Ruelle’s Perron-Frobenius Theorem: Recall the statement: Let $\Sigma^+_A$ be a topologically mixing subshift of finite type and $\varphi: \Sigma^+_A \to \mathbb{R}$ Hölder. Then there is $\lambda > 0$, $\nu \in M(\Sigma^+_A)$ a shift-invariant measure, and $h \in \Sigma^+_A$ positive such that

$$L^*_\varphi(\nu) = \lambda \nu$$
$$L^*_\varphi(h) = \lambda h$$
$$\nu(h) = 1$$

and for every $g \in \Sigma^+_A$,

$$\left\| \frac{1}{\lambda^m} L^m_\varphi(g) - \nu(g)h \right\| \to 0, \text{ as } m \to \infty$$

where $\| \cdot \|$ denotes the supremum norm on continuous functions.

Some new notation: $C_m(\omega) = \{ \omega' \in \Sigma^+_A : \omega_i = \omega'_i, i = 0, \ldots, m \}$. This is $C[0; \omega_0, \ldots, \omega_m]$ in our old notation. Recall that we were working with the function space $\Lambda = \{ u \in C(\Sigma^+_A) | u \geq 0, \nu(u) = 1, \text{ and if } \omega_i = \omega'_i \text{ for } i \in [0, m] \text{ then } u(\omega) \leq B_m u(\omega') \}$ where $B_m = e^{\sum_{k=0}^{\infty} 2^b \alpha^k}$, where $\text{var}_k(\varphi) \leq 2^b \alpha^k$. Let $M$ be the smallest positive integer such that $A^M$ has all positive entries. Recall $Q = \frac{1}{\lambda} L_\varphi$. We showed last time that $Q(\Lambda) \subset \Lambda$. We also showed that if $u \in \Lambda$ then $\inf Q^M(u) > \frac{1}{K}$, where $K = \lambda^M e^{M||\varphi||} B_0$. By the Schauder-Tychonoff theorem, there was an $h \in \Lambda$ such that $Q(h) = h > \frac{1}{K}$ since also $Q^M(h) = h$. Last time we also proved the inequality for $u \in \Lambda, \omega' \in C_m(\omega)$,

$$[L_\varphi u](\omega) \leq e^{2b \alpha^m+1} B_{m+1} [L_\varphi u](\omega')$$

It only remains to prove the convergence statement $\left\| \frac{1}{\lambda^m} L^m_\varphi(g) - \nu(g)h \right\| \to 0$ for $g \in C(\Sigma^+_A)$. We will prove this statement for $g = u \in \Lambda$. The general case can be found in Bowen’s book; it proceeds by an approximation argument (all functions in $C(\Sigma^+_A)$ can be suitably approximated by functions in $\Lambda$).

The main step is to show that there is some $\eta \in (0, 1)$ such that for every $u \in \Lambda$,

$$Q^M(u) = \eta h + (1 - \eta) u',$$

for some $u' \in \Lambda$. This lemma can be iterated:

$$Q^{2M}(u) = \eta h + (1 - \eta) Q^M(u') = \eta h + (1 - \eta) \eta h + (1 - \eta)^2 u''$$

and in general

$$Q^{kM}(u) = C_k h + (1 - \eta)^k u_k$$

where $u_k \in \Lambda$ and $C_k = (1 - (1 - \eta)^k)$. This proves that

$$\|Q^{kM}(u) - h\| \leq (1 - \eta)^k (\|h\| + \|u_k\|) \leq (1 - \eta)^k (\|h\| + \|u_k\|) \leq (1 - \eta)^k (\|h\| + K)$$
Recall from last time we also proved an upper bound \( \|u\| \leq K \) for \( u \in \Lambda \). Thus \( \|Q^k(u) - h\| \to 0 \) exponentially fast, which proves the claim since \( \nu(u) = 1 \). (Technically, we have only shown this along a subsequence, but it’s easy to upgrade this to convergence along the full sequence using the estimates coming from \( u \in \Lambda \)).

We now prove the lemma: Fix \( \eta \in (0, 1) \), whose exact value we will determine later. Given \( u \in \Lambda \), set \( u' = \frac{Q^M(u) - \eta h}{1 - \eta} \).

We need to show that for \( \eta \) small enough, \( u' \in \Lambda \). It’s easy to check that \( \nu(u') = 1 \). If \( \eta < \frac{1}{\|h\|K} \), then \( u' \geq 0 \), using the lower bound \( Q^M(u) \geq \frac{1}{K} \). It suffices to show that for \( v = (1 - \eta)u' \), for every \( m \geq 0 \), and for every \( \omega, \omega' \in C_m(\omega) \), \( v(\omega) \leq B_m v(\omega') \)

We can rewrite this as \( \eta (B_m h(\omega) - h(\omega')) \leq B_m Q^M u(\omega) - Q^M u(\omega') \) \( (*) \)

Let \( u_1 = Q^{M-1}(u) \). Using the inequality we proved to show that \( Q(\Lambda) \subset \Lambda \),

\[
L_\varphi u_1(\omega) \leq e^{b\alpha + \eta} B_{m+1} L_\varphi u_1(\omega')
\]

\[
\Rightarrow Q^M u(\omega) \leq e^{b\alpha + \eta} B_{m+1} Q^M u(\omega')
\]

Now \( h \in \Lambda \) implies that \( h(\omega) \leq B_m h(\omega') \), and so inequality \( (*) \) can be proved by finding an \( \eta \) such that

\[
\eta \leq \frac{B_m - e^{b\alpha + \eta} B_{m+1}}{K \|h\| \left( B_m - B_m^{-1} \right)}.
\]

Set \( c_m := e^{b\alpha m} \). Then we want to choose \( \eta \) such that:

\[
\eta \leq \frac{B_m - c_m+1 B_{m+1}}{K \|h\| \left( B_m - B_m^{-1} \right)}.
\]

Now:

\[
\frac{B_m - c_m+1 B_{m+1}}{K \|h\| \left( B_m - B_m^{-1} \right)} \leq \frac{B_m - c_m B_{m+1}}{K \|h\| \left( B_m - B_m^{-1} \right)}
\]

\[
= \frac{B_m \left( 1 - c_m^{-1} \right)}{K \|h\| \left( B_m - B_m^{-1} \right)}
\]

\[
= \frac{\left( 1 - c_m^{-1} \right)}{K \|h\| \left( 1 - B_m^{-2} \right)}
\]

But as \( m \to \infty \),

\[
\frac{\left( 1 - c_m^{-1} \right)}{K \left( 1 - B_m^{-2} \right) \|h\|} \to \frac{1}{K \|h\|}.
\]

Hence for \( m \) large enough, choosing \( \eta \leq \frac{1}{2 K \|h\|} \) is sufficient, and multiplying by a sufficiently small constant gives the inequality for all \( m \).

An alternate proof of Ruelle’s Perron-Frobenius can be found in an article of Pollicott titled A complex Ruelle-Perron-Frobenius theorem and two counterexamples.
**Gibbs measures:** We keep the same assumptions as those for Ruelle’s Perron-Frobenius theorem. Last time we showed that if we set \( d\mu = hd\nu \), \( \mu_\varphi \) is a \( \sigma \)-invariant measure. We now want to show this measure has some nice properties. First we have a detour on ergodicity and mixing.

Let \( T : (X, \mu) \to (X, \mu) \) be a measure preserving transformation. \( T \) is ergodic if for every measurable \( B \),

\[
T^{-1}(B) = B \Rightarrow \mu(B) = 0 \text{ or } 1.
\]

This means that the only \( T \)-invariant sets are trivial, i.e., have full measure or zero measure. \( T \) is mixing if for every \( B \) and \( C \) measurable,

\[
\lim_{m \to \infty} \mu(T^{-m}(B) \cap C) = \mu(B)\mu(C).
\]

It’s easy to see that mixing implies ergodic: If \( B \) is \( T \)-invariant then taking \( B = C \) in the above formula gives \( \mu(B) = \mu(B)^2 \), so that \( \mu(B) = 0 \) or 1. Mixing can be reinterpreted by defining

\[
\mu(T^{-m}(B) \mid C) := \frac{\mu(T^{-m}(B) \cap C)}{\mu(C)}
\]

to be the conditional measure of \( T^{-m}(B) \) in \( C \). This is the percentage of \( C \) that \( T^{-m}(B) \) occupies. Mixing then implies that as \( m \to \infty \), \( B \) and \( C \) become asymptotically independent; the percentage of \( T^{-m}(B) \) in \( C \) is the same as the percentage of \( B \) in \( X \).

**Decay of Correlations:** It’s an easy exercise to show that \( T \) is ergodic if and only if for every \( f, g \in L^2(X, \mu) \), \( f \circ T = f \) implies that \( f \) is constant (where both of these statements should be taken \( \mu \)-a.e.).

A similar alternative definition of mixing is the following: \( T \) is mixing if and only if for every \( f, g \in L^2(X, \mu) \), as \( m \to \infty \),

\[
\int (f \circ T^m)(x)g(x)d\mu(x) \to \left( \int f(x)d\mu(x) \right) \left( \int g(x)d\mu(x) \right).
\]

This is sometimes called a ”decay of correlations” condition, in analogy with probability theory, thinking of \( \{f \circ T^m\}_{m \in \mathbb{N}} \) as a sequence of random variables and \( \int (f \circ T^m)(x)g(x)d\mu(x) \) as the correlations between these random variables and \( g \). The mixing condition then implies that the random variables \( \{f \circ T^m\}_{m \in \mathbb{N}} \) become asymptotically independent of any given function \( g \).

Back to subshifts of finite type: We claim that \( \sigma \) acting on \( (\Sigma_1^+, \mu_\varphi) \) is a mixing transformation (hence is also ergodic). It suffices to show that decay of correlations holds for a dense set in \( L^2(\Sigma_1^+, \mu_\varphi) \); e.g. that for every \( u, v \in C(\Sigma_1^+) \), we have

\[
\lim_{m \to \infty} \int (u \circ \sigma^m) v d\mu \to \left( \int ud\mu_\varphi \right) \left( \int vd\mu_\varphi \right).
\]

Let \( u, v \) be continuous. Then

\[
\mu(u \cdot (v \circ \sigma^m)) = \nu(h \cdot u \cdot (v \circ \sigma^m))
= \frac{1}{\lambda^m} L_\varphi^m \nu(h \cdot u \cdot (v \circ \sigma^m))
= \frac{1}{\lambda^m} \nu(L_\varphi^m (h \cdot u \cdot (v \circ \sigma^m)))
= \frac{1}{\lambda^m} \nu(L_\varphi^m (h \cdot u) \cdot v)
= \nu\left( \frac{1}{\lambda^m} L_\varphi^m (h \cdot u) \cdot v \right).
\]
and therefore

\[
\mu(u \cdot (v \circ \sigma^m)) - \mu(u) \mu(v) = \nu \left( \frac{1}{\lambda^m} L^m(\phi) \cdot v \right) - \nu(hu) \nu(hv)
\]

\[
= \nu \left( \frac{1}{\lambda^m} L^m(\phi) \cdot v - \nu(hu) \cdot v \right) - \nu(hu) \cdot v 
\]

which goes to 0 uniformly as \(m \to \infty\) by the Ruelle-Perron-Frobenius theorem.