Notes on Chain Recurrence and Lyapunov Functions

John Franks

April 25, 2017

Abstract

This short expository note provides an introduction to the concept of chain recurrence in topological dynamics and a proof of the existence of complete Lyapunov functions for homeomorphisms of compact metric spaces due to Charles Conley [C]. I have used it as supplementary material in introductory dynamics courses.

1 Epsilon Chains

We briefly review the definition of \( \varepsilon \)-chains and chain recurrence developed by Charles Conley in [C]. In the following \( f : X \to X \) will denote a homeomorphism of a compact metric space \( X \).

**Definition 1.1.** An \( \varepsilon \)-chain from \( x \) to \( y \) for \( f \) is a sequence of points in \( X \), \( x = x_0, x_1, \ldots, x_n = y \), with \( n \geq 1 \), such that

\[
d(f(x_i), x_{i+1}) < \varepsilon \quad \text{for } 0 \leq i \leq n - 1.
\]

A point \( x \in X \) is called chain recurrent if for every \( \varepsilon > 0 \) there is an \( \varepsilon \)-chain from \( x \) to itself. The set \( \mathcal{R}(f) \) of chain recurrent points is called the chain recurrent set of \( f \).

**Exercise 1.2.** Let \( f : X \to X \) be a homeomorphism of a compact metric space.

1. The set \( \mathcal{R}(f) \) is closed (hence compact) and invariant under \( f \).
2. If \( x_0, x_1, \ldots, x_n \) is an \( \varepsilon \)-chain from \( x \) to \( y \) and \( y_0, y_1, \ldots, y_m \) is an \( \varepsilon \)-chain from \( y \) to \( z \), then \( x_0, x_1, \ldots, x_n = y_0, y_1, \ldots, y_m \) is an \( \varepsilon \)-chain from \( x \) to \( z \).

3. If for every \( \varepsilon > 0 \) there is an \( \varepsilon \)-chain from \( x \) to \( y \) for \( f \) then for every \( \varepsilon > 0 \) there is a \( \varepsilon \)-chain from \( y \) to \( x \) for \( f^{-1} \).

Recall that a point \( x \) is called recurrent for \( f : X \to X \) if \( x \) is a limit point of the sequence \( x, f(x), f^2(x), \ldots f^n(x), \ldots \). Clearly any recurrent point is also chain recurrent. The converse is not true.

Recall that if \( \mu \) is a finite Borel measure on \( X \) and \( f : X \to X \) is a, not necessarily invertible, function then we say \( \mu \) is \( f \)-invariant provided \( \mu(E) = \mu(f^{-1}(E)) \) for every measurable subset \( E \subset X \).

If there is a finite \( f \)-invariant measure on \( X \) then almost every point of \( X \) (in the measure sense) is recurrent.

**Theorem 1.3** (Poincaré Recurrence Theorem). Suppose \( \mu \) is a finite Borel measure on \( X \) and \( f : X \to X \) is a measure preserving transformation. If \( E \subset X \) is measurable and \( \mathcal{N} \) is the subset of \( E \) given by

\[
\mathcal{N} = \{ x \in E \mid f^k(x) \notin E \text{ for at most finitely many } k \geq 1 \},
\]

then \( \mathcal{N} \) is measurable and \( \mu(\mathcal{N}) = 0 \).

**Proof.** Define

\[
E_N = \bigcup_{n=0}^{\infty} f^{-n}(E) \text{ and } F = \bigcap_{n=0}^{\infty} E_n.
\]

Then \( F \) is the set of points whose forward orbit hits \( E \) infinitely often so \( \mathcal{N} = E \setminus F \) and \( \mathcal{N} \) is measurable. Since \( E_{n+1} = f^{-1}(E_n) \) we have \( \mu(E_{n+1}) = \mu(E_n) \) for all \( n \geq 0 \). Since \( E_0 \supset E_1 \supset E_2 \ldots \) we have

\[
\mu(F) = \mu\left(\bigcap_{n=0}^{\infty} E_n\right) = \lim_{n \to \infty} \mu(E_n) = \mu(E_0).
\]

Hence \( \mu(E_0 \setminus F) = \mu(E_0) - \mu(F) = 0 \). Since \( E \setminus F \subset E_0 \setminus F \) we conclude \( \mu(\mathcal{N}) = \mu(E \setminus F) = 0 \).

**Corollary 1.4.** Suppose \( \mu \) is a probability measure on \( X \). If \( \mu \) is \( f \)-invariant then the set \( \mathcal{N} \) of points which are not recurrent has measure \( \mu(\mathcal{N}) = 0 \).
Proof. Let \( \mathcal{N}_n \) denote the set of points \( x \in X \) such that \( d(x, f^k(x)) > 1/n \) for all \( k > 0 \). We wish first to show \( \mu(\mathcal{N}_n) = 0 \) for all \( n > 0 \).

To do this suppose \( B \) is an open ball in the metric space \( X \) of radius \( 1/2n \) so the distance between any two points of \( B \) is less than \( 1/n \). We conclude from Theorem (1.3) that \( \mu(B \cap \mathcal{N}_n) = 0 \). But since \( X \) is compact it can be covered by finitely many balls \( B \) of radius \( 1/2n \) so we conclude \( \mu(\mathcal{N}_n) = 0 \).

Since \( \mathcal{N}_n = \bigcup_{n=1}^{\infty} \mathcal{N}_n \) we conclude \( \mu(\mathcal{N}) = 0 \). \qed

We have the following immediate corollary.

**Corollary 1.5.** Suppose \( f : X \to X \) preserves a finite Borel measure \( \mu \) and \( \mu(U) > 0 \) for every non-empty open set \( U \). Then there are recurrent points in every such \( U \). I.e. the recurrent points are dense in \( X \).

It is easy to see that in this case the chain recurrent set \( \mathcal{R}(f) \) is all of \( X \). Also, as we now show, in this circumstance if \( X \) is connected, then for any points \( x, y \in X \) there is an \( \varepsilon \)-chain from \( x \) to \( y \).

**Proposition 1.6.** Suppose \( \mu \) is an \( f \)-invariant measure on \( X \) satisfying \( \mu(X) = 1 \) and \( \mu(U) > 0 \) for every non-empty open set \( U \subset X \) and suppose that \( X \) is connected. Then for any \( x, y \in X \) and any \( \varepsilon > 0 \), there is an \( \varepsilon \)-chain from \( x \) to \( y \).

**Proof.** Fix a value of \( \varepsilon > 0 \). We construct an equivalence relation \( \sim_\varepsilon \) on the space \( X \) as follows. Let \( x \sim_\varepsilon y \) provided there is an \( \varepsilon \)-chain from \( x \) to \( y \) and one from \( y \) to \( x \). This clearly defines a symmetric and transitive relation. It is reflexive as well, however. To see this, let \( U \) be a neighborhood of \( x \) such that \( U \) and \( f(U) \) have diameter less than \( \varepsilon \). Clearly if \( f^n(U) \cap U \neq \emptyset \) for some \( n > 0 \) then there is an \( \varepsilon \)-chain from \( x \) to \( x \). In fact, if \( x_0, f^n(x_0) \in U \), we can define \( x_1 = x, x_i = f^i(x_0), 1 < i < n \) and the only “jumps” needed are from \( f(x) \) to \( x_2 = f(x_0) \) and from \( f(x_{n-1}) = f^n(x_0) \) to \( x_n = x \).

But for any open \( U \) it must be the case that \( f^n(U) \cap U \neq \emptyset \) for some \( n > 0 \) since otherwise the sets \( f^i(U) \) are pairwise disjoint and all have the same positive measure which would mean the measure of \( X \) is infinite. Hence the relation \( \sim_\varepsilon \) is reflexive and thus an equivalence relation.
From the definition of $\varepsilon$-chain it is immediate that the equivalence classes are open sets in $X$. Since the equivalence classes form a partition of $X$ into pairwise disjoint open sets and $X$ is connected, there must be a single equivalence class. Thus for any $x, y \in X$ there is an $\varepsilon$-chain for $X$ from $x$ to $y$. Since $\varepsilon$ was arbitrary the result follows. 

2 The “Fundamental Theorem of Dynamical Systems”

In this section we briefly review the elementary theory of attractor-repeller pairs and complete Lyapunov functions developed by Charles Conley in [C]. We give Conley’s proof of the the existence of complete Lyapunov functions, (which is sometimes called the “Fundamental Theorem of Dynamical Systems”)

If $A \subset X$ is a compact subset and there is an open neighborhood $U$ of $A$ such that $f(\text{cl}(U)) \subset U$ and $\cap_{n \geq 0} f^n(\text{cl}(U)) = A$, then $A$ is called an attractor and $U$ is an isolating neighborhood. It is easy to see that if $V = X \setminus \text{cl}(U)$ and $A^* = \cap_{n \geq 0} f^{-n}(\text{cl}(V))$, then $A^*$ is an attractor for $f^{-1}$ with isolating neighborhood $V$. The set $A^*$ is called the repeller dual to $A$. It is clear that $A^*$ is independent of the choice of isolating neighborhood $U$ for $A$. Obviously $f(A) = A$ and $f(A^*) = A^*$.

Lemma 2.1. The set of attractors for $f$ is countable.

Proof. Choose a countable basis $\mathcal{B} = \{V_n\}_{n=1}^\infty$ for the topology of $X$. If $A$ is an attractor with open isolating neighborhood $U$, then $U$ is a union of sets in $\mathcal{B}$. Hence, since $A$ is compact, there are $V_{i_1}, \ldots, V_{i_k}$ such that $A \subset V_{i_1} \cup \cdots \cup V_{i_k} \subset U$. Clearly $A = \cap_{n \geq 0} f^n(U) = \cap_{n \geq 0} f^n(V_{i_1} \cup \cdots \cup V_{i_k})$. Consequently there are at most as many attractors as finite subsets of $\mathcal{B}$, i.e., the set of attractors is countable. 

Lemma 2.2. If $\{A_n\}_{n=1}^\infty$ are the attractors of $f$ and $\{A_n^*\}$ their dual repellers, then the chain recurrent set $\mathcal{R}(f) = \cap_{n=1}^\infty (A_n \cup A_n^*)$.

Proof. We first show $\mathcal{R}(f) \subset \cap (A_n \cup A_n^*)$. This is equivalent to showing that if $x \notin A \cup A^*$ for some attractor $A$, then $x \notin \mathcal{R}(f)$. If $U$ is an open isolating neighborhood of $A$ and $x \notin A \cup A^*$, then $x \in f^{-n}(U)$ for some $n$. Let $m$ be the smallest such $n$. Replacing $U$ with $f^{-m}(U)$ we can assume $x \in U \setminus f(U)$. 

4
Now choose \( \varepsilon_0 > 0 \) so that any \( \varepsilon_0 \)-chain \( x = x_1, x_2, x_3 \) must have \( x_3 \in f^2(U) \). If \( \varepsilon_1 = d(X \setminus f(U), \text{cl}(f^2(U))) \) and \( \varepsilon = \frac{1}{2} \min \{ \varepsilon_0, \varepsilon_1 \} \), then no \( \varepsilon \)-chain can start and end at \( x \), since no \( \varepsilon \)-chain from a point of \( f^2(U) \) can reach a point of \( X \setminus f(U) \). Thus \( x \notin \mathcal{R}(f) \). We have shown \( \mathcal{R}(f) \subset \cap (A_n \cup A_n^*) \).

We next show the reverse inclusion. Suppose \( x \in \bigcap_{n=1}^{\infty} (A_n \cup A_n^*) \). If \( x \) is not in \( \mathcal{R}(f) \), there is an \( \varepsilon_0 > 0 \) such that no \( \varepsilon_0 \)-chain from \( x \) to itself exists. Let \( \Omega(x, \varepsilon) \) denote the set of \( y \in X \) such that there is an \( \varepsilon \)-chain from \( x \) to \( y \).

By definition, the set \( V = \Omega(x, \varepsilon_0) \) is open. Moreover, \( f(\text{cl}(V)) \subset V \), because if \( z \in \text{cl}(V) \), there is \( z_0 \in V \) such that \( d(f(z), f(z_0)) < \varepsilon_0 \) and consequently an \( \varepsilon_0 \)-chain from \( x \) to \( z_0 \), gives an \( \varepsilon_0 \)-chain \( x = x_0, x_1, \ldots, x_k, z_0, f(z) \) from \( x \) to \( f(z) \). Hence \( A = \bigcap_{n \geq 0} f^n(\text{cl}(V)) \) is an attractor with isolating neighborhood \( V \). By assumption either \( x \in A \) or \( x \in A^* \). Since there is no \( \varepsilon_0 \)-chain from \( x \) to \( x \), \( x \notin A \). On the other hand, if \( \omega(x) \) denotes the limit points of \( \{ f^n(x) \mid n \geq 0 \} \), then clearly \( \omega(x) \subset V \), but this is not possible if \( x \in A^* \) since \( A^* \) is closed and \( x \in A^* \) would imply \( \omega(x) \subset A^* \). Thus we have contradicted the assumption that \( x \notin \mathcal{R} \).

**Exercise 2.3.** Let \( f = \text{id} : X \to X \) be the identity homeomorphism of a compact metric space. Find all attractors of \( f \) and their dual repellers.

If we define a relation \( \sim \) on \( \mathcal{R} \) by \( x \sim y \) if for every \( \varepsilon > 0 \) there is an \( \varepsilon \)-chain from \( x \) to \( y \) and another from \( y \) to \( x \), then it is clear that \( \sim \) is an equivalence relation.

**Definition 2.4.** The equivalence classes in \( \mathcal{R}(f) \) for the equivalence relation \( \sim \) above are called the chain transitive components of \( \mathcal{R}(f) \).

**Proposition 2.5.** If \( x, y \in \mathcal{R}(f) \), then \( x \) and \( y \) are in the same chain transitive component if and only if there is no attractor \( A \) with \( x \in A \), \( y \in A^* \) or with \( y \in A \), \( x \in A^* \).

**Proof.** Suppose first that \( x \) and \( y \) are in the same chain transitive component, i.e., \( x \sim y \), and \( x \in A \). If \( U \) is an open isolating neighborhood for \( A \), let \( \varepsilon = \text{dist}(X \setminus U, \text{cl}(f(U))) \). There can be no \( \varepsilon/2 \)-chain from a point in \( f(U) \) to a point in \( X \setminus U \), hence none from a point in \( A \) to a point in \( A^* \). By Lemma 2.2 we know \( y \in A \cup A^* \), but \( x \sim y \) implies \( y \notin A^* \), so \( y \in A \). This proves one direction of our result.

To show the converse, suppose that for every attractor \( A \), \( x \in A \) if and only if \( y \in A \) (and hence \( x \in A^* \) if and only if \( y \in A^* \)). Given \( \varepsilon > 0 \) let \( V = \Omega(x, \varepsilon) = \{ z \in X \mid \text{for which there is an } \varepsilon \text{-chain from } x \text{ to } z \} \).
to z. Since x is chain recurrent \( x \in V \). Also as in the proof of Lemma 2.2, V is an isolating neighborhood for an attractor \( A_0 \). Since \( x \in A_0 \cup A_0^* \) and \( x \in V \) we have \( x \in A_0 \). Thus \( y \in A_0 \subset V \) so there is an \( \varepsilon \)-chain from \( x \) to \( y \). A similar argument shows there is an \( \varepsilon \)-chain from \( y \) to \( x \) so \( x \sim y \).

We are now prepared to present Conley’s proof of the existence of a complete Lyapunov function.

**Definition 2.6.** A complete Lyapunov function for \( f : X \to X \) is a continuous function \( g : X \to \mathbb{R} \) satisfying:

1. If \( x \notin \mathcal{R}(f) \), then \( g(f(x)) < g(x) \)
2. If \( x, y \in \mathcal{R}(f) \), then \( g(x) = g(y) \) if and only if \( x \sim y \) (i.e., \( x \) and \( y \) are in the same chain transitive component).
3. \( g(\mathcal{R}(f)) \) is a compact nowhere dense subset of \( \mathbb{R} \).

By analogy with the smooth setting, elements of \( g(\mathcal{R}(f)) \) are called critical values of \( g \).

**Lemma 2.7.** There is a continuous function \( g : X \to [0, 1] \) such that \( g^{-1}(0) = A, g^{-1}(1) = A^* \) and \( g \) is strictly decreasing on orbits of points in \( X \setminus (A \cup A^*) \).

*Proof.* Define \( g_0 : X \to [0, 1] \) by

\[
g_0(x) = \frac{d(x, A)}{d(x, A) + d(x, A^*)}.
\]

Let \( g_1(x) = \sup\{g_0(f^n(x)) \mid n \geq 0\} \). Then \( g_1 : X \to [0, 1] \) and \( g_1(f(x)) \leq g_1(x) \) for all \( x \). We must show \( g_1 \) is continuous. If \( \lim x_i = x \in A \), then clearly \( \lim g_1(x_i) = 0 \) so \( g_1 \) is continuous at points of \( A \) and the same argument shows it is continuous at points of \( A^* \). If \( U \) is an open isolating neighborhood as above, let \( N = \text{cl}(U) \setminus f(U) \). Let \( x \in N \) and \( r = \inf\{g_0(x) \mid x \in N\} \). Since \( f^n(U) \subset f^n(\text{cl}(U)) \) and \( \bigcap_{n \geq 0} f^n(\text{cl}(U)) = A \), it follows that there is \( n_0 > 0 \) such that \( g_0(f^n(U)) \subset [0, r/2] \) whenever \( n > n_0 \). Hence for \( x \in N \),

\[
g_1(x) = \max\{g_0(f^n(x)) \mid 0 \leq n \leq n_0\}
\]

6
so \( g_1 \) is continuous on \( N \). Since \( \bigcup_{n=-\infty}^{\infty} f^n(N) = X \setminus (A \cup A^*) \), \( g_1 \) is continuous. Finally, letting

\[
g(x) = \sum_{n=0}^{\infty} \frac{g_1(f^n(x))}{2^n+1}
\]

we obtain a continuous function \( g : X \to [0,1] \) such that \( g^{-1}(0) = A \), \( g^{-1}(1) = A^* \). Also

\[
g(f(x)) - g(x) = \sum_{n=0}^{\infty} \frac{g_1(f^{n+1}(x)) - g_1(f^n(x))}{2^{n+1}}
\]

which is negative if \( x \notin A \cup A^* \), since \( g_1(f(y)) \leq g_1(y) \) for all \( y \) and \( g_1 \) is not constant on the orbit of \( x \). \( \Box \)

The following theorem is essentially a result of \([C]\). We have changed the setting from flows to homeomorphisms.

**Theorem 2.8** (Fundamental Theorem of Dynamical Systems). If \( f : X \to X \) is a homeomorphism of a compact metric space, then there is a complete Lyapunov function \( g : X \to \mathbb{R} \) for \( f \).

**Proof.** By Lemma (2.4) there are only countably many attractors \( \{A_n\} \) for \( f \). By Lemma (2.7) we can find \( g_n : X \to \mathbb{R} \) with \( g_n^{-1}(0) = A_n \), \( g_n^{-1}(1) = A^*_n \) and \( g_n \) strictly decreasing on \( X \setminus (A_n \cup A^*_n) \). Define \( g : X \to \mathbb{R} \) by

\[
g(x) = \sum_{n=1}^{\infty} \frac{2g_n(x)}{3^n}.
\]

The series converges uniformly so \( g(x) \) is continuous. Clearly if \( x \notin \mathcal{R}(f) \), then there is an \( A_i \) with \( x \notin (A_i \cup A^*_i) \) so \( g(f(x)) < g(x) \).

Also, if \( x \in \mathcal{R}(f) \), then \( x \in (A_n \cup A^*_n) \) for every \( n \), so \( g_n(x) = 0 \) or 1 for all \( n \). It follows that the ternary expansion of \( g(x) \) can be written with only the digits 0 and 2, and hence \( g(x) \in C \), the Cantor middle third set. Thus \( g(\mathcal{R}(f)) \subset C \) so \( g(\mathcal{R}(f)) \) is compact and nowhere dense. This proves (3) of the definition.

Finally, if \( x, y \in \mathcal{R}(f) \) then \( g(x) = g(y) \) if and only if \( g_n(x) = g_n(y) \) for all \( n \). This is true since \( 2g_n(x) \) is the \( n \)th digit of the ternary expansion of \( g(x) \) so \( g(x) = g(y) \) implies \( g_n(x) = g_n(y) \) for all \( n \). But \( g_n(x) = g_n(y) \) for all \( n \) if and only if there is no \( n \) with \( x \in A_n \), \( y \in A^*_n \) or with \( x \in A^*_n \), \( y \in A_n \). Thus by Proposition (2.6), \( g(x) = g(y) \) if and only if \( x \) and \( y \) are in the same chain transitive component. \( \Box \)
References

[C] C. Conley, *Isolated Invariant Sets and the Morse index*, C.B.M.S. Regional Conference Series in Math \textbf{38}, Amer. Math. Soc., Providence, RI, 1978.