\xi).
\end{equation}
Applying the continuous linear map $(\pi_0)_*$, we get
\begin{equation}
\mu = (\pi_0)_*(\nu) = \int (\pi_0)_*(\mu_\xi) \nu(d\xi).
\end{equation}
Because $\mu$ is an extreme point it follows that the measures $(\pi_0)_*(\mu_\xi)$ in $P_F(X)$ are equal to $\mu$ for $\nu$ a.e. $\xi$. That is, $\mu_\xi \in C_\mu$ for $\nu$ a.e. $\xi$. The $L^1$ result follows similarly.

**Acknowledgement**

The authors are grateful to F. Y. Hunt for bringing the paper Aubin, Frankowska and Lasota (1991) to our attention.

**References**


**Department of Mathematics, Howard University, Washington, D.C. 20059**

**Department of Mathematics, The City College, New York, New York 10031**