Definition 26 (Harris recurrence) *A set A is called Harris recurrent if*

$$Q(x, A) = P_x(\eta_A = \infty) = 1, \quad x \in A.$$

A chain $\mathbf{\Phi}$ is called Harris (recurrent) if it is ψ-irreducible and every set $A \in \mathcal{B}^+(\mathcal{X})$ is Harris recurrent.

A standard alternative definition of an Harris recurrence set is that $L(x, A) = 1$ for $x \in A$, however they are equivalent. Nevertheless, using $Q(x, A) = 1$ highlights the strengthening of recurrence to Harris recurrence from an expected infinite number of visits to a set to an almost surely infinite number of visits.

Proposition 17 *Suppose for some one set $A \in \mathcal{B}(\mathcal{X})$ we have $L(x, A) \equiv 1$ for $x \in A$. Then $Q(x, A) = L(x, A)$ for all $x \in \mathcal{X}$ and A is Harris recurrent.*

Proof:

The most difficult proof of this section is the following.

Theorem 11 *(i) Suppose that $D \rightsquigarrow A$ for any D, $A \in \mathcal{B}(\mathcal{X})$. Let μ be any initial distribution of the chain $\mathbf{\Phi}$. Then*

$$\{\mathbf{\Phi} \in D\, i.o.\} \subseteq \{\mathbf{\Phi} \in A\, i.o.\} \quad a.s.\ [P_\mu]$$

and $Q(y, D) \leq Q(y, A)$ for all $y \in \mathcal{X}$.

(ii) If $\mathcal{X} \rightsquigarrow A$, then A is Harris recurrent and $Q(x, A) \equiv 1$ for all $x \in \mathcal{X}$.

Proof: See Meyn & Tweedie, p. 202. The proof involves Martingales and would take too much time to give the necessary background to show/prove the Martingale Convergence Theorem.

This leads us to the following strengthening of Harris recurrence...

Theorem 12 *If $\mathbf{\Phi}$ is Harris recurrent, then $Q(x, B) = 1$ for every $x \in \mathcal{X}$ and every $B \in \mathcal{B}^+(\mathcal{X})$.*

Proof:

Let D be any Harris recurrent set. Let $D^\infty = \{y : L(y, D) = 1\}$. Then $D \subseteq D^\infty$ and D^∞ is absorbing. We call D a maximal absorbing set if $D = D^\infty$.

Definition 27 (Maximal Harris sets) *A set H is called maximal Harris if H is a maximal absorbing set such that $\mathbf{\Phi}$ restricted to H is Harris recurrent.*

Recall the following definitions

Definition 28 (Full and Absorbing Sets) *A set $A \in \mathcal{B}(\mathcal{X})$ is said to be*

(i) full if $\psi(A^c) = 0$.

(ii) absorbing if $P(x, A) = 1$ for $x \in A$.

In order to prove Theorem 13 we need the following three lemmas:

Lemma 5 *Suppose that $\mathbf{\Phi}$ is ψ-irreducible. Then*

(i) every absorbing set is full.

(ii) every full set contains a non-empty absorbing set.

Proof:

Lemma 6 *If $\mathbf{\Phi}$ is a Markov chain and if $A \in \mathcal{B}(\mathcal{X})$ satisfies $L(x, A) \leq \epsilon < 1$ for $x \in A$, then $U(x, A) \leq (1-\epsilon)^{-1}$ for all $x \in \mathcal{X}$.*

Proof:

Lemma 7 *If $\mathbf{\Phi}$ is ψ-irreducible and $A \in \mathcal{B}(\mathcal{X})$ with $\psi(A) = 0$, then A is transient.*

Proof: See Meyn & Tweedie, p. 186.

Theorem 13 *If $\mathbf{\Phi}$ is recurrent, then*

$$\mathcal{X} = H \cup N$$

where H is a non-empty maximal Harris set and N is transient. Furthermore, $\psi(N) = 0$.

Proof:

This theorem states that a recurrent Markov chain and a Harris recurrent Markov chain differ only by the existence of a ψ-null set N on which recurrence does not hold.

Theorem 14 *Suppose that* $\mathbf{\Phi}$ *is* ψ*-irreducible and aperiodic. Then* $\mathbf{\Phi}$ *is Harris if and only if each skeleton is Harris.*

Proof:

The existence of a stationary distribution

Now we investigate the conditions necessary for a Markov chain to possess a stationary distribution. This is quite important for the development of MCMC theory—for we want to be able to sample from the posterior distribution, which is stationary. It is quite obvious that if we construct a Markov chain that does not possess a stationary distribution, we will not be sampling from the posterior. More importantly, we need to also show that the n-step transition kernels converge to this stationary distribution, no matter what the initial distribution, and will do so in the next section. Thus once $P^n(\mu, A) = \pi(A)$ we are guaranteed that $P^{n+1}(\mu, A) = \pi(A)$ as well (once we start sampling from the posterior distribution, we will always be sampling from the posterior).

Definition 29 (Invariant measure) *If a σ-finite measure π on $\mathcal{B}(\mathcal{X})$ has the property*

$$\pi(A) = \int_{\mathcal{X}} \pi(dx)P(x, A), \quad A \in \mathcal{B}(\mathcal{X}),$$

then we call π invariant.

If an invariant measure is finite, then we can renormalize it to be a probability measure. Obviously, this is the situation in which we are most interested.

Definition 30 (Positive and null chains) *If $\mathbf{\Phi}$ is ψ-irreducible and admits an invariant probability measure π, then we call $\mathbf{\Phi}$ a positive chain. If $\mathbf{\Phi}$ does not admit such a measure, we call it null.*

Recall that a process is stationary if for any k, the marginal distribution of $\{\Phi_n, \ldots, \Phi_{n+k}\}$ does not change as n varies. In general, Markov chains are not stationary (e.g. consider the a chain with initial distribution δ_x). However, with an appropriate choice for the initial distribution for Φ_0 we can produce a stationary process $\{\Phi_n, n \in \mathbb{N}_+\}$.

For Markov chains we only need to consider first step stationarity to generate an entire stationary process. Suppose π is the initial invariant measure (initial distribution of the chain) with

$$\pi(A) = \int_{\mathcal{X}} \pi(dw)P(w, A).$$

Now, by iterating and the Chapman-Kolmogorov equations we have all n and $A \in \mathcal{B}(\mathcal{X})$

$$
\begin{aligned}
\pi(A) &= \int_{\mathcal{X}} \pi(dw)P(w,A) \\
&= \int_{\mathcal{X}} \left[\int_{\mathcal{X}} \pi(dx)P(x,dw) \right] P(w,A) \\
&= \int_{\mathcal{X}} \pi(dx) \int_{\mathcal{X}} P(x,dw)P(w,A) \\
&= \int_{\mathcal{X}} \pi(dx)P^2(x,A) \\
&= \vdots \\
&= \int_{\mathcal{X}} \pi(dx)P^n(x,A) = P_\pi(\Phi_n \in A).
\end{aligned}
$$

From the Markov property $\mathbf{\Phi}$ is stationary if and only if the distribution of Φ_n does not depend on n (time).

Lemma 8 *Let $\mathbf{\Phi}$ be a Markov chain and if $A \in \mathcal{B}(\mathcal{X})$ is uniformly transient with $U(x,A) \leq M$ for $x \in A$, then $U(x,A) \leq 1 + M$ for all $x \in \mathcal{X}$.*

Proof:

Proposition 18 *If the chain $\mathbf{\Phi}$ is positive, then it is recurrent.*

Proof:

Positive chains are often referred to as positive recurrent to reinforce the fact that they are recurrent.

Definition 31 (Positive Harris chains) *If $\mathbf{\Phi}$ is Harris recurrent and positive, then $\mathbf{\Phi}$ is called a positive Harris (recurrent) chain.*

Now we set out to show that an invariant probability measure exists and that it is unique, up to a multiplicative constant, for certain chains. We will begin by showing that chains that admit atoms are positive, and then extend to strongly aperiodic chains and then to recurrent chains.

Definition 32 (Subinvariant measures) *if μ is σ-finite and satisfies*

$$\mu(A) \geq \int_{\mathcal{X}} \mu(dw)P(w, A), \quad A \in \mathcal{B}(\mathcal{X}) \tag{11}$$

then μ is called subinvariant.

Iterating we get

$$\mu(A) \geq \int_{\mathcal{X}} \mu(dx)P^n(x, A). \tag{12}$$

Multiplying by $a(n)$, where a is a sampling distribution on $\mathbb{N}_+$, and then summing we get

$$\mu(A) \geq \int_{\mathcal{X}} \mu(dx)K_a(x, A). \tag{13}$$

Equations (12) and (13) tell us, respectively, that if μ is a subinvariant measure for $\mathbf{\Phi}$ is it also a subinvariant measure for any m-skeleton and for any sampled chain.

Proposition 19 *If $\mathbf{\Phi}$ is ψ-irreducible and if μ is any measure satisfying (11) with $\mu(A) < \infty$ for some $A \in \mathcal{B}^+(\mathcal{X})$, then*

(i) μ is σ-finite and thus μ is a subinvariant measure;

(ii) $\psi \prec \mu$;

(iii) if C is ν_a-petite, then $\mu(C) < \infty$;

(iv) if $\mu(\mathcal{X}) < \infty$, then μ is invariant.

Proof:

Proof: Use the first-entrance decomposition. In the limit we have

$$\begin{aligned} U(x,A) &= U_A(x,A) + \int_A U_A(x,dy)U(y,A) \\ &\leq L(x,A) + \int_A U_A(x,dy) \sup_{y\in A} U(y,A) \\ &= L(x,A) + L(x,A) \sup_{y\in A} U(y,A) \\ &\leq 1 + M. \end{aligned}$$

Proposition 18 *If the chain $\mathbf{\Phi}$ is positive, then it is recurrent.*

Proof: Suppose that the chain is transient. Then there is a countable cover of $\mathcal{X}$ with uniformly transient sets A_j. Hence, there exists an M_j such that $U(x, A_j) \leq M_j$ by the previous lemma. Now for any j, k we have

$$\pi(A_j) = k^{-1} \sum_{n=1}^{k} \int_{\mathcal{X}} \pi(dx) P^n(x, A_j) \leq k^{-1} M_j$$

As $k \uparrow \infty$ we have $\pi(A_j) = 0$. Therefore π cannot be a probability measure and $\mathbf{\Phi}$ is null. $\Rightarrow\Leftarrow$. □

Positive chains are often referred to as positive recurrent to reinforce the fact that they are recurrent.

Definition 31 (Positive Harris chains) *If $\mathbf{\Phi}$ is Harris recurrent and positive, then $\mathbf{\Phi}$ is called a positive Harris (recurrent) chain.*

Now we set out to show that an invariant probability measure exists and that it is unique, up to a multiplicative constant, for certain chains. We will begin by showing that chains that admit atoms are positive, and then extend to strongly aperiodic chains and then to recurrent chains.

Definition 32 (Subinvariant measures) *if μ is σ-finite and satisfies*

$$\mu(A) \geq \int_{\mathcal{X}} \mu(dw) P(w, A), \quad A \in \mathcal{B}(\mathcal{X}) \tag{13}$$

then μ is called subinvariant.

Iterating we get

$$\mu(A) \geq \int_{\mathcal{X}} \mu(dx) P^n(x, A). \tag{14}$$

Multiplying by $a(n)$, where a is a sampling distribution on $\mathbb{N}_+$, and then summing we get

$$\mu(A) \geq \int_{\mathcal{X}} \mu(dx) K_a(x, A). \tag{15}$$

Equations (14) and (15) tell us, respectively, that if μ is a subinvariant measure for $\mathbf{\Phi}$ is it also a subinvariant measure for any m-skeleton and for any sampled chain.

Proposition 19 *If $\mathbf{\Phi}$ is ψ-irreducible and if μ is any measure satisfying (13) with $\mu(A) < \infty$ for some $A \in \mathcal{B}^+(\mathcal{X})$, then*

(i) μ is σ-finite and thus μ is a subinvariant measure;

(ii) $\psi \prec \mu$;

(iii) if C is ν_a-petite, then $\mu(C) < \infty$;

(iv) if $\mu(\mathcal{X}) < \infty$, then μ is invariant.

Proof:
(i) Suppose $A \in \mathcal{B}^+(\mathcal{X})$ and $\mu(A) < \infty$. Consider the sets

$$A^*(j) = \{y : K_{1/2}(y, A) > j^{-1}\}.$$

Then

$$\begin{aligned} \infty > \mu(A) &\geq \int_{\mathcal{X}} \mu(dx) K_{1/2}(x, A) \\ &\geq \int_{A^*(j)} \mu(dx) K_{1/2}(x, A) \\ &> j^{-1} \mu(A^*(j)). \end{aligned}$$

So, each $A^*(j)$ has μ-finite measure. Furthermore since, $\lim_{j \uparrow \infty} A^*(j) = \cup_j A^*(j) = \mathcal{X}$, μ is σ-finite.

(ii) Let $A \in \mathcal{B}^+(\mathcal{X})$, i.e. $\psi(A) > 0$. Since $\mathbf{\Phi}$ is ψ-irreducible, $K_{1/2}(x, A) > 0$ which implies

$$\mu(A) \geq \int_{\mathcal{X}} \mu(dx) K_{1/2}(x, A) > 0.$$

Hence, $\mu(A) > 0$ whenever $\psi(A) > 0$, or $\psi \prec \mu$.

(iii) Suppose C is ν_a-petite. Then ν_a is a non-trivial measure and

$$K_a(x, B) \geq \nu_a(B)$$

for all $B \in \mathcal{B}(\mathcal{X})$ and $x \in \mathcal{C}$. Hence, there exists a set $A \in \mathcal{B}(\mathcal{X})$ with $\nu_a(A) > 0$ and, by assumption, $\mu(A) < \infty$. So, by (i),

$$\infty > \mu(A) \geq \int_{\mathcal{X}} \mu(dx) K_a(x, A) \geq \int_C \mu(dx) K_a(x, A) \geq \mu(C)\nu_a(A)$$

so that $\mu(C) < \infty$.

(iv) Suppose not. Suppose $\mu(\mathcal{X}) < \infty$ and μ is not invariant. Then there exists an A such that $\mu(A) > \int_{\mathcal{X}} \mu(dx) P(x, A)$.

$$\begin{aligned}
\mu(\mathcal{X}) = \mu(A) + \mu(A^c) &> \int_{\mathcal{X}} \mu(dx) P(x, A) + \int_{\mathcal{X}} \mu(dx) P(x, A^c) \\
&= \int_{\mathcal{X}} \mu(dx) P(x, \mathcal{X}) \\
&= \mu(\mathcal{X}).
\end{aligned}$$

This implies that $\mu(\mathcal{X}) = \infty$. $\Rightarrow\Leftarrow$. Hence, μ must be invariant. □

1.6.2 The existence of an invariant measure—chains with atoms

We are interested in Harris recurrent ψ-irreducible chains for MCMC theory. However, to show the existence of an invariant measure for recurrent ψ-irreducible chain, we will first proof the existence for chains with atoms (not necessarily recurrent) and then use Nummelin's splitting technique to extend the results to recurrent chains.

Lemma 9 *Suppose Φ is a Markov chain. Let $A \in \mathcal{B}(\mathcal{X})$. If $L(x, A) = 1$ for all $x \in A$, then A is a recurrent set.*

Proof: Suppose $L(x, A) = 1$ for all $x \in A$. Use the last-exit decomposition to get

$$U^{(z)}(x, A) = U_A^{(z)}(x, A) + \int_A U^{(z)}(x, dy) U_A^{(z)}(y, A).$$

Now take the limit as $z \uparrow 1$. Then

$$U(x, A) = L(x, A) + L(x, A)U(x, A) = 1 + U(x, A).$$

Therefore $U(x, A) = \infty$ for all $x \in A$ and A is recurrent by definition. □

Theorem 15 *Let $\mathbf{\Phi}$ be ψ-irreducible and suppose $\mathcal{X}$ contains an accessible atom $\boldsymbol{\alpha}$.*

(i) There exists a subinvariant measure $\mu^o_{\boldsymbol{\alpha}}$ for $\mathbf{\Phi}$ given by

$$\mu^o_{\boldsymbol{\alpha}}(A) = U_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}, A) = \sum_{n=1}^{\infty} {}_{\boldsymbol{\alpha}}P^n(\boldsymbol{\alpha}, A), \quad \forall\, A \in \mathcal{B}(\mathcal{X}),$$

where $\mu^o_{\boldsymbol{\alpha}}$ is invariant if and only if $\mathbf{\Phi}$ is recurrent.

(ii) The measure $\mu^o_{\boldsymbol{\alpha}}$ is minimal in the sense that if μ is subinvariant with $\mu(\boldsymbol{\alpha}) = 1$ then

$$\mu(A) \geq \mu^o_{\boldsymbol{\alpha}}(A), \quad \forall\, A \in \mathcal{B}(\mathcal{X}).$$

When $\mathbf{\Phi}$ is recurrent, $\mu^o_{\boldsymbol{\alpha}}$ is the unique (sub)invariant measure with $\mu(\boldsymbol{\alpha}) = 1$.

(iii) The subinvariant measure $\mu^o_{\boldsymbol{\alpha}}$ is a finite measure if and only if

$$\mathbb{E}_{\boldsymbol{\alpha}}(\tau_{\boldsymbol{\alpha}}) < \infty,$$

in which case $\mu^o_{\boldsymbol{\alpha}}$ is invariant.

Proof: (i) Let $A \in \mathcal{B}(\mathcal{X})$. Then

$$\begin{aligned}
\int_{\mathcal{X}} \mu^o_{\boldsymbol{\alpha}}(dy)P(y, A) &= \int_{\boldsymbol{\alpha}} \mu^o_{\boldsymbol{\alpha}}(dy)P(y, A) + \int_{\boldsymbol{\alpha}^c} \mu^o_{\boldsymbol{\alpha}}(dy)P(y, A) \\
&= \mu^o_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})P(\boldsymbol{\alpha}, A) + \int_{\boldsymbol{\alpha}^c} \sum_{n=1}^{\infty} {}_{\boldsymbol{\alpha}}P^n(\boldsymbol{\alpha}, dy)P(y, A) \\
&\leq P(\boldsymbol{\alpha}, A) + \sum_{n=1}^{\infty} \int_{\boldsymbol{\alpha}^c} {}_{\boldsymbol{\alpha}}P^n(\boldsymbol{\alpha}, dy)P(y, A) \\
&= {}_{\boldsymbol{\alpha}}P(\boldsymbol{\alpha}, A) + \sum_{n=2}^{\infty} {}_{\boldsymbol{\alpha}}P^n(\boldsymbol{\alpha}, A) \\
&= \mu^o_{\boldsymbol{\alpha}}(A).
\end{aligned}$$

Hence, $\mu^o_{\boldsymbol{\alpha}}$ is a subinvariant measure.

Now, $\mu^o_{\boldsymbol{\alpha}}$ is invariant if and only if $\mu^o_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) = 1$. But by definition, $\mu^o_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) = U_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = L(\boldsymbol{\alpha}, \boldsymbol{\alpha})$. But if $L(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = 1$, then $\mathbf{\Phi}$ is recurrent by Lemma 9.

(ii) Let μ be any subinvariant measure with $\mu(\boldsymbol{\alpha}) = 1$. We have

$$\begin{aligned}
\mu(A) &\geq \int_{\mathcal{X}} \mu(dy)P(y, A) \\
&\geq \int_{\boldsymbol{\alpha}} \mu(dy)P(y, A) \\
&= \mu(\boldsymbol{\alpha})P(\boldsymbol{\alpha}, A) = P(\boldsymbol{\alpha}, A).
\end{aligned}$$

Now assume that $\mu(A) \geq \sum_{m=1}^{n} {}_{\boldsymbol{\alpha}}P^m(\boldsymbol{\alpha}, A)$ for all A. Then by subinvariance

$$\begin{aligned}
\mu(A) &\geq \mu(\boldsymbol{\alpha})P(\boldsymbol{\alpha}, A) + \int_{\boldsymbol{\alpha}^c} \mu(dx)P(x, A) \\
&\geq P(\boldsymbol{\alpha}, A) + \int_{\boldsymbol{\alpha}^c} \left(\sum_{m=1}^{n} {}_{\boldsymbol{\alpha}}P^m(\boldsymbol{\alpha}, A) \right) P(x, A) \\
&= \sum_{m=1}^{n+1} {}_{\boldsymbol{\alpha}}P^m(\boldsymbol{\alpha}, A).
\end{aligned}$$

Taking the limit as $n \uparrow \infty$ shows that $\mu(A) \geq \mu^o_{\boldsymbol{\alpha}}(A), \quad \forall\, A \in \mathcal{B}(\mathcal{X})$.

Next, suppose $\boldsymbol{\Phi}$ is recurrent so that $\mu^o_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) = 1$. If μ and $\mu^o_{\boldsymbol{\alpha}}$ differ, then $\mu(A) > \mu^o_{\boldsymbol{\alpha}}(A)$ for some $A \in \mathcal{B}(\mathcal{X})$. By ψ-irreducibility there exists an n such that $P^n(x, \boldsymbol{\alpha}) > 0$ for all $x \in \mathcal{X}$, since $\psi(\boldsymbol{\alpha}) > 0$ ($\boldsymbol{\alpha}$ is an accessible atom). Then

$$\begin{aligned}
1 = \mu(\boldsymbol{\alpha}) &\geq \int_{\mathcal{X}} \mu(dx)P^n(x, \boldsymbol{\alpha}) \\
&= \int_{A} \mu(dx)P^n(x, \boldsymbol{\alpha}) + \int_{A^c} \mu(dx)P^n(x, \boldsymbol{\alpha}) \\
&> \int_{A} \mu^o_{\boldsymbol{\alpha}}(dx)P^n(x, \boldsymbol{\alpha}) + \int_{A^c} \mu^o_{\boldsymbol{\alpha}}(dx)P^n(x, \boldsymbol{\alpha}) \\
&= \int_{\mathcal{X}} \mu^o_{\boldsymbol{\alpha}}(dx)P^n(x, \boldsymbol{\alpha}) = \mu^o_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) = 1.
\end{aligned}$$

This leads to a contradiction so that $\mu(A) = \mu^o_{\boldsymbol{\alpha}}(A)$. Therefore, $\mu = \mu^o_{\boldsymbol{\alpha}}$ and $\mu^o_{\boldsymbol{\alpha}}$ is the unique (sub)invariant measure.

(*iii*) If $\mu^o_{\boldsymbol{\alpha}}$ is finite, then it is invariant (by Proposition 19(*iv*) on page 45). Also by the construction of $\mu^o_{\boldsymbol{\alpha}}$ we have

$$\mu^o_{\boldsymbol{\alpha}}(\mathcal{X}) = \sum_{n=1}^{\infty} {}_{\boldsymbol{\alpha}}P^n(\boldsymbol{\alpha}, \mathcal{X}) = \sum_{n=1}^{\infty} P_{\boldsymbol{\alpha}}(\tau_{\boldsymbol{\alpha}} \geq n) = \mathbb{E}_{\boldsymbol{\alpha}}(\tau_{\boldsymbol{\alpha}}).$$

Therefore an invariant probability measure exists if and only if the mean return time to $\boldsymbol{\alpha}$ is finite. In the above the first equality holds by definition. The second equality holds because

$${}_{\boldsymbol{\alpha}}P^n(\boldsymbol{\alpha}, \mathcal{X}) := P_{\boldsymbol{\alpha}}(\Phi_n \in \mathcal{X}; \tau_{\boldsymbol{\alpha}} \geq n) = P_{\boldsymbol{\alpha}}(\tau_{\boldsymbol{\alpha}} \geq n).$$

The third equality holds because

$$\sum_{n=1}^{\infty} P_{\boldsymbol{\alpha}}(\tau_{\boldsymbol{\alpha}} \geq n) = \sum_{n=1}^{\infty} \mathbb{E}_{\boldsymbol{\alpha}}[\mathbb{I}(\tau_\alpha \geq n)] = \mathbb{E}_{\boldsymbol{\alpha}}\left[\sum_{n=1}^{\infty} \mathbb{I}(\tau_\alpha \geq n)\right] = \mathbb{E}_{\boldsymbol{\alpha}}(\tau_{\boldsymbol{\alpha}}).$$