10

The Existence of π

In our treatment of the structure and stability concepts for irreducible chains we have to this point considered only the dichotomy between transient and recurrent chains.

For transient chains there are many areas of theory that we shall not investigate further, despite the flourishing research that has taken place in both the mathematical development and the application of transient chains in recent years. Areas which are notable omissions from our treatment of Markovian models thus include the study of potential theory and boundary theory [223], as well as the study of renormalized models approximated by diffusions and the quasi-stationary theory of transient processes [71, 5].

Rather, we concentrate on recurrent chains which have stable properties without renormalization of any kind, and develop the consequences of the concept of recurrence.

In this chapter we further divide recurrent chains into *positive* and *null* recurrent chains, and show here and in the next chapter that the former class provide stochastic stability of a far stronger kind than the latter.

For many purposes, the strongest possible form of stability that we might require in the presence of persistent variation is that the distribution of Φ_n does not change as n takes on different values. If this is the case, then by the Markov property it follows that the finite dimensional distributions of $\boldsymbol{\Phi}$ are invariant under translation in time. Such considerations lead us to the consideration of *invariant measures.*

Invariant measures

A σ-finite measure π on $\mathcal{B}(\mathsf{X})$ with the property

$$\pi(A) = \int_{\mathsf{X}} \pi(dx)P(x,A), \qquad A \in \mathcal{B}(\mathsf{X}) \tag{10.1}$$

will be called *invariant.*

Although we develop a number of results concerning invariant measures, the key conclusion in this chapter is undoubtedly

Theorem 10.0.1 *If the chain $\boldsymbol{\Phi}$ is recurrent then it admits a unique (up to constant multiples) invariant measure π, and the measure π has the representation, for any $A \in \mathcal{B}^+(\mathsf{X})$*

$$\pi(B) = \int_A \pi(dw) \mathsf{E}_w \Big[\sum_{n=1}^{\tau_A} \mathbb{1}\{\Phi_n \in B\}\Big], \qquad B \in \mathcal{B}(\mathsf{X}). \tag{10.2}$$

The invariant measure π is finite (rather than merely σ-finite) if there exists a petite set C such that

$$\sup_{x \in C} \mathsf{E}_x[\tau_C] < \infty.$$

Proof The existence and representation of invariant measures for recurrent chains is proved in full generality in Theorem 10.4.9: the proof exploits, via the Nummelin splitting technique, the corresponding theorem for chains with atoms as in Theorem 10.2.1, in conjunction with a representation for invariant measures given in Theorem 10.4.9. The criterion for finiteness of π is in Theorem 10.4.10. □

If an invariant measure is finite, then it may be normalized to a stationary probability measure, and in practice this is the main stable situation of interest. If an invariant measure has infinite total mass, then its probabilistic interpretation is much more difficult, although for recurrent chains, there is at least the interpretation as described in (10.2).

These results lead us to define the following classes of chains.

> Positive and Null Chains
>
> Suppose that $\boldsymbol{\Phi}$ is ψ-irreducible, and admits an invariant probability measure π. Then $\boldsymbol{\Phi}$ is called a *positive* chain.
>
> If $\boldsymbol{\Phi}$ does not admit such a measure, then we call $\boldsymbol{\Phi}$ *null*.

10.1 Stationarity and Invariance

10.1.1 Invariant measures

Processes with the property that for any k, the marginal distribution of $\{\Phi_n, \dots, \Phi_{n+k}\}$ does not change as n varies are called *stationary processes*, and whilst it is clear that in general a Markov chain will not be stationary, since in a particular realization we may have $\Phi_0 = x$ with probability one for some fixed x, it is possible that with an appropriate choice of the initial distribution for Φ_0 we may produce a stationary process $\{\Phi_n, n \in \mathbb{Z}_+\}$.

It is immediate that we only need to consider a form of first step stationarity in order to generate an entire stationary process. Given an initial invariant probability measure π such that

$$\pi(A) = \int_{\mathsf{X}} \pi(dw) P(w, A), \tag{10.3}$$

we can iterate to give

$$\begin{aligned}
\pi(A) &= \int_{\mathsf{X}} \left[\int_{\mathsf{X}} \pi(dx) P(x,dw)\right] P(w,A) \\
&= \int_{\mathsf{X}} \pi(dx) \int_{\mathsf{X}} P(x,dw) P(w,A) \\
&= \int_{\mathsf{X}} \pi(dx) P^2(x,A) \\
&\ \ \vdots \\
&= \int_{\mathsf{X}} \pi(dx) P^n(x,A) = \mathsf{P}_\pi(\Phi_n \in A),
\end{aligned} \tag{10.4}$$

for any n and all $A \in \mathcal{B}(\mathsf{X})$.

From the Markov property, it is clear that $\boldsymbol{\Phi}$ is stationary if and only if the distribution of Φ_n does not vary with time. We have immediately

Proposition 10.1.1 *If the chain $\boldsymbol{\Phi}$ is positive then it is recurrent.*

PROOF Suppose that the chain is positive and let π be a invariant probability measure. If the chain is also transient, let A_j be a countable cover of X with uniformly transient sets, as guaranteed by Theorem 8.3.4, with $U(x, A_j) \le M_j$, say.

Using (10.4) we have for any j, k

$$k\pi(A_j) = \sum_{n=1}^{k} \int \pi(dw) P^n(w, A_j) \le M_j$$

and since the left hand side remains finite as $k \to \infty$, we have $\pi(A_j) = 0$. This implies π is trivial so we have a contradiction. □

Positive chains are often called "positive recurrent" to reinforce the fact that they are recurrent. This also naturally gives the definition

---

Positive Harris chains

If $\boldsymbol{\Phi}$ is Harris recurrent and positive, then $\boldsymbol{\Phi}$ is called a *positive Harris* chain.

It is of course not yet clear that an invariant probability measure π ever exists, or whether it will be unique when it does exist. It is the major purpose of this chapter to find conditions for the existence of π, and to prove that for any positive (and indeed recurrent) chain, π is essentially unique.

Invariant probability measures are important not merely because they define stationary processes. They will also turn out to be the measures which define the long term or ergodic behavior of the chain. To understand why this should be plausible, consider $\mathsf{P}_\mu(\Phi_n \in \cdot\,)$ for any starting distribution μ. If a limiting measure γ_μ exists in a suitable topology on the space of probability measures, such as

$$\mathsf{P}_\mu(X_n \in A) \to \gamma_\mu(A)$$

for all $A \in \mathcal{B}(\mathsf{X})$, then

$$\begin{aligned}\gamma_\mu(A) &= \lim_{n\to\infty} \int \mu(dx)P^n(x,A) \\ &= \lim_{n\to\infty} \int_{\mathsf{X}} \mu(dx) \int P^{n-1}(x,dw)P(w,A) \\ &= \int_{\mathsf{X}} \gamma_\mu(dw)P(w,A), \end{aligned} \tag{10.5}$$

since setwise convergence of $\int \mu(dx)P^n(x,\cdot)$ implies convergence of integrals of bounded measurable functions such as $P(w,A)$.

Hence if a limiting distribution exists, it is an invariant probability measure; and obviously, if there is a unique invariant probability measure, the limit γ_μ will be independent of μ whenever it exists.

We will not study the existence of such limits properly until Part III, where our goal will be to develop asymptotic properties of $\boldsymbol{\Phi}$ in some detail. However, motivated by these ideas, we will give in Section 10.5 one example, the linear model, where this route leads to the existence of an invariant probability measure.

10.1.2 Subinvariant measures

The easiest way to investigate the existence of π is to consider a yet wider class of measures, satisfying inequalities related to the invariant equation (10.1).

Subinvariant measures

If μ is σ-finite and satisfies

$$\mu(A) \geq \int_{\mathsf{X}} \mu(dx)P(x,A), \quad A \in \mathcal{B}(\mathsf{X}) \tag{10.6}$$

then μ is called *subinvariant*.

The following generalization of the subinvariance equation (10.6) is often useful: we have, by iterating (10.6),

$$\mu(B) \geq \int \mu(dw)P^n(w,B)$$

and hence, multiplying by $a(n)$ and summing,

$$\mu(B) \geq \int \mu(dw)K_a(w,B), \tag{10.7}$$

for any sampling distribution a.

We begin with some structural results for arbitrary subinvariant measures.

Proposition 10.1.2 *Suppose that $\boldsymbol{\Phi}$ is ψ-irreducible. If μ is any measure satisfying (10.6) with $\mu(A) < \infty$ for some one $A \in \mathcal{B}^+(\mathsf{X})$, then*

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

(ii) *$\mu \succ \psi$;*

(iii) *if C is petite then $\mu(C) < \infty$;*

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

PROOF Suppose $\mu(A) < \infty$ for some A with $\psi(A) > 0$. Using $A^*(j) = \{y : K_{a_{1/2}}(y, A) > j^{-1}\}$, we have by (10.7),

$$\infty > \mu(A) \geq \int_{A^*(j)} \mu(dw) K_{a_{1/2}}(w, A) \geq j^{-1} \mu(A^*(j));$$

since $\bigcup A^*(j) = \mathsf{X}$ when $\psi(A) > 0$, such a μ must be σ-finite.

To prove (ii) observe that, by (10.7), if $B \in \mathcal{B}^+(\mathsf{X})$ we have $\mu(B) > 0$, so $\mu \succ \psi$.

Thirdly, if C is ν_a-petite then there exists a set B with $\nu_a(B) > 0$ and $\mu(B) < \infty$, from (i). By (10.7) we have

$$\mu(B) \geq \int \mu(dw) K_a(w, B) \geq \mu(C)\nu_a(B) \tag{10.8}$$

and so $\mu(C) < \infty$ as required.

Finally, if there exists some A such that $\mu(A) > \int \mu(dy) P(y, A)$ then we have

$$\begin{aligned} \mu(\mathsf{X}) = \mu(A) + \mu(A^c) &> \int \mu(dy) P(y, A) + \int \mu(dy) P(y, A^c) \\ &= \int \mu(dy) P(y, \mathsf{X}) \\ &= \mu(\mathsf{X}) \end{aligned} \tag{10.9}$$

and if $\mu(\mathsf{X}) < \infty$ we have a contradiction.

□

The major questions of interest in studying subinvariant measures lie with recurrent chains, for we always have

Proposition 10.1.3 *If the chain $\boldsymbol{\Phi}$ is transient then there exists a strictly subinvariant measure for $\boldsymbol{\Phi}$.*

PROOF Suppose that $\boldsymbol{\Phi}$ is transient: then by Theorem 8.3.4, we have that the measures μ_x given by

$$\mu_x(A) = U(x, A), \qquad A \in \mathcal{B}(\mathsf{X})$$

are σ-finite; and trivially

$$\mu_x(A) = P(x, A) + \int \mu_x(dy) P(y, A) \geq \int \mu_x(dy) P(y, A), \qquad A \in \mathcal{B}(\mathsf{X}) \tag{10.10}$$

so that each μ_x is subinvariant (and obviously strictly subinvariant, since there is some A with $\mu_x(A) < \infty$ such that $P(x, A) > 0$). □

We now move on to study recurrent chains, where the existence of a subinvariant measure is less obvious.

10.2 The existence of π: chains with atoms

Rather than pursue the question of existence of invariant and subinvariant measures on a fully countable space in the first instance, we prove here that the existence of just one atom α in the space is enough to describe completely the existence and structure of such measures.

The following theorem obviously incorporates countable space chains as a special case; but the main value of this presentation will be in the development of a theory for general space chains via the split chain construction of Section 5.1.

Theorem 10.2.1 *Suppose Φ is ψ-irreducible, and X contains an accessible atom α.*

(i) *There is always a subinvariant measure μ_α° for Φ given by*

$$\mu_\alpha^\circ(A) = U_\alpha(\alpha, A) = \sum_{n=1}^{\infty} {}_\alpha P^n(\alpha, A), \qquad A \in \mathcal{B}(\mathsf{X}); \tag{10.11}$$

and μ_α° is invariant if and only if Φ is recurrent.

(ii) *The measure μ_α° is minimal in the sense that if μ is subinvariant with $\mu(\alpha) = 1$, then*

$$\mu(A) \geq \mu_\alpha^\circ(A), \qquad A \in \mathcal{B}(\mathsf{X}).$$

When Φ is recurrent, μ_α° is the unique (sub)invariant measure with $\mu(\alpha) = 1$.

(iii) *The subinvariant measure μ_α° is a finite measure if and only if*

$$\mathsf{E}_\alpha[\tau_\alpha] < \infty,$$

in which case μ_α° is invariant.

PROOF **(i)** By construction we have for $A \in \mathcal{B}(\mathsf{X})$

$$\begin{aligned} \int_{\mathsf{X}} \mu_\alpha^\circ(dy) P(y, A) &= \mu_\alpha^\circ(\alpha) P(\alpha, A) + \int_{\alpha^c} \sum_{n=1}^{\infty} {}_\alpha P^n(\alpha, dy) P(y, A) \\ &\leq {}_\alpha P(\alpha, A) + \sum_{n=2}^{\infty} {}_\alpha P^n(\alpha, A) \qquad (10.12) \\ &= \mu_\alpha^\circ(A), \end{aligned}$$

where the inequality comes from the bound $\mu_\alpha^\circ(\alpha) \leq 1$. Thus μ_α° is subinvariant, and is invariant if and only if $\mu_\alpha^\circ(\alpha) = \mathsf{P}_\alpha(\tau_\alpha < \infty) = 1$; that is, from Proposition 8.3.1, if and only if the chain is recurrent.

(ii) Let μ be any subinvariant measure with $\mu(\alpha) = 1$. By subinvariance,

$$\begin{aligned} \mu(A) &\geq \int_{\mathsf{X}} \mu(dw) P(w, A) \\ &\geq \mu(\alpha) P(\alpha, A) = P(\alpha, A). \end{aligned}$$

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

$$\begin{aligned}\mu(A) &\geq \mu(\boldsymbol{\alpha})P(\boldsymbol{\alpha}, A) + \int_{\alpha^c} \mu(dw)P(w, A) \\ &\geq P(\boldsymbol{\alpha}, A) + \int_{\alpha^c} \Big[\sum_{m=1}^{n} {}_{\alpha}P^m(\boldsymbol{\alpha}, dw)\Big] P(w, A) \\ &= \sum_{m=1}^{n+1} {}_{\alpha}P^m(\boldsymbol{\alpha}, A).\end{aligned}$$

Taking $n \uparrow \infty$ shows that $\mu(A) \geq \mu^\circ_\alpha(A)$ for all $A \in \mathcal{B}(\mathsf{X})$.

Suppose $\boldsymbol{\Phi}$ is recurrent, so that $\mu^\circ_\alpha(\boldsymbol{\alpha}) = 1$. If μ°_α differs from μ, there exists A and n such that $\mu(A) > \mu^\circ_\alpha(A)$ and $P^n(w, \boldsymbol{\alpha}) > 0$ for all $w \in A$, since $\psi(\boldsymbol{\alpha}) > 0$. By minimality, subinvariance of μ, and invariance of μ°_α,

$$\begin{aligned}1 = \mu(\boldsymbol{\alpha}) &\geq \int_{\mathsf{X}} \mu(dw)P^n(w, \boldsymbol{\alpha}) \\ &> \int_{\mathsf{X}} \mu^\circ_\alpha(dw)P^n(w, \boldsymbol{\alpha}) \\ &= \mu^\circ_\alpha(\boldsymbol{\alpha}) = 1.\end{aligned}$$

Hence we must have $\mu = \mu^\circ_\alpha$, and thus when $\boldsymbol{\Phi}$ is recurrent, μ°_α is the unique (sub) invariant measure.

(iii) If μ°_α is finite it follows from Proposition 10.1.2 (iv) that μ°_α is invariant. Finally

$$\mu^\circ_\alpha(\mathsf{X}) = \sum_{n=1}^{\infty} \mathsf{P}_\alpha(\tau_\alpha \geq n) \tag{10.13}$$

and so an invariant probability measure exists if and only if the mean return time to $\boldsymbol{\alpha}$ is finite, as stated. □

We shall use π to denote the unique invariant measure in the recurrent case. Unless stated otherwise we will assume π is normalized to be a probability measure when $\pi(\mathsf{X})$ is finite.

The invariant measure μ°_α has an equivalent sample path representation for recurrent chains:

$$\mu^\circ_\alpha(A) = \mathsf{E}_\alpha\Big[\sum_{n=1}^{\tau_\alpha} \mathbb{1}\{\Phi_n \in A\}\Big], \qquad A \in \mathcal{B}(\mathsf{X}). \tag{10.14}$$

This follows from the definition of the taboo probabilities ${}_{\alpha}P^n$.

As an immediate consequence of this construction we have the following elegant criterion for positivity.

Theorem 10.2.2 (Kac's Theorem) *If $\boldsymbol{\Phi}$ is ψ-irreducible and admits an atom $\boldsymbol{\alpha} \in \mathcal{B}^+(\mathsf{X})$, then $\boldsymbol{\Phi}$ is positive recurrent if and only if $\mathsf{E}_\alpha[\tau_\alpha] < \infty$; and if π is the invariant probability measure for $\boldsymbol{\Phi}$ then*

$$\pi(\boldsymbol{\alpha}) = (\mathsf{E}_\alpha[\tau_\alpha])^{-1}. \tag{10.15}$$

PROOF If $\mathsf{E}_\alpha[\tau_\alpha] < \infty$, then also $L(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = 1$, and by Proposition 8.3.1 $\boldsymbol{\Phi}$ is recurrent; it follows from the structure of π in (10.11) that π is finite so that the chain is positive.

Conversely, $\mathsf{E}_\alpha[\tau_\alpha] < \infty$ when the chain is positive from the structure of the unique invariant measure.

By the uniqueness of the invariant measure normalized to be a probability measure π we have

$$\pi(\boldsymbol{\alpha}) = \frac{\mu^{\circ}_{\alpha}(\boldsymbol{\alpha})}{\mu^{\circ}_{\alpha}(\mathsf{X})} = \frac{U_{\alpha}(\boldsymbol{\alpha},\boldsymbol{\alpha})}{U_{\alpha}(\boldsymbol{\alpha},\mathsf{X})} = \frac{1}{\mathsf{E}_{\alpha}[\tau_{\alpha}]}$$

which is (10.15). □

The relationship (10.15) is often known as Kac's Theorem. For countable state space models it immediately gives us

Proposition 10.2.3 *For a positive recurrent irreducible Markov chain on a countable space, there is a unique (up to constant multiples) invariant measure π given by*

$$\pi(x) = [\mathsf{E}_x[\tau_x]]^{-1}$$

for every $x \in \mathsf{X}$. □

We now illustrate the use of the representation of π for a number of countable space models.

10.3 Invariant measures: countable space models

10.3.1 Renewal chains

Forward recurrence time chains Consider the forward recurrence time process $\mathbf{V}^{+}$ with

$$P(1,j) = p(j), \qquad j \geq 1; \qquad P(j,j-1) = 1, \qquad j > 1. \tag{10.16}$$

As noted in Section 8.1.2, this chain is always recurrent since $\sum p(j) = 1$.

By construction we have that

$${}_1P^n(1,j) = p(j+n-1), \qquad j \leq n$$

and zero otherwise; thus the minimal invariant measure satisfies

$$\pi(j) = U_1(1,j) = \sum_{n \geq j} p(n) \tag{10.17}$$

which is finite if and only if

$$\sum_{j=1}^{\infty} \pi(j) = \sum_{j=1}^{\infty} \sum_{n=j}^{\infty} p(n) = \sum_{n=1}^{\infty} np(n) < \infty : \tag{10.18}$$

that is, if and only if the renewal distribution $\{p(i)\}$ has finite mean.

It is, of course, equally easy to deduce this formula by solving the invariant equations themselves, but the result is perhaps more illuminating from this approach.

Linked forward recurrence time chains Consider the forward recurrence chain with transition law (10.16), and define the bivariate chain $\mathbf{V}^* = (V_1^+(n), V_2^+(n))$ on the space $\mathsf{X}^* := \{1,2,\ldots\} \times \{1,2,\ldots\}$, with the transition law

$$\begin{array}{rcllcl} P((i,j),(i-1,j-1)) &=& 1, & i,j &>& 1; \\ P((1,j),(k,j-1)) &=& p(k), & k,j &>& 1; \\ P((i,1),(i-1,k)) &=& p(k), & i,k &>& 1; \\ P((1,1),(j,k)) &=& p(j)p(k), & j,k &>& 1. \end{array} \tag{10.19}$$

This chain is constructed by taking the two independent copies $V_1^+(n), V_2^+(n)$ of the forward recurrence chain and running them independently.

Now suppose that the distribution $\{p(j)\}$ is *periodic with period d*: that is, the greatest common divisor d of the set $N_p = \{n : p(n) > 0\}$ is d. We show that $\mathbf{V}^*$ is ψ-irreducible and positive if $\{p(j)\}$ has period $d = 1$ and $\sum_{n\geq 1} np(n) < \infty$.

By the definition of d we have that there must exist $r, s \in N_p$ with greatest common divisor d, and by Lemma D.7.3 there exist integers n, m such that

$$nr = ms + d:$$

without loss of generality we can assume $n, m > 0$.

We show that the bivariate chain $\mathbf{V}^*$ is $\delta_{1,1}$-irreducible if $d = 1$.

To see this, note that for any pair (i,j) with $i \geq j$ we have

$$P^{j+(i-j)nr}((i,j),(1,1)) \geq [p(r)]^{(i-j)nr-1}[p(s)]^{(i-j)ms-1} > 0$$

since

$$j + (i-j)nr = i + (i-j)ms.$$

Moreover $\mathbf{V}^*$ is positive Harris recurrent on X^* provided only $\sum_k kp(k) < \infty$, as was the case for the single copy of the forward recurrence time chain. To prove this we need only note that the product measure $\pi^*(i,j) = \pi(i)\pi(j)$ is invariant for $\mathbf{V}^*$, where

$$\pi(j) = \sum_{k\geq j} p(k) / \sum_k kp(k)$$

is the invariant probability measure for the forward recurrence time process from (10.17) and (10.18); positive Harris recurrence follows since $\pi^*(\mathsf{X}^*) = [\pi(\mathsf{X})]^2 = 1$.

These conditions for positive recurrence of the bivariate forward time process will be of critical use in the development of the asymptotic properties of general chains in Part III.

10.3.2 The number in an M/G/1 queue

Recall from Section 3.3.3 that $\mathbf{N}^*$ is a modified random walk on a half line with increment distribution concentrated on the integers $\{\ldots,-1,0,1\}$ having the transition probability matrix of the form

$$P = \begin{pmatrix} q_0 & q_1 & q_2 & q_3 & \cdots \\ q_0 & q_1 & q_2 & q_3 & \cdots \\ & q_0 & q_1 & q_2 & \cdots \\ & & q_0 & q_1 & \cdots \\ & & & q_0 & \cdots \end{pmatrix}$$

where $q_i = \mathsf{P}(Z = i - 1)$ for the increment variable in the chain when the server is busy; that is, for transitions from states other than $\{0\}$. The chain $\mathbf{N}^*$ is always ψ-irreducible if $q_0 > 0$, and irreducible in the standard sense if also $q_0 + q_1 < 1$, and we shall assume this to be the case to avoid trivialities.

In this case, we can actually solve the invariant equations explicitly. For $j \geq 1$, (10.1) can be written

$$\pi(j) = \sum_{k=0}^{j+1} \pi(k) q_{j+1-k}. \tag{10.20}$$

and if we define

$$\bar{q}_j = \sum_{n=j+1}^{\infty} q_n$$

we get the system of equations

$$\begin{array}{rcl} \pi(1)q_0 & = & \pi(0)\bar{q}_0 \\ \pi(2)q_0 & = & \pi(0)\bar{q}_1 + \pi(1)\bar{q}_1 \\ \pi(3)q_0 & = & \pi(0)\bar{q}_2 + \pi(1)\bar{q}_2 + \pi(2)\bar{q}_1 \\ & & \dots \end{array} \tag{10.21}$$

In this case, therefore, we always get a unique invariant measure, regardless of the transience or recurrence of the chain.

The criterion for positivity follows from (10.21). Note that the mean increment β of Z satisfies

$$\beta = \sum_{j \geq 0} \bar{q}_j - 1$$

so that formally summing both sides of (10.21) gives, since $q_0 = 1 - \bar{q}_0$

$$(1 - \bar{q}_0) \sum_{j=1}^{\infty} \pi(j) = (\beta + 1)\pi(0) + (\beta + 1 - \bar{q}_0) \sum_{j=1}^{\infty} \pi(j). \tag{10.22}$$

If the chain is positive, this implies

$$\infty > \sum_{j=1}^{\infty} \pi(j) = -\pi(0)(\beta + 1)/\beta$$

so, since $\beta > -1$, we must have $\beta < 0$. Conversely, if $\beta < 0$, and we take

$$\pi(0) = -\beta$$

then the same summation (10.22) indicates that the invariant measure π is finite.

Thus we have

Proposition 10.3.1 *The chain $\mathbf{N}^*$ is positive if and only if the increment distribution satisfies $\beta = \sum j q_j < 1$.*

This same type of direct calculation can be carried out for any so called "skip-free" chain with $P(i, j) = 0$ for $j < i - 1$, such as the forward recurrence time chain above. For other chains it can be far less easy to get a direct approach to the invariant measure through the invariant equations, and we turn to the representation in (10.11) for our results.

10.3.3 The number in a GI/M/1 queue

We illustrate the use of the structural result in giving a novel interpretation of an old result for the specific random walk on a half line $\mathbf{N}$ corresponding to the number in a GI/M/1 queue.

Recall from Section 3.3.3 that $\mathbf{N}$ has increment distribution concentrated on the integers $\{\ldots,-1,0,1\}$ giving the transition probability matrix of the form

$$P = \begin{pmatrix} \sum_1^\infty p_i & p_0 & & & \\ \sum_2^\infty p_i & p_1 & p_0 & & 0 \\ \sum_3^\infty p_i & p_2 & p_1 & p_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

where $p_i = \mathsf{P}(Z = 1-i)$. The chain $\mathbf{N}$ is ψ-irreducible if $p_0 + p_1 < 1$, and irreducible if $p_0 > 0$ also. Assume these inequalities hold, and let $\{0\} = \boldsymbol{\alpha}$ be our atom.

To investigate the existence of an invariant measure for $\mathbf{N}$, we know from Theorem 10.2.1 that we should look at the quantities ${}_\alpha P^n(\boldsymbol{\alpha}, j)$.

Write $[k] = \{0,\ldots,k\}$. Because the chain can only move up one step at a time, so the last visit to $[k]$ is at k itself, we have on decomposing over the last visit to $[k]$, for $k \geq 1$

$$ {}_\alpha P^n(\boldsymbol{\alpha}, k+1) = \sum_{r=1}^{n} {}_\alpha P^r(\boldsymbol{\alpha}, k)\, {}_{[k]}P^{n-r}(k, k+1). \tag{10.23}$$

Now the translation invariance property of P implies that for $j > k$

$$ {}_{[k]}P^r(k, j) = {}_\alpha P^r(\boldsymbol{\alpha}, j-k). \tag{10.24}$$

Thus, summing (10.23) from 1 to ∞ gives

$$\begin{aligned} \sum_{n=1}^{\infty} {}_\alpha P^n(\boldsymbol{\alpha}, k+1) &= \Big[\sum_{n=1}^{\infty} {}_\alpha P^n(\boldsymbol{\alpha}, k)\Big] \Big[\sum_{n=1}^{\infty} {}_{[k]}P^n(k, k+1)\Big] \\ &= \Big[\sum_{n=1}^{\infty} {}_\alpha P^n(\boldsymbol{\alpha}, k)\Big] \Big[\sum_{n=1}^{\infty} {}_\alpha P^n(\boldsymbol{\alpha}, 1)\Big]. \end{aligned}$$

Using the form (10.11) of μ_α°, we have now shown that

$$\mu_\alpha^\circ(k+1) = \mu_\alpha^\circ(k)\mu_\alpha^\circ(1),$$

and so the minimal invariant measure satisfies

$$\mu_\alpha^\circ(k) = s_\alpha^k \tag{10.25}$$

where $s_\alpha = \mu_\alpha^\circ(1)$.

The chain then has an invariant probability measure if and only if we can find $s_\alpha < 1$ for which the measure μ_α° defined by the geometric form (10.25) is a solution to the subinvariant equations for P: otherwise the minimal subinvariant measure is not summable.

We can go further and identify these two cases in terms of the underlying parameters p_j. Consider the second (that is, the $k = 1$) invariant equation

$$\mu^\circ_\alpha(1) = \sum \mu^\circ_\alpha(k) P(k,1).$$

This shows that s_α must be a solution to

$$s = \sum_0^\infty p_j s^j, \tag{10.26}$$

and since μ°_α is minimal it must the smallest solution to (10.26). As is well-known, there are two cases to consider: since the function of s on the right hand side of (10.26) is strictly convex, a solution $s \in (0,1)$ exists if and only if

$$\sum_0^\infty j p_j > 1,$$

whilst if $\sum_j j\, p_j \le 1$ then the minimal solution to (10.26) is $s_\alpha = 1$.

One can then verify directly that in each of these cases μ°_α solves all of the invariant equations, as required. In particular, if $\sum_j j\, p_j = 1$ so that the chain is recurrent from the remarks following Proposition 9.1.2, the unique invariant measure is $\mu_\alpha(x) \equiv 1, x \in \mathsf{X}$: note that in this case, in fact, the first invariant equation is exactly

$$1 = \sum_{j \ge 0} \sum_{n > j} p_n = \sum_j j\, p_j.$$

Hence for recurrent chains (those for which $\sum_j j\, p_j \ge 1$) we have shown

Proposition 10.3.2 *The unique subinvariant measure for* $\mathbf{N}$ *is given by* $\mu_\alpha(k) = s_\alpha^k$, *where* s_α *is the minimal solution to (10.26) in* $(0,1]$; *and* $\mathbf{N}$ *is positive recurrent if and only if* $\sum_j j\, p_j > 1$. □

The geometric form (10.25), as a "trial solution" to the equation (10.1), is often presented in an arbitrary way: the use of Theorem 10.2.1 motivates this solution, and also shows that s_α in (10.25) has an interpretation as the expected number of visits to state $k+1$ from state k, for any k.

10.4 The existence of π: ψ-irreducible chains

10.4.1 Invariant measures for recurrent chains

We prove in this section that a general recurrent ψ-irreducible chain has an invariant measure, using the Nummelin splitting technique.

First we show how subinvariant measures for the split chain correspond with subinvariant measures for $\boldsymbol{\Phi}$.

Proposition 10.4.1 *Suppose that* $\boldsymbol{\Phi}$ *is a strongly aperiodic Markov chain and let* $\check{\boldsymbol{\Phi}}$ *denote the split chain. Then*

(i) *If the measure* $\check{\pi}$ *is invariant for* $\check{\boldsymbol{\Phi}}$, *then the measure* π *on* $\mathcal{B}(\mathsf{X})$ *defined by*

$$\pi(A) = \check{\pi}(A_0 \cup A_1), \qquad A \in \mathcal{B}(\mathsf{X}), \tag{10.27}$$

is invariant for $\boldsymbol{\Phi}$, *and* $\check{\pi} = \pi^*$.

(ii) *If* μ *is any subinvariant measure for* $\boldsymbol{\Phi}$ *then* μ^* *is subinvariant for* $\check{\boldsymbol{\Phi}}$, *and if* μ *is invariant then so is* μ^*.

PROOF To prove (i) note that by (5.5), (5.6), and (5.7), we have that the measure $\check{P}(x_i, \cdot)$ is of the form $\mu^*_{x_i}$ for any $x_i \in \check{\mathsf{X}}$, where μ_{x_i} is a probability measure on X. By linearity of the splitting and invariance of $\check{\pi}$, for any $\check{A} \in \mathcal{B}(\check{\mathsf{X}})$,

$$\check{\pi}(\check{A}) = \int \check{\pi}(dx_i)\check{P}(x_i, \check{A}) = \int \check{\pi}(dx_i)\mu^*_{x_i}(\check{A}) = \Big(\int \check{\pi}(dx_i)\mu_{x_i}(\,\cdot\,)\Big)^*(\check{A})$$

Thus $\check{\pi} = \pi_0^*$, where $\pi_0 = \int \check{\pi}(dx_i)\mu_{x_i}(\,\cdot\,)$.

By (10.27) we have that $\pi(A) = \pi_0^*(A_0 \cup A_1) = \pi_0(A)$, so that in fact $\check{\pi} = \pi^*$. This proves one part of (i), and we now show that π is invariant for $\boldsymbol{\Phi}$. For any $A \in \mathcal{B}(\mathsf{X})$ we have by invariance of π^* and (5.10),

$$\pi(A) = \pi^*(A_0 \cup A_1) = \pi^*\check{P}\,(A_0 \cup A_1) = \Big(\pi P\Big)^*(A_0 \cup A_1) = \pi P\,(A)$$

which shows that π is invariant and completes the proof of (i).

The proof of (ii) also follows easily from (5.10): if the measure μ is subinvariant then

$$\mu^*\check{P} = (\mu P)^* \le \mu^*,$$

which establishes subinvariance of μ^*, and similarly, $\mu^*\check{P} = \mu^*$ if μ is strictly invariant. □

We can now give a simple proof of

Proposition 10.4.2 *If $\boldsymbol{\Phi}$ is recurrent and strongly aperiodic then $\boldsymbol{\Phi}$ admits a unique (up to constant multiples) subinvariant measure which is invariant.*

PROOF Assume that $\boldsymbol{\Phi}$ is strongly aperiodic, and split the chain as in Section 5.1.

If $\boldsymbol{\Phi}$ is recurrent then it follows from Proposition 8.2.2 that $\check{\boldsymbol{\Phi}}$ is also recurrent. We have from Theorem 10.2.1 that $\check{\boldsymbol{\Phi}}$ has a unique subinvariant measure $\check{\pi}$ which is invariant. Thus we have from Proposition 10.4.1 that $\boldsymbol{\Phi}$ also has an invariant measure.

The uniqueness is equally easy. If $\boldsymbol{\Phi}$ has another subinvariant measure μ, then by Proposition 10.4.1 the split measure μ^* is subinvariant for $\check{\boldsymbol{\Phi}}$, and since from Theorem 10.2.1, the invariant measure $\check{\pi}$ is unique (up to constant multiples) for $\check{\boldsymbol{\Phi}}$, we must have for some $c > 0$ that $\mu^* = c\check{\pi}$. By linearity this gives $\mu = c\pi$ as required. □

We can, quite easily, lift this result to the whole chain even in the case where we do not have strong aperiodicity by considering the resolvent chain, since the chain and the resolvent share the same invariant measures.

Theorem 10.4.3 *For any $\varepsilon \in (0,1)$, a measure π is invariant for the resolvent K_{a_ε} if and only if it is invariant for P.*

PROOF If π is invariant with respect to P then by (10.4) it is also invariant for K_a, for any sampling distribution a.

To see the converse, suppose that π satisfies $\pi K_{a_\varepsilon} = \pi$ for some $\varepsilon \in (0,1)$, and consider the chain of equalities

$$\pi P = (1-\varepsilon)\sum_{k=0}^{\infty} \varepsilon^k \pi P^{k+1}$$

$$\begin{aligned} &= (1-\varepsilon)\varepsilon^{-1}(\sum_{k=0}^{\infty}\varepsilon^k \pi P^k - \pi) \\ &= \varepsilon^{-1}(\pi K_{a_\varepsilon} - (1-\varepsilon)\pi) \\ &= \pi. \end{aligned}$$

□

This now gives us immediately

Theorem 10.4.4 *If $\boldsymbol{\Phi}$ is recurrent then $\boldsymbol{\Phi}$ has a unique (up to constant multiples) subinvariant measure which is invariant.*

Proof Using Theorem 5.2.3, we have that the K_{a_ε}-chain is strongly aperiodic, and from Theorem 8.2.4 we know that the K_{a_ε}-chain is recurrent. Let π be the unique invariant measure for the K_{a_ε}-chain, guaranteed from Proposition 10.4.2. From Theorem 10.4.3 π is also invariant for $\boldsymbol{\Phi}$.

Suppose that μ is subinvariant for $\boldsymbol{\Phi}$. Then by (10.7) we have that μ is also subinvariant for the K_{a_ε}-chain, and so there is a constant $c > 0$ such that $\mu = c\pi$. Hence we have shown that π is the unique (up to constant multiples) invariant measure for $\boldsymbol{\Phi}$. □

We may now equate positivity of $\boldsymbol{\Phi}$ to positivity for its skeletons as well as the resolvent chains.

Theorem 10.4.5 *Suppose that $\boldsymbol{\Phi}$ is ψ-irreducible and aperiodic. Then, for each m, a measure π is invariant for the m-skeleton if and only if it is invariant for $\boldsymbol{\Phi}$.*

Hence, under aperiodicity, the chain $\boldsymbol{\Phi}$ is positive if and only if each of the m-skeletons $\boldsymbol{\Phi}^m$ is positive.

Proof If π is invariant for $\boldsymbol{\Phi}$ then it is obviously invariant for $\boldsymbol{\Phi}^m$, by (10.4).

Conversely, if π_m is invariant for the m-skeleton then by aperiodicity the measure π_m is the unique invariant measure (up to constant multiples) for $\boldsymbol{\Phi}^m$. In this case write

$$\pi(A) = \frac{1}{m}\sum_{k=0}^{m-1}\int \pi_m(dw)P^k(w,A), \qquad A \in \mathcal{B}(\mathsf{X}).$$

From the P^m-invariance we have, using operator theoretic notation,

$$\pi P = \frac{1}{m}\sum_{k=0}^{m-1}\pi_m P^{k+1} = \pi$$

so that π is an invariant measure for P. Moreover, since π is invariant for P, it is also invariant for P^m from (10.4), and so by uniqueness of π_m, for some $c > 0$ we have $\pi = c\pi_m$. But as π is invariant for P^j for every j, we have from the definition that

$$\pi = c^{-1}\frac{1}{m}\sum_{k=0}^{m-1}\int \pi P^{k+1} = c^{-1}\pi$$

and so $\pi_m = \pi$. □

10.4.2 Minimal subinvariant measures

In order to use invariant measures for recurrent chains, we shall study in some detail the structure of the invariant measures we have now proved to exist in Theorem 10.2.1. We do this through the medium of subinvariant measures, and we note that, in this section at least, we do not need to assume any form of irreducibility. Our goal is essentially to give a more general version of Kac's Theorem.

Assume that μ is an arbitrary subinvariant measure, and let $A \in \mathcal{B}(\mathsf{X})$ be such that $0 < \mu(A) < \infty$. Define the measure μ_A° by

$$\mu_A^\circ(B) = \int_A \mu(dy) U_A(y,B), \quad B \in \mathcal{B}(\mathsf{X}). \tag{10.28}$$

Proposition 10.4.6 *The measure μ_A° is subinvariant, and minimal in the sense that $\mu(B) \geq \mu_A^\circ(B)$ for all $B \in \mathcal{B}(\mathsf{X})$.*

PROOF If μ is subinvariant, then we have first that

$$\mu(B) \geq \int_A \mu(dw) P(w,B);$$

assume inductively that $\mu(B) \geq \int_A \mu(dw) \sum_{m=1}^n {}_AP^m(w,B)$, for all B. Then, by subinvariance,

$$\begin{aligned} \mu(B) &\geq \int_{A^c} \left[\int_A \mu(dw) \sum_{m=1}^n {}_AP^m(w,dv) \right] P(v,B) + \int_A \mu(dw) P(w,B) \\ &= \int_A \mu(dw) \sum_{m=1}^{n+1} {}_AP^m(w,B). \end{aligned}$$

Hence the induction holds for all n, and taking $n \uparrow \infty$ shows that

$$\mu(B) \geq \int_A \mu(dw) U_A(w,B)$$

for all B. Now by this minimality of μ_A°

$$\begin{aligned} \mu_A^\circ(B) &= \int_A \mu(dw) P(w,B) + \int_A \mu(dw) \sum_{m=2}^\infty {}_AP^m(w,B) \\ &\geq \int_A \mu_A^\circ(dw) P(w,B) + \int_{A^c} [\int_A \mu(dw) \sum_{m=1}^\infty {}_AP^m(w,dv)] P(v,B) \\ &= \int_\mathsf{X} \mu_A^\circ(dw) P(w,B). \end{aligned}$$

Hence μ_A° is subinvariant also. □

Recall that we define $\overline{A} := \{x : L(x,A) > 0\}$. We now show that if the set A in the definition of μ_A° is Harris recurrent, the minimal subinvariant measure is in fact invariant and identical to μ itself on $\overline{A}$.

Theorem 10.4.7 *If $L(x,A) \equiv 1$ for μ-almost all $x \in A$, then we have*

(i) *$\mu(B) = \mu_A^\circ(B)$ for $B \subset \overline{A}$;*

(ii) *μ_A° is invariant and $\mu_A^\circ(\overline{A}^c) = 0$.*

PROOF **(i)** We first show that $\mu(B) = \mu_A^\circ(B)$ for $B \subseteq A$.

For any $B \subseteq A$, since $L(x, A) \equiv 1$ for μ-almost all $x \in A$, we have from minimality of μ_A°

$$\begin{aligned} \mu(A) &= \mu(B) + \mu(A \cap B^c) \\ &\geq \mu_A^\circ(B) + \mu_A^\circ(A \cap B^c) \\ &= \int_A \mu(dw) U_A(w, B) + \int_A \mu(dw) U_A(w, A \cap B^c) \\ &= \int_A \mu(dw) U_A(w, A) = \mu(A). \end{aligned} \tag{10.29}$$

Hence, the inequality $\mu(B) \geq \mu_A^\circ(B)$ must be an equality for all $B \subseteq A$. Thus the measure μ satisfies

$$\mu(B) = \int_A \mu(dw) U_A(w, B) \tag{10.30}$$

whenever $B \subseteq A$.

We now use (10.30) to prove invariance of μ_A°. For any $B \in \mathcal{B}(\mathsf{X})$,

$$\begin{aligned} \int_{\mathsf{X}} \mu_A^\circ(dy) P(y, B) &= \int_A \mu_A^\circ(dy) P(y, B) \\ &\quad + \int_{A^c} \left[\int_A \mu_A^\circ(dw) U_A(w, dy) \right] P(y, B) \\ &= \int_A \mu_A^\circ(dy) \left[P(y, B) + \sum_2^\infty {}_A P^n(y, B) \right] \\ &= \mu_A^\circ(B) \end{aligned} \tag{10.31}$$

and so μ_A° is invariant for $\boldsymbol{\Phi}$. It follows by definition that $\mu_A^\circ(\overline{A}^c) = 0$, so (ii) is proved.

We now prove (i) by contradiction. Suppose that $B \subseteq \overline{A}$ with $\mu(B) > \mu_A^\circ(B)$. Then we have from invariance of the resolvent chain in Proposition 10.4.3 and minimality of μ_A°, and the assumption that $K_{a_\varepsilon}(x, A) > 0$ for $x \in B$,

$$\mu(A) \geq \int_{\mathsf{X}} \mu(dy) K_{a_\varepsilon}(y, A) > \int_{\mathsf{X}} \mu_A^\circ(dy) K_{a_\varepsilon}(y, A) = \mu_A^\circ(A) = \mu(A),$$

and we thus have a contradiction. □

An interesting consequence of this approach is the identity (10.30). This has the following interpretation. Assume A is Harris recurrent, and define the *process on* A, denoted by $\boldsymbol{\Phi}^A = \{\Phi_n^A\}$, by starting with $\Phi_0 = x \in A$, then setting Φ_1^A as the value of $\boldsymbol{\Phi}$ at the next visit to A, and so on. Since return to A is sure for Harris recurrent sets, this is well defined.

Formally, $\boldsymbol{\Phi}^A$ is actually constructed from the transition law

$$U_A(x, B) = \sum_{n=1}^\infty {}_A P^n(x, B) = \mathsf{P}_x\{\Phi_{\tau_A} \in B\},$$

$B \subseteq A$, $B \in \mathcal{B}(\mathsf{X})$. Theorem 10.4.7 thus states that for a Harris recurrent set A, any subinvariant measure restricted to A is actually invariant for the process on A.

One can also go in the reverse direction, starting off with an invariant measure for the process on A. The following result is proved using the same calculations used in (10.31):

Proposition 10.4.8 *Suppose that ν is an invariant probability measure supported on the set A with*

$$\int_A \nu(dx) U_A(x, B) = \nu(B), \qquad B \subseteq A.$$

Then the measure ν° defined as

$$\nu^\circ(B) := \int_A \nu(dx) U_A(x, B) \qquad B \in \mathcal{B}(\mathsf{X})$$

is invariant for $\boldsymbol{\Phi}$. □

10.4.3 The structure of π for recurrent chains

These preliminaries lead to the following key result.

Theorem 10.4.9 *Suppose $\boldsymbol{\Phi}$ is recurrent. Then the unique (up to constant multiples) invariant measure π for $\boldsymbol{\Phi}$ is equivalent to ψ and satisfies for any $A \in \mathcal{B}^+(\mathsf{X})$, $B \in \mathcal{B}(\mathsf{X})$,*

$$\begin{aligned} \pi(B) &= \int_A \pi(dy) U_A(y, B) \\ &= \int_A \pi(dy) \mathsf{E}_y\Big[\sum_{k=1}^{\tau_A} \mathbb{1}\{\Phi_k \in B\}\Big] \\ &= \int_A \pi(dy) \mathsf{E}_y\Big[\sum_{k=0}^{\tau_A - 1} \mathbb{1}\{\Phi_k \in B\}\Big] \end{aligned} \tag{10.32}$$

Proof The construction in Theorem 10.2.1 ensures that the invariant measure π exists. Hence from Theorem 10.4.7 we see that $\pi = \pi_A^\circ$ for any Harris recurrent set A, and π then satisfies the first equality in (10.32) by construction. The second equality is just the definition of U_A. To see the third equality,

$$\int_A \pi(dy) \mathsf{E}_y\Big[\sum_{k=1}^{\tau_A} \mathbb{1}\{\Phi_k \in B\}\Big] = \int_A \pi(dy) \mathsf{E}_y\Big[\sum_{k=0}^{\tau_A - 1} \mathbb{1}\{\Phi_k \in B\}\Big],$$

apply (10.30) which implies that

$$\int_A \pi(dy) \mathsf{E}_y[\mathbb{1}\{\Phi_{\tau_A} \in B\}] = \int_A \pi(dy) \mathsf{E}_y[\mathbb{1}\{\Phi_0 \in B\}].$$

We finally prove that $\pi \cong \psi$. From Proposition 10.1.2 we need only show that if $\psi(B) = 0$ then also $\pi(B) = 0$. But since $\psi(\bar{B}) = 0$, we have that $B^0 \in \mathcal{B}^+(\mathsf{X})$, and so from the representation (10.32),

$$\pi(B) = \int_{B^0} \pi(dy) U_{B_0}(y, B) = 0,$$

which is the required result. □

The interpretation of (10.32) is this: for a fixed set $A \in \mathcal{B}^+(\mathsf{X})$, the invariant measure $\pi(B)$ is proportional to the amount of time spent in B between visits to A, provided the chain starts in A with the distribution π_A which is invariant for the chain $\boldsymbol{\Phi}^A$ on A.

When A is a single point, $\boldsymbol{\alpha}$, with $\pi(\boldsymbol{\alpha}) > 0$ then each visit to $\boldsymbol{\alpha}$ occurs at $\boldsymbol{\alpha}$. The chain $\boldsymbol{\Phi}^\alpha$ is hence trivial, and its invariant measure π_α is just δ_α. The representation (10.32) then reduces to μ_α given in Theorem 10.2.1.

We will use these concepts systematically in building the asymptotic theory of positive chains in Chapter 13 and later work, and in Chapter 11 we develop a number of conditions equivalent to positivity through this representation of π. The next result is a foretaste of that work.

Theorem 10.4.10 *Suppose that Φ is ψ-irreducible, and let μ denote any subinvariant measure.*

(i) *The chain Φ is positive if and only if for one, and then every, set with $\mu(A) > 0$*

$$\int_A \mu(dy)\mathsf{E}_y[\tau_A] < \infty. \tag{10.33}$$

(ii) *The measure μ is finite and thus Φ is positive recurrent if for some petite set $C \in \mathcal{B}^+(\mathsf{X})$*

$$\sup_{y \in C} \mathsf{E}_y[\tau_C] < \infty. \tag{10.34}$$

The chain Φ is positive Harris if also

$$\mathsf{E}_x[\tau_C] < \infty, \qquad x \in \mathsf{X}. \tag{10.35}$$

PROOF The first result is a direct consequence of (10.28), since we have

$$\mu_A^\circ(\mathsf{X}) = \int_A \mu(dy) U_A(y, \mathsf{X}) = \int_A \mu(dy)\mathsf{E}_y[\tau_A];$$

if this is finite then μ_A° is finite and the chain is positive by definition. Conversely, if the chain is positive then by Theorem 10.4.9 we know that μ must be a finite invariant measure and (10.33) then holds for every A.

The second result now follows since we know from Proposition 10.1.2 that $\mu(C) < \infty$ for petite C; and hence we have positive recurrence from (10.34) and (i), whilst the chain is also Harris if (10.35) holds from the criterion in Theorem 9.1.7. □

In Chapter 11 we find a variety of usable and useful conditions for (10.34) and (10.35) to hold, based on a drift approach which strengthens those in Chapter 8.

10.5 Invariant Measures: General Models

The constructive approach to the existence of invariant measures which we have featured so far enables us either to develop results on invariant measures for a number of models, based on the representation in (10.32), or to interpret the invariant measure probabilistically once we have determined it by some other means.

We now give a variety of examples of this.

10.5.1 Random walk

Consider the random walk on the line, with increment measure Γ, as defined in (RW1). Then by Fubini's Theorem and the translation invariance of μ^{Leb} we have for any $A \in \mathcal{B}(\mathsf{X})$

$$\begin{aligned}\int_{\mathbb{R}} \mu^{\mathrm{Leb}}(dy)P(y,A) &= \int_{\mathbb{R}} \mu^{\mathrm{Leb}}(dy)\Gamma(A-y)\\ &= \int_{\mathbb{R}} \mu^{\mathrm{Leb}}(dy)\int_{\mathbb{R}} \mathbb{1}_{A-y}(x)\Gamma(dx)\\ &= \int_{\mathbb{R}} \Gamma(dx)\int_{\mathbb{R}} \mathbb{1}_{A-x}(y)\mu^{\mathrm{Leb}}(dy) \qquad (10.36)\\ &= \mu^{\mathrm{Leb}}(A)\end{aligned}$$

since $\Gamma(\mathbb{R}) = 1$. We have already used this formula in (6.8): here it shows that Lebesgue measure is invariant for unrestricted random walk in either the transient or the recurrent case.

Since Lebesgue measure on $\mathbb{R}$ is infinite, we immediately have from Theorem 10.4.9 that there is no finite invariant measure for this chain: this proves

Proposition 10.5.1 *The random walk on $\mathbb{R}$ is never positive recurrent.* □

If we put this together with the results in Section 9.5, then we have that when the mean β of the increment distribution is zero, then the chain is null recurrent.

Finally, we note that this is one case where the interpretation in (10.32) can be expressed in another way. We have, as an immediate consequence of this interpretation

Proposition 10.5.2 *Suppose $\boldsymbol{\Phi}$ is a random walk on $\mathbb{R}$, with spread out increment measure Γ having zero mean and finite variance.*

Let A be any bounded set in $\mathbb{R}$ with $\mu^{\mathrm{Leb}}(A) > 0$, and let the initial distribution of Φ_0 be the uniform distribution on A. If we let $N_A(B)$ denote the mean number of visits to a set B prior to return to A then for any two bounded sets B, C with $\mu^{\mathrm{Leb}}(C) > 0$ we have

$$\mathsf{E}[N_A(B)]/\mathsf{E}[N_A(C)] = \mu^{\mathrm{Leb}}(B)/\mu^{\mathrm{Leb}}(C).$$

PROOF Under the given conditions on Γ we have from Proposition 9.4.5 that the chain is non-evanescent, and hence recurrent.

Using (10.36) we have that the unique invariant measure with $\pi(A) = 1$ is $\pi = \mu^{\mathrm{Leb}}/\pi(A)$, and then the result follows from the form (10.32) of π. □

10.5.2 Forward recurrence time chains

Let us consider the forward recurrence time chain $\mathbf{V}_\delta^+$ defined in Section 3.5 for a renewal process on $\mathbb{R}_+$. For any fixed δ consider the expected number of visits to an interval strictly outside $[0,\delta]$. Exactly as we reasoned in the discrete time case studied in Section 10.3, we have

$$F[y,\infty)dy \le U_{[0,\delta]}(x,dy) \le F[y-\delta,\infty)dy.$$

Thus, if π_δ is to be the invariant probability measure for $\mathbf{V}_\delta^+$, by using the normalized version of the representation (10.32)

$$\frac{F[y,\infty)dy}{[\int_0^\infty F(w,\infty)dw]} \le \pi_\delta(dy) \le \frac{F[y-\delta,\infty)dy}{[\int_\delta^\infty F(w,\infty)dw]}.$$

Now we use uniqueness of the invariant measure to note that, since the chain $\mathbf{V}_\delta^+$ is the "two-step" chain for the chain $\mathbf{V}_{\delta/2}^+$, the invariant measures π_δ and $\pi_{\delta/2}$ must coincide. Thus letting δ go to zero through the values $\delta/2^n$ we find that for any δ the invariant measure is given by

$$\pi_\delta(dy) = m^{-1}F[y,\infty)dy \tag{10.37}$$

where $m = \int_0^\infty tF(dt)$; and π_δ is a probability measure provided $m < \infty$.

By direct integration it is also straightforward to show that this is indeed the invariant measure for $\mathbf{V}_\delta^+$.

This form of the invariant measure thus reinforces the fact that the quantity $F[y,\infty)dy$ is the expected amount of time spent in the infinitesimal set dy on each excursion from the point $\{0\}$, even though in the discretized chain $\mathbf{V}_\delta^+$ the point $\{0\}$ is never actually reached.

10.5.3 Ladder chains and GI/G/1 queues

General ladder chains We will now turn to a more complex structure and see how far the representation of the invariant measure enables us to carry the analysis.

Recall from Section 3.5.4 the Markov chain constructed on $\mathbb{Z}_+ \times \mathbb{R}$ to analyze the GI/G/1 queue, with the "ladder-invariant" transition kernel

$$\begin{array}{rcl} P(i,x;j\times A) &=& 0, \quad j > i+1 \\ P(i,x;j\times A) &=& \Lambda_{i-j+1}(x,A), \quad j = 1,\dots,i+1 \\ P(i,x;0\times A) &=& \Lambda_i^*(x,A). \end{array} \tag{10.38}$$

Let us consider the general chain defined by (10.38), where we can treat x and A as general points in and subsets of X, so that the chain $\boldsymbol{\Phi}$ now moves on a ladder whose (countable number of) rungs are general in nature. In the special case of the GI/G/1 model the results specialize to the situation where $\mathsf{X} = \mathbb{R}_+$, and there are many countable models where the rungs are actually finite and matrix methods are used to achieve the following results.

Using the representation of π, it is possible to construct an invariant measure for this chain in an explicit way; this then gives the structure of the invariant measure for the GI/G/1 queue also.

Since we are interested in the structure of the invariant probability measure we make the assumption in this section that the chain defined by (10.38) is positive Harris and $\psi([0]) > 0$, where $[0] := \{0 \times \mathsf{X}\}$ is the bottom "rung" of the ladder. We shall explore conditions for this to hold in Chapter 19.

Our assumption ensures we can reach the bottom of the ladder with probability one. Let us denote by π_0 the invariant probability measure for the process on $[0]$, so that π_0 can be thought of as a measure on $\mathcal{B}(\mathsf{X})$.

Our goal will be to prove that the structure of the invariant measure for $\boldsymbol{\Phi}$ is an "operator-geometric" one, mimicking the structure of the invariant measure developed in Section 10.3 for skip-free random walk on the integers.

Theorem 10.5.3 *The invariant measure π for $\boldsymbol{\Phi}$ is given by*

$$\pi(k\times A) = \int_{\mathsf{X}} \pi_0(dy) S^k(y,A) \tag{10.39}$$

where

$$S^k(y,A) = \int_{\mathsf{X}} S(y,dz)S^{k-1}(z,A) \tag{10.40}$$

for a kernel S which is the minimal solution of the operator equation

$$S(y,B) = \sum_{k=0}^{\infty} \int_{\mathsf{X}} S^k(y,dz)\Lambda_k(z,B), \qquad x \in \mathsf{X}, B \in \mathcal{B}(\mathsf{X}). \tag{10.41}$$

PROOF Using the structural result (10.32) we have

$$\pi(k \times A) = \int_{[0]} \pi_0(dy) U_{[0]}(0,y;k \times B) \tag{10.42}$$

so that if we write

$$S^{(k)}(y,A) := U_{[0]}(0,y;k \times A) \tag{10.43}$$

we have by definition

$$\pi(k \times A) = \int_{[0]} \pi_0(dy) S^{(k)}(y,A). \tag{10.44}$$

Now if we define the set $[n] = \{0,1,\dots,n\} \times \mathsf{X}$, by the fact that the chain is translation invariant above the zero level we have that the functions

$$U_{[n]}(n,y;(n+k) \times B) = U_{[0]}(0,y;k \times B) = S^{(k)}(y,A) \tag{10.45}$$

are independent of n. Using a last exit decomposition over visits to $[k]$, together with the "skip-free" property which ensures that the last visit to $[k]$ prior to reaching $(k+1) \times \mathsf{X}$ takes place at the level $k \times \mathsf{X}$, we find

$$\begin{aligned} &{}_{[0]}P^{\ell}(0,x;(k+1)\times A) \\ &= \textstyle\sum_{j=1}^{\ell-1} \int_{\mathsf{X}} {}_{[0]}P^j(0,x;k \times dy)\, {}_{[k]}P^{\ell-j}(k,y;(k+1) \times A) \\ &= \textstyle\sum_{j=1}^{\ell-1} \int_{\mathsf{X}} {}_{[0]}P^j(0,x;k \times dy)\, {}_{[0]}P^{\ell-j}(0,y;1 \times A) \end{aligned} \tag{10.46}$$

Summing over ℓ and using (10.45) shows that the operators $S^{(k)}(y,A)$ have the geometric form in (10.40) as stated.

To see that the operator S satisfies (10.41), we decompose ${}_{[0]}P^n$ over the position at time $n-1$. By construction ${}_{[0]}P^1(0,x;1 \times B) = \Lambda_0(x,B)$, and for $n > 1$,

$${}_{[0]}P^n(0,x;1 \times B) = \sum_{k \geq 1} \int_{\mathsf{X}} {}_{[0]}P^{n-1}(0,x;k \times dy)\Lambda_k(y,B); \tag{10.47}$$

summing over n and using (10.40) gives the result (10.41).

To prove minimality of the solution S to (10.41), we first define, for $N \geq 1$, the partial sums

$$S_N(x;k \times B) := \sum_{j=1}^{N} {}_{[0]}P^j(0,x;k \times B) \tag{10.48}$$

so that as $N \to \infty, S_N(x;1 \times B) \to S(x;B)$.

Using (10.46) these partial sums also satisfy

$$S_{N-1}(x;k+1 \times B) \leq \int S_{N-1}(x;k \times dy) S_{N-1}(y;1 \times B)$$

so that

$$S_{N-1}(x;k+1\times B)\le\int S^k_{N-1}(x;1\times dy)S_{N-1}(y;1\times B). \tag{10.49}$$

Moreover from (10.47)

$$S_N(x;1\times B)=\Lambda_0(x,B)+\sum_{k\ge1}\int_{\mathsf X}S_{N-1}(x;k\times dy)\Lambda_k(y,B). \tag{10.50}$$

Substituting from (10.49) in (10.50) shows that

$$S_N(x;1,B)\le\sum_k\int_{\mathsf X}S^k_{N-1}(x;1,dy)\Lambda_k(y,B). \tag{10.51}$$

Now let S^* be any other solution of (10.41). Notice that $S_1(x;1\times B)=\Lambda_0(x,B)\le S^*(x,B)$, from (10.41). Assume inductively that $S_{N-1}(x;1\times B)\le S^*(x,B)$ for all x,B: then we have from (10.51) that

$$S_N(x;1\times B)\le\sum_k\int_{\mathsf X}[S^*]^k(x,dy)\Lambda_k(y,B)=S^*(x,B). \tag{10.52}$$

Taking limits as $N\to\infty$ gives $S(x,B)\le S^*(x,B)$ for all x,B as required. □

This result is a generalized version of (10.25) and (10.26), where the "rungs" on the ladder were singletons.

The GI/G/1 queue Note that in the ladder processes above, the returns to the bottom rung of the ladder, governed by the kernels Λ^*_i in (10.38), only appear in the representation (10.39) implicitly, through the form of the invariant measure π_0 for the process on the set [0].

In particular cases it is of course of critical importance to identify this component of the invariant measure also. In the case of a singleton rung, this is trivial since the rung is an atom. This gives the explicit form in (10.25) and (10.26).

We have seen in Section 3.5 that the general ladder chain is a model for the GI/G/1 queue, if we make the particular choice of

$$\Phi_n=(N_n,R_n),\quad n\ge1$$

where N_n is the number of customers at T'_n- and R_n is the residual service time at T'_n+. In this case the representation of $\pi_{[0]}$ can also be made explicit.

For the GI/G/1 chain we have that the chain on [0] has the distribution of R_n at a time-point $\{T'_n+\}$ where there were no customers at $\{T'_n-\}$: so at these time points R_n has precisely the distribution of the service brought by the customer arriving at T'_n, namely H.

So in this case we have that the process on [0], provided [0] is recurrent, is a process of i.i.d random variables with distribution H, and thus is very clearly positive Harris with invariant probability H.

Theorem 10.5.3 then gives us

Theorem 10.5.4 *The ladder chain Φ describing the GI/G/1 queue has an invariant probability if and only if the measure π given by*

$$\pi(k\times A)=\int_{\mathsf X}H(dy)S^k(y,A) \tag{10.53}$$

is a finite measure, where S is the minimal solution of the operator equation

$$S(y,B) = \sum_{k=0}^{\infty} \int_{\mathsf{X}} S^k(y,dz) \Lambda_k(z,B), \qquad x \in \mathsf{X}, B \in \mathcal{B}(\mathsf{X}). \tag{10.54}$$

In this case π suitably normalized is the unique invariant probability measure for $\boldsymbol{\Phi}$.

PROOF Using the proof of Theorem 10.5.3 we have that π is the minimal subinvariant measure for the GI/G/1 queue, and the result is then obvious. □

10.5.4 Linear state space models

We now consider briefly a chain where we utilize the property (10.5) to develop the form of the invariant measure. We will return in much more detail to this approach in Chapter 12.

We have seen in (10.5) that limiting distributions provide invariant probability measures for Markov chains, provided such limits exist. The linear model has a structure which makes it easy to construct an invariant probability through this route, rather than through the minimal measure construction above.

Suppose that (LSS1) and (LSS2) are satisfied, and observe that since $\mathbf{W}$ is assumed i.i.d. we have for each initial condition $X_0 = x_0 \in \mathbb{R}^n$,

$$\begin{aligned} X_k &= F^k x_0 + \sum_{i=0}^{k-1} F^i G W_{k-i} \\ &\sim F^k x_0 + \sum_{i=0}^{k-1} F^i G W_i. \end{aligned}$$

This says that for any continuous, bounded function $g\colon \mathbb{R}^n \to \mathbb{R}$,

$$P^k g\,(x_0) = \mathsf{E}_{x_0}[g(X_k)] = \mathsf{E}[g(F^k x_0 + \sum_{i=0}^{k-1} F^i G W_i)].$$

Under the additional hypothesis that the eigenvalue condition (LSS5) holds, it follows from Lemma 6.3.4 that $F^i \to 0$ as $i \to \infty$ at a geometric rate. Since W has a finite mean then it follows from Fubini's Theorem that the sum

$$X_\infty := \sum_{i=0}^{\infty} F^i G W_i$$

converges absolutely, with $\mathsf{E}[|X_\infty|] \le \mathsf{E}[|W|] \sum_{i=0}^{\infty} \|F^i G\| < \infty$, with $\|\cdot\|$ an appropriate matrix norm. Hence by the Dominated Convergence Theorem, and the assumption that g is continuous,

$$\lim_{k\to\infty} P^k g\,(x_0) = \mathsf{E}[g(X_\infty)].$$

Let us write π_∞ for the distribution of X_∞. Then π_∞ is an invariant probability. For take g bounded and continuous as before, so that using the Feller property for $\mathbf{X}$ in Chapter 6 we have that Pg is continuous. For such a function g

$$\begin{aligned}\pi_\infty(Pg) = \mathsf{E}[Pg(X_\infty)] &= \lim_{k\to\infty} P^k(x_0, Pg) \\ &= \lim_{k\to\infty} P^{k+1}g\,(x_0) \\ &= \mathsf{E}[g(X_\infty)] = \pi_\infty(g).\end{aligned}$$

Since π is determined by its values on continuous bounded functions, this proves that π is invariant.

In the Gaussian case (LSS3) we can express the invariant probability more explicitly. In this case X_∞ itself is Gaussian with mean zero and covariance

$$\mathsf{E}[X_\infty X_\infty^\top] = \sum_{k=0}^{\infty} F^i GG^\top F^{i\top}$$

That is, $\pi = N(0,\Sigma)$ where Σ is equal to the controllability grammian for the linear state space model, defined in (4.17).

The covariance matrix Σ is full rank if and only if the controllability condition (LCM3) holds, and in this case, for any k greater than or equal to the dimension of the state space, $P^k(x,dy)$ possesses the density $p_k(x,y)dy$ given in (4.18). It follows immediately that when (LCM3) holds, the probability π possesses the density p on $\mathbb{R}^n$ given by

$$p(y) = (2\pi|\Sigma|)^{-n/2}\exp\{-\tfrac{1}{2}y^T\Sigma^{-1}y\}, \tag{10.55}$$

while if the controllability condition (LCM3) fails to hold then the invariant probability is concentrated on the controllable subspace $\mathsf{X}_0 = \mathcal{R}(\Sigma) \subset \mathsf{X}$ and is hence singular with respect to Lebesgue measure.

10.6 Commentary

The approach to positivity given here is by no means standard. It is much more common, especially with countable spaces, to classify chains either through the behavior of the sequence P^n, with null chains being those for which $P^n(x,A) \to 0$ for, say, petite sets A and all x, and positive chains being those for which such limits are not always zero; a limiting argument such as that in (10.5), which we have illustrated in Section 10.5.4, then shows the existence of π in the positive case.

Alternatively, positivity is often defined through the behavior of the expected return times to petite or other suitable sets.

We will show in Chapter 11 and Chapter 18 that even on a general space all of these approaches are identical. Our view is that the invariant measure approach is much more straightforward to understand than the P^n approach, and since one can now develop through the splitting technique a technically simple set of results this gives an appropriate classification of recurrent chains.

The existence of invariant probability measures has been a central topic of Markov chain theory since the inception of the subject. Doob [68] and Orey [208] give some good background. The approach to countable recurrent chains through last exit probabilities as in Theorem 10.2.1 is due to Derman [61], and has not changed much since, although the uniqueness proofs we give owe something to Vere-Jones [284]. The construction of π given here is of course one of our first serious uses of the splitting method of Nummelin [200]; for strongly aperiodic chains the result is also derived in Athreya and Ney [12]. The fact that one identifies the actual structure of π in

Theorem 10.4.9 will also be of great use, and Kac's Theorem [114] provides a valuable insight into the probabilistic difference between positive and null chains: this is pursued in the next chapter in considerably more detail.

Before the splitting technique, verifying conditions for the existence of π had appeared to be a deep and rather difficult task. It was recognized in the relatively early development of general state space Markov chains that one could prove the existence of an invariant measure for $\boldsymbol{\Phi}$ from the existence of an invariant probability measure for the "process on A". The approach pioneered by Harris [95] for finding the latter involves using deeper limit theorems for the "process on A" in the special case where A is a ν_n-small set, (called a C-set in Orey [208]) if $a_n = \delta_n$ and $\nu_n\{A\} > 0$. In this methodology, it is first shown that limiting probabilities for the process on A exist, and the existence of such limits then provides an invariant measure for the process on A: by the construction described in this chapter this can be lifted to an invariant measure for the whole chain. Orey [208] remains an excellent exposition of the development of this approach.

This "process on A" method is still the only one available without some regeneration, and we will develop this further in a topological setting in Chapter 12, using many of the constructions above.

We have shown that invariant measures exist without using such deep asymptotic properties of the chain, indicating that the existence and uniqueness of such measures is in fact a result requiring less of the detailed structure of the chain.

The minimality approach of Section 10.4.2 of course would give another route to Theorem 10.4.4, provided we had some method of proving that a "starting" subinvariant measure existed. There is one such approach, which avoids splitting and remains conceptually simple. This involves using the kernels

$$U^{(r)}(x, A) = \sum_{n=1}^{\infty} P^n(x, A) r^n \geq r \int_{\mathsf{X}} U^{(r)}(x, dy) P(y, A) \tag{10.56}$$

defined for $0 < r < 1$. One can then define a subinvariant measure for $\boldsymbol{\Phi}$ as a limit

$$\lim_{r \uparrow 1} \pi_r(\,\cdot\,) := \lim_{r \uparrow 1} [\int_C \nu_n(dy) U^{(r)}(y, \,\cdot\,)] / [\int_C \nu_n(dy) U^{(r)}(y, C)]$$

where C is a ν_n-small set. The key is the observation that this limit gives a non-trivial σ-finite measure due to the inequalities

$$M_j \geq \pi_r(\bar{C}(j)) \tag{10.57}$$

and

$$\pi_r(A) \geq r^n \nu_n(A), \qquad A \in \mathcal{B}(\mathsf{X}), \tag{10.58}$$

which are valid for all r large enough. Details of this construction are in Arjas and Nummelin [8], as is a neat alternative proof of uniqueness.

All of these approaches are now superseded by the splitting approach, but of course only when the chain is ψ-irreducible. If this is not the case then the existence of an invariant measure is not simple. The methods of Section 10.4.2, which are based on Tweedie [280], do not use irreducibility, and in conjunction with those in Chapter 12 they give some ways of establishing uniqueness and structure for the invariant measures from limiting operations, as illustrated in Section 10.5.4.

The general question of existence and, more particularly, uniqueness of invariant measures for non-irreducible chains remains open at this stage of theoretical development.

The invariance of Lebesgue measure for random walk is well known, as is the form (10.37) for models in renewal theory. The invariant measures for queues are derived directly in [40], but the motivation through the minimal measure of the geometric form is not standard. The extension to the operator-geometric form for ladder chains is in [277], and in the case where the rungs are finite, the development and applications are given by Neuts [194, 195].

The linear model is analyzed in Snyders [250] using ideas from control theory, and the more detailed analysis given there allows a generalization of the construction given in Section 10.5.4. Essentially, if the noise does not enter the "unstable" region of the state space then the stability condition on the driving matrix F can be slightly weakened.