1.5.3 Cyclic Behavior

An artificial example of cyclic behavior on a continuous state space is the following. Let $\mathcal{X} = [0, d)$ and U_i denote the uniform distribution on $[i, i+1)$. Let $\mathbf{\Phi}$ be a chain with transition probability kernel

$$P(x, \cdot) := \mathbb{I}_{[i-1,i)}(x) U_i(\cdot), \quad i = 0, 1, \ldots, d-1 \pmod d.$$

Then it is easy to see that $\mathbf{\Phi}$ cycles through the subsets $[i, i+1)$: if $\Phi_n = x \in [i-1, i)$, then $P(\Phi_{n+1} \in [i, i+1) \mid \Phi_n = x) \equiv 1$.

We will now show that this finite cyclic behavior is the worst behavior to be found for a ψ-irreducible chain. This is due to existence of small sets.

Suppose C is any ν_M-small set. Without loss of generality, we can assume $\nu_M(C) > 0$ by Proposition 9(iii). This small set C and the corresponding measure ν_M will be used to define a cycle for a general irreducible Markov chain. Since C is ν_M-small, $P^M(x, C) \geq \nu_M(C) > 0$, for $x \in C$, so when the chain starts in C it returns with positive probability at time M. We will define the set E_C to be the set of time points for which C is a small set with minorizing measure proportional to ν_M:

$$E_C = \{n \geq 1 : C \text{ is } \nu_n\text{-small, with } \nu_n = \delta_n \nu_M \text{ for some } \delta_n > 0\}.$$

E_C is closed under addition: Let $B \in \mathcal{B}(\mathcal{X})$, n, $m \in E_C$. Then, for $x \in C$,

$$P^{n+m}(x, B) \geq \int_C P^n(x, dy) P^m(y, B) \geq [\delta_m \delta_n \nu_M(C)] \nu_M(B).$$

Hence, there is a natural period for the set C: the greatest common divisor of E_C. Furthermore, by a number theory result (see, Billingsley, *Probability and Measure*, Theorem A21, 1995), C is ν_{nd}-small for all n larger than some n_0.

The following theorem states that this value is a property of the chain $\mathbf{\Phi}$ and does not depend on the C chosen.

Theorem 6 *Let $\mathbf{\Phi}$ be ψ-irreducible Markov chain on $\mathcal{X}$. Let $C \in \mathcal{B}^+(\mathcal{X})$ be ν_M-small and let d be the greatest common divisor of the set E_C. Then there exist disjoint sets $D_0, \ldots, D_{d-1} \in \mathcal{B}(\mathcal{X})$ (called a "d-cycle") such that*

(i) if $x \in D_i$, then $P(x, D_{i+1}) = 1, \quad i = 0, \ldots, d-1 \pmod d$;

(ii) $\psi(N) = 0$ where $N = \left(\cup_{i=0}^{d-1} D_i\right)^c$.

The d-cycle $\{D_i\}$ is maximal in the sense that for any other collection $\{d', D'_k, k = 0, \ldots, d'-1\}$ that satisfy the above conditions, d' divides d. If $d = d'$, then, $D'_i = D_i$ a.e. $[\psi]$, by a reordering if necessary.

Proof: (Meyn & Tweedie, pp. 111-112).

Definition 19 (Periodic and Aperiodic Chains) *Let $\mathbf{\Phi}$ be a φ-irreducible Markov chain.*

- *The largest d for which a d-cycle occurs for $\mathbf{\Phi}$ is called the period of $\mathbf{\Phi}$.*
- *If $d = 1$, then $\mathbf{\Phi}$ is called aperiodic.*
- *If there exists a ν_1-small set A with $\nu_1(A) > 0$, then $\mathbf{\Phi}$ is called strongly aperiodic.*

Proposition 10 *Let $\mathbf{\Phi}$ be ψ-irreducible.*

(i) If $\mathbf{\Phi}$ is strongly aperiodic, the minorization condition holds.

(ii) For all $\epsilon \in (0, 1)$, the K_ϵ-chain is strongly aperiodic.

(iii) If $\mathbf{\Phi}$ is aperiodic, then every skeleton is ψ-irreducible and aperiodic, and there exists an m such that the m-skeleton is strongly aperiodic.

Proof:

This theorem shows it is desirable to work with strongly aperiodic, irreducible chains as then the minorization condition holds and we can take advantage of the split chain. However, this condition is not met in general and so we will restrict ourselves to aperiodic chains and prove results for strongly aperiodic chains and then use special methods to extend them to general chains through the m-skeleton or K_ϵ-chain.

Proposition 11 *If $\mathbf{\Phi}$ is ψ-irreducible with period d and d-cycle $\{D_i,\, i = 0, \ldots, d-1\}$, then each of the sets D_i is an absorbing ψ-irreducible set for the chain $\mathbf{\Phi}_d$ corresponding to the transition probability kernel P^d. Furthermore, $\mathbf{\Phi}_d$ is aperiodic on each D_i.*

Proof:

1.5.4 Petite sets

A convenient tool for the analysis of Markov chains is the sampled chain. This idea extends the idea of the m-skeleton and K_ϵ chain. Petite sets are defined in terms of the sampled chain and are a generalization of small sets.

Let $a = \{a(n)\}$ be a distribution (or probability measure) on $\mathbb{N}_+$. Consider the Markov chain $\mathbf{\Phi}_a$ with probability transition kernel

$$K_a(x, A) := \sum_{n=0}^{\infty} P^n(x, A)a(n), \quad x \in \mathcal{X},\, A \in \mathcal{B}(\mathcal{X}).$$

We call $\mathbf{\Phi}_a$ the K_a-chain with sampling distribution a. Probabilistically, $\mathbf{\Phi}_a$ has the interpretation of being the chain $\mathbf{\Phi}$ "sampled" at time points drawn successively according to the distribution a. Note that if $a = \delta_m$ is the Dirac measure with $\delta_m(m) = 1$, then the K_{δ_m}-chain is the m-skeleton with transition probability kernel P^m. If a is the geometric distribution with

$$a(n) = (1-\epsilon)\epsilon^n, \quad n \in \mathbb{N}_+,$$

then we have the resolvent, K_ϵ. This concept of sampled chains allows us to determine conditions under which one set in uniformly accessible from another set.

Definition 20 *We say that a set $B \in \mathcal{B}(\mathcal{X})$ is 'uniformly accessible using a' (the distribution a) from another set $A \in \mathcal{B}(\mathcal{X})$ if there exists $\delta > 0$ such that*

$$\inf_{x \in A} K_a(x, B) \geq \delta. \tag{8}$$

We then write $A \overset{a}{\rightsquigarrow} B$.

Note: Compare this to the definition of uniformly accessible and make sure you understand the difference. The difference being that *uniformly accessible* is in relation to the entire chain $\mathbf{\Phi}$ and its probability transition kernel P and the condition that $\inf_{x\in A} L(x, A) \geq \delta > 0$ while *uniformly accessible using* a uses the sampled chain $\mathbf{\Phi}_a$ and its probability transition kernel K_a.

Lemma 1 *If* $A \overset{a}{\rightsquigarrow} B$ *for some distribution* a*, then* $A \rightsquigarrow B$.

Proof:

Lemma 2

(i) If a *and* b *are distributions on* $\mathbb{N}_+$*, then the sampled chains* K_a *and* K_b *satisfy the generalized Chapman-Kolmogorov equations*

$$K_{a*b}(x, A) = \int_{\mathcal{X}} K_a(x, dy) K_b(y, A)$$

where $*$ *is the convolution operator.*

(ii) If $A \overset{a}{\rightsquigarrow} B$ *and* $B \overset{b}{\rightsquigarrow} C$*, then* $A \overset{a*b}{\rightsquigarrow} C$.

(iii) If a *is a distribution on* $\mathbb{N}_+$*, then the sampled chain with transition probability kernel* K_a *satisfies the relation*

$$U(x, A) \geq \int_{\mathcal{X}} U(x, dy) K_a(y, A).$$

Proof:

Small sets always exist for ψ-irreducible chains. We will also show that every small set is petite and if the chain is aperiodic, then every petite set is small. For MCMC theory we will only be concerned with aperiodic chains. But for MCMC theory we also need the concepts of recurrence and Harris recurrence, which we will study on their own merit (independent of whether the chain is aperiodic). Some of the results of recurrence will rely on petite sets. Hence,