Symbolic dynamics emerges from a dynamical system $((X,T)$ if we code the T-orbits of the points $x \in X$. To this end, we let $\mathcal{J} = \{J_a\}_{a\in\mathcal{A}}$ (for a finite or countable alphabet $\mathcal{A}$) be a partition of X, and to each $x \in X$ we assign an **itinerary** $i(x) \in \mathcal{A}^{\mathbb{N}_0}$:

$$i_n(x) = a \quad \text{ if } T^n(x) \in J_a.$$

If T is invertible, then we can extend sequences to $\mathcal{A}^{\mathbb{Z}}$. It is clear that $i \circ T(x) = \sigma \circ i(x)$. Therefore, if we set $\Sigma = i(X)$, then $\sigma(\Sigma) \subset \Sigma$ and if $T : X \to X$ is surjective, then $\sigma(\Sigma) = \Sigma$. But (Σ, σ) is in general not a subshift, because Σ is not closed.

Example 0.1. *Let $X = [0,1]$ and $T(x) = Q_4(x) = 4x(1-x)$. Let $J_0 = [0, \frac12]$ and $J_1 = (\frac12, 1]$. Then $i(X)$ is not closed, because there is no $x \in [0,1]$ such that $i(x) = 1100000\dots$, while $1100000\dots = \lim_{x\searrow\frac12} i(x)$. Naturally, redefining the partition to $J_0 = [0,\frac12)$ and $J_1 = [\frac12, 1]$ doesn't help, because then there is no $x \in [0,1]$ such that $i(x) = 0100000\dots$, while $0100000\dots = \lim_{x\nearrow\frac12} i(x)$.*

Other "solutions" that one sees in the literature are:

- *Assigning a different symbol (often $*$ or C) to $\frac12$. That is, using the partition $J_0 = [0, \frac12)$, $J_* = \{\frac12\}$ and $J_1 = (\frac12, 1]$. This resolves the "ambiguity" about which symbol to give to $\frac12$, but it doesn't make the shift space closed.*

- *Assigning the two symbols to $\frac12$, so $J_0 = [0, \frac12]$ and $J_1 = [\frac12, 1]$ are noo longer a partition, but have $\frac12$ in common. Therefore $\frac12$ will have two itineraries, and so will every point in the backward orbit of $\frac12$. With all these extra itineraries, $i(X)$ becomes closed. But this doesn't work in all cases, see Exercise 0.2.*

- *Taking a quotient space $i(X)/\sim$ where in this case $x \sim y$ if there is $n \in \mathbb{N}_0$ such that*

$$x_0 \dots x_{n-1} = y_0 \dots y_{n-1} \text{ and } \begin{cases} x_n x_{n+1} x_{n+2} x_{n+3} x_{n+4} \dots = 11000\dots, \\ y_n y_{n+1} y_{n+2} y_{n+3} y_{n+4} \dots = 01000\dots \end{cases}$$

  or vice versa. This quotient space adapts the quotient topology (so $/\sim$ is not a Cantor set anymore), and it turns the coding map into a genuine homeomorphism.

Exercise 0.2. *Let $a = 3.83187405528332\dots$ and $T(x) = Q_a(x) = ax(1-x)$. For this parameter, $T^3(\frac12) = \frac12$. Let $\mathcal{J}' = \{[0, \frac12], (\frac12, 1]\}$ and $\mathcal{J} = \{[0, \frac12], [\frac12, 1]\}$, so $\frac12$ get two symbols. Let $\Sigma' = i(X)$ w.r.t. $\mathcal{J}'$ and $\Sigma = i(X)$ w.r.t. $\mathcal{J}$. Show that $\overline{\Sigma'} \neq \Sigma$.*

From now on, assume that X is compact metric space without isolated points. We will now discuss the properties of the coding map i itself. First of all, for i to be continuous it is crucial that $T|_{J_a}$ is continuous on each element $J_a \in \mathcal{J}$. But this is not enough: if x is a common boundary of two element of $\mathcal{J}$ then (no matter how you assign the symbol to x in Example 0.1), for each neighborhood $U \ni x$, $\operatorname{diam}(i(U)) = 1$, so continuity fails at x. It is only by using quotient spaces of $i(X)$ (so changing the topology of $i(X)$) that can make i continuous. Normally, we choose to live with the discontinuity, because it affects only few points:

Lemma 0.3. *Let $\partial\mathcal{J}$ denote the collection of common boundary points of different elements in $\mathcal{J}$. If $orb(x) \cap \partial J = \varnothing$, then the coding map $i : X \to \mathcal{A}^{\mathbb{N}_0}$ or $\mathcal{A}^{\mathbb{Z}}$ is continuous at x.*

Proof. We carry out the proof for invertible maps. Let $\varepsilon > 0$ be arbitrary and fix $N \in \mathbb{N}$ such that $2^{-N} < \varepsilon$. For each $n \in \mathbb{Z}$ with $|n| \leq N$, let $U_n \ni T^n(x)$ be such a small neighborhood that it is contained in a single partition element $J_{i_n(x)}$. Since $\mathrm{orb}(x) \cap \partial J = \varnothing$, this is possible. Then $U := \cap_{|n|\leq N} T^n(U_n)$ is an open neighborhood of x and $i_n(y) = i_n(x)$ for all $|n| \leq N$ and $y \in U$. Therefore $\mathrm{diam}(i(U)) \leq 2^{-N} < \varepsilon$, and continuity at x follows. □

Definition 0.4. *A system (X,T) is called* **expansive** *if there exists $\delta > 0$ such that for all distinct $x, y \in X$, there is $n \geq 0$ (or $n \in \mathbb{Z}$ if $T$is invertible) such that $d(T^n x, T^n y) > \delta$. We call δ the* **expansivity constant**.

Lemma 0.5. *Suppose that T is a continuous expansive dynamical system and injective on each $J_a \in \mathcal{J}$. If the expansivity constant is larger than $\sup_{a\in\mathcal{A}} \mathrm{diam}(J_a)$, then the coding map $i : X \to \mathcal{A}^{\mathbb{N}_0}$ or $\mathcal{A}^{\mathbb{Z}}$ is injective.*

Proof. Suppose that there are $x \neq y \in X$ such that $i(x) = i(y)$. Since $T|_{J_a}$ is injective for each $a \in \mathcal{A}$, $T^n(x) \geq T^n(y)$ for all $n \geq 0$. Let $\delta > 0$ be an expansitivity constant of T. Thus, there is $n \in \mathbb{Z}$ such that $d(T^n(x), T^n(y)) > \delta$, so, by assumption, they cannot lie in the same element of $\mathcal{J}$. Hence x and y cannot have the same itinerary after all. □

To obtain injectivity of the coding map, it often suffices that T is expanding on each partition element J_a. Expanding should not be confused with expansive.

Definition 0.6. *Given a metric space X and $Y \subset X$, a map $T : Y \to T(Y)$ is* **expanding** *if $d(T(x), T(y)) \geq \rho d(x, y)$ for all $x, y \in Y$, and* **uniformly expanding** *there is $\rho > 1$ such that $d(T(x), T(y)) \geq \rho d(x, y)$ for all $x, y \in Y$.*

Example 0.7. *Let $T : \mathbb{S}^1 \to \mathbb{S}^1$, $x \mapsto 2x \bmod 1$, be the doubling map, and $J_0 = (\frac{1}{4}, \frac{3}{4})$ and $J_1 = \mathbb{S}^1 \setminus J_0$. Clearly $T'(x) = 2$ for all $x \in \mathbb{S}^1$, but T is not expanding on the whole of $\mathbb{S}^1$, because for instance $d(T(\frac{1}{4}), T(\frac{3}{4})) = 0 < \frac{1}{2} = d(\frac{1}{4}, \frac{3}{4})$. More importantly, T is not expanding on the either J_a; for example $d(T(\frac{1}{4}+\varepsilon), T(\frac{3}{4}-\varepsilon)) = 4\varepsilon < \frac{1}{2} - 2\varepsilon = d(\frac{1}{4}+\varepsilon, \frac{3}{4}-\varepsilon)$ for each $\varepsilon \in (0, \frac{1}{12})$. The corresponding coding map is* **not** *injective. The way to see this by noting that the involution $S(x) = 1 - x$ commutes with T and also preserves each J_a. It follows that $i(x) = i(S(x))$ for all $x \in \mathbb{S}^1$, and only $x = 0$ and $x = \frac{1}{2}$ have unique itineraries.*