13

Ergodicity

In Part II we developed the ideas of stability largely in terms of recurrence structures. Our concern was with the way in which the chain returned to the "center" of the space, how sure we could be that this would happen, and whether it might happen in a finite mean time.

Part III is devoted to the perhaps even more important, and certainly deeper, concepts of the chain "settling down", or converging, to a stable or stationary regime.

In our heuristic introduction to the various possible ideas of stability in Section 1.3, such convergence was presented as a fundamental idea, related in the dynamical systems and deterministic contexts to asymptotic stability. We noted briefly, in (10.5) in Chapter 10, that the existence of a finite invariant measure was a necessary condition for such a stationary regime to exist as a limit. In Chapter 12 we explored in much greater detail the way in which convergence of P^n to a limit, on topological spaces, leads to the existence of invariant measures.

In this chapter we begin a systematic approach to this question from the other side. Given the existence of π, when do the n-step transition probabilities converge in a suitable way to π?

We will prove that for positive recurrent ψ-irreducible chains, such limiting behavior takes place with no topological assumptions, and moreover the limits are achieved in a much stronger way than under the tightness assumptions in the topological context. The Aperiodic Ergodic Theorem, which unifies the various definitions of positivity, summarizes this asymptotic theory. It is undoubtedly the outstanding achievement in the general theory of ψ-irreducible Markov chains, even though we shall prove some considerably stronger variations in the next two chapters.

Theorem 13.0.1 (Aperiodic Ergodic Theorem) *Suppose that $\boldsymbol{\Phi}$ is an aperiodic Harris recurrent chain, with invariant measure π. The following are equivalent:*

(i) *The chain is positive Harris: that is, the unique invariant measure π is finite.*

(ii) *There exists some ν-small set $C \in \mathcal{B}^+(\mathsf{X})$ and some $P^\infty(C) > 0$ such that as $n \to \infty$, for all $x \in C$*

$$P^n(x, C) \to P^\infty(C). \tag{13.1}$$

(iii) *There exists some regular set in $\mathcal{B}^+(\mathsf{X})$: equivalently, there is a petite set $C \in \mathcal{B}(\mathsf{X})$ such that*

$$\sup_{x \in C} \mathsf{E}_x[\tau_C] < \infty. \tag{13.2}$$

(iv) *There exists some petite set C, some $b < \infty$ and a non-negative function V finite at some one $x_0 \in \mathsf{X}$, satisfying*

$$\Delta V(x) := PV(x) - V(x) \le -1 + b\mathbb{1}_C(x), \qquad x \in \mathsf{X}. \tag{13.3}$$

Any of these conditions is equivalent to the existence of a unique invariant probability measure π such that for every initial condition $x \in \mathsf{X}$,

$$\sup_{A\in\mathcal{B}(\mathsf{X})} |P^n(x,A) - \pi(A)| \to 0 \tag{13.4}$$

as $n \to \infty$, and moreover for any regular initial distributions λ, μ,

$$\sum_{n=1}^{\infty} \int\int \lambda(dx)\mu(dy) \sup_{A\in\mathcal{B}(\mathsf{X})} |P^n(x,A) - P^n(y,A)| < \infty. \tag{13.5}$$

PROOF That $\pi(\mathsf{X}) < \infty$ in (i) is equivalent to the finiteness of hitting times as in (iii) and the existence of a mean drift test function in (iv) is merely a restatement of the overview Theorem 11.0.1 in Chapter 11.

The fact that any of these positive recurrence conditions imply the uniform convergence over all sets A from all starting points x as in (13.4) is of course the main conclusion of this theorem, and is finally shown in Theorem 13.3.3.

That (ii) holds from (13.4) is obviously trivial by dominated convergence. The cycle is completed by the implication that (ii) implies (13.4), which is in Theorem 13.3.5.

The extension from convergence to summability provided the initial measures are regular is given in Theorem 13.4.4. Conditions under which π itself is regular are also in Section 13.4.2. □

There are four ideas which should be born in mind as we embark on this third part of the book, especially when coming from a countable space background. The first two involve the types of limit theorems we shall address; the third involves the method of proof of these theorems; and the fourth involves the nomenclature we shall use.

Modes of Convergence The first is that we will be considering, in this and the next three chapters, convergence of a chain in terms of its transition probabilities. Although it is important also to consider convergence of a chain along its sample paths, leading to strong laws, or of normalized variables leading to central limit theorems and associated results, we do not turn to this until Chapter 17.

This is in contrast to the traditional approach in the countable state space case. Typically, there, the search is for conditions under which there exist pointwise limits of the form

$$\lim_{n\to\infty} |P^n(x,y) - \pi(y)| = 0; \tag{13.6}$$

but the results we derive are related to the signed measure $(P^n - \pi)$, and so concern not merely such pointwise or even setwise convergence, but a more global convergence in terms of the total variation norm.

Total Variation Norm

If μ is a signed measure on $\mathcal{B}(\mathsf{X})$ then the *total variation norm* $\|\mu\|$ is defined as

$$\|\mu\| := \sup_{f:|f|\leq 1} |\mu(f)| = \sup_{A\in\mathcal{B}(\mathsf{X})} \mu(A) - \inf_{A\in\mathcal{B}(\mathsf{X})} \mu(A) \tag{13.7}$$

The key limit of interest to us in this chapter will be of the form

$$\lim_{n\to\infty} \|P^n(x,\,\cdot\,) - \pi\| = 2 \lim_{n\to\infty} \sup_A |P^n(x,A) - \pi(A)| = 0. \tag{13.8}$$

Obviously when (13.8) holds on a countable space, then (13.6) also holds and indeed holds uniformly in the end-point y. This move to the total variation norm, necessitated by the typical lack of structure of pointwise transitions in the general state space, will actually prove exceedingly fruitful rather than restrictive.

When the space is topological, it is also the case that total variation convergence implies weak convergence of the measures in question.

This is clear since (see Chapter 12) the latter is defined as convergence of expectations of functions which are not only bounded but also continuous. Hence the weak convergence of P^n to π as in Proposition 12.1.4 will be subsumed in results such as (13.4) provided the chain is suitably irreducible and positive.

Thus, for example, asymptotic properties of T-chains will be much stronger than those for arbitrary weak Feller chains even when a unique invariant measure exists for the latter.

Independence of initial and limiting distributions The second point to be made explicitly is that the limits in (13.8), and their refinements and extensions in Chapters 14–16, will typically be found to hold independently of the particular starting point x, and indeed we will be seeking conditions under which this is the case.

Having established this, however, the identification of the class of starting distributions for which particular asymptotic limits hold becomes a question of some importance, and the answer is not always obvious: in essence, if the chain starts with a distribution “too near infinity” then it may never reach the expected stationary distribution.

This is typified in (13.5), where the summability holds only for regular initial measures.

The same type of behavior, and the need to ensure that initial distributions are appropriately “regular” in extended ways, will be a highly visible part of the work in Chapters 14 and 15.

The role of renewal theory and splitting Thirdly, in developing the ergodic properties of ψ-irreducible chains we will use the splitting techniques of Chapter 5 in a systematic and fruitful way, and we will also need the properties of renewal sequences associated with visits to the atom in the split chain.

Up to now the existence of a "pseudo-atom" has not generated many results that could not have been derived (sometimes with considerable but nevertheless relatively elementary work) from the existence of petite sets: the only real "atom-based" result has been the existence of regular sets in Chapter 11. We have not given much reason for the reader to believe that the atom-based constructions are other than a gloss on the results obtainable through petite sets.

In Part III, however, we will find that the existence of atoms provides a critical step in the development of asymptotic results. This is due to the many limit theorems available for renewal processes, and we will prove such theorems as they fit into the Markov chain development.

We will also see that several generalizations of regular sets also play a key role in such results: the essential equivalence of regularity and positivity, developed in Chapter 11, becomes of far more than academic value in developing ergodic structures.

Ergodic chains Finally, a word on the term *ergodic*. We will adopt this term for chains where the limit in (13.6) or (13.8) holds as the time sequence $n \to \infty$, rather than as $n \to \infty$ through some subsequence.

Unfortunately, we know that in complete generality Markov chains may be periodic, in which case the limits in (13.6) or (13.8) can hold at best as we go through a periodic sequence nd as $n \to \infty$. Thus by definition, ergodic chains will be aperiodic, and a minor, sometimes annoying but always vital change to the structure of the results is needed in the periodic case.

We will therefore give results, typically, for the aperiodic context and give the required modification for the periodic case following the main statement when this seems worthwhile.

13.1 Ergodic chains on countable spaces

13.1.1 First-entrance last-exit decompositions

In this section we will approach the ergodic question for Markov chains in the countable state space case, before moving on to the general case in later sections. The methods are rather similar: indeed, given the splitting technique there will be a relatively small amount of extra work needed to move to the more general context.

Even in the countable case, the technique of proof we give is simpler and more powerful than that usually presented. One real simplification of the analysis through the use of total variation norm convergence results comes from an extension of the first-entrance and last-exit decompositions of Section 8.2, together with the representation of the invariant probability given in Theorem 10.2.1.

The *first-entrance last-exit decomposition*, for any states $x, y, \boldsymbol{\alpha} \in \mathsf{X}$ is given by

$$P^n(x,y) = {}_{\alpha}P^n(x,y) + \sum_{j=1}^{n-1}\Big[\sum_{k=1}^{j} {}_{\alpha}P^k(x,\boldsymbol{\alpha})P^{j-k}(\boldsymbol{\alpha},\boldsymbol{\alpha})\Big]{}_{\alpha}P^{n-j}(\boldsymbol{\alpha},y), \qquad (13.9)$$

where we have used the notation $\boldsymbol{\alpha}$ to indicate that the specific state being used for the decomposition is distinguished from the more generic states x, y which are the starting and end points of the decomposition.

We will wish in what follows to concentrate on the time variable rather than a particular starting point or end point, and it will prove particularly useful to have notation that reflects this. Let us hold the reference state $\boldsymbol{\alpha}$ fixed and introduce the three forms

$$a_x(n) := \mathsf{P}_x(\tau_\alpha = n) \tag{13.10}$$

$$u(n) := \mathsf{P}_\alpha(\Phi_n = \boldsymbol{\alpha}) \tag{13.11}$$

$$t_y(n) := {}_\alpha P^n(\boldsymbol{\alpha}, y). \tag{13.12}$$

This notation is designed to stress the role of $a_x(n)$ as a delay distribution in the renewal sequence of visits to $\boldsymbol{\alpha}$, and the "tail properties" of $t_y(n)$ in the representation of π: recall from (10.11) that

$$\begin{aligned} \pi(y) &= (\mathsf{E}_\alpha[\tau_\alpha])^{-1} \textstyle\sum_{j=1}^\infty {}_\alpha P^j(\boldsymbol{\alpha}, y) \\ &= \pi(\boldsymbol{\alpha}) \textstyle\sum_{j=1}^\infty t_y(j). \end{aligned} \tag{13.13}$$

Using this notation the first entrance and last exit decompositions become

$$\begin{aligned} P^n(x, \boldsymbol{\alpha}) &= \textstyle\sum_{j=0}^n \mathsf{P}_x(\tau_\alpha = j) P^{n-j}(\boldsymbol{\alpha}, \boldsymbol{\alpha}) \\ &= \textstyle\sum_{j=0}^n a_x(j) u(n-j) \\ P^n(\boldsymbol{\alpha}, y) &= \textstyle\sum_{j=0}^n P^j(\boldsymbol{\alpha}, \boldsymbol{\alpha}) {}_\alpha P^{n-j}(\boldsymbol{\alpha}, y) \\ &= \textstyle\sum_{j=0}^n u(j) t_y(n-j) \end{aligned}$$

or, using the convolution notation $a * b\,(n) = \sum_0^n a(j)b(n-j)$ introduced in Section 2.4.1,

$$P^n(x, \boldsymbol{\alpha}) = a_x * u\,(n) \tag{13.14}$$

$$P^n(\boldsymbol{\alpha}, y) = u * t_y\,(n). \tag{13.15}$$

The first-exit last-entrance decomposition (13.9) can be written similarly as

$$P^n(x, y) = {}_\alpha P^n(x, y) + a_x * u * t_y\,(n). \tag{13.16}$$

The power of these forms becomes apparent when we link them to the representation of the invariant measure given in (13.13). The next decomposition underlies all ergodic theorems for countable space chains.

Proposition 13.1.1 *Suppose that $\boldsymbol{\Phi}$ is a positive Harris recurrent chain on a countable space, with invariant probability π. Then for any $x, y, \boldsymbol{\alpha} \in \mathsf{X}$*

$$|P^n(x, y) - \pi(y)| \le {}_\alpha P^n(x, y) + |a_x * u - \pi(\boldsymbol{\alpha})| * t_y\,(n) + \pi(\boldsymbol{\alpha}) \sum_{j=n+1}^\infty t_y(j). \tag{13.17}$$

PROOF From the decomposition (13.16) we have

$$\begin{aligned} |P^n(x,y) - \pi(y)| \quad \leq \quad & {}_\alpha P^n(x,y) \\ & +|a_x * u * t_y\,(n) - \pi(\boldsymbol{\alpha}) \textstyle\sum_{j=1}^n t_y(j)| \\ & +|\pi(\boldsymbol{\alpha}) \textstyle\sum_{j=1}^n t_y(j) - \pi(y)|. \end{aligned} \tag{13.18}$$

Now we use the representation (13.13) for π and (13.17) is immediate. □

13.1.2 Solidarity from one ergodic state

If the three terms in (13.17) can all be made to converge to zero, we will have shown that $P^n(x,y) \to \pi(y)$ as $n \to \infty$. The two extreme terms involve the convergence of simple positive expressions, and finding bounds for both of these is at the level of calculation we have already used, especially in Chapters 10 and 11. The middle term involves a deeper limiting operation, and showing that this term does indeed converge is at the heart of proving ergodic theorems.

We can reduce the problem of this middle term entirely to one independent of the initial state x and involving only the reference state $\boldsymbol{\alpha}$. Suppose we have

$$|u(n) - \pi(\boldsymbol{\alpha})| \to 0, \qquad n \to \infty. \tag{13.19}$$

Then using Lemma D.7.1 we find

$$\lim_{n\to\infty} a_x * u\,(n) = \pi(\boldsymbol{\alpha}) \tag{13.20}$$

provided we have (as we do for a Harris recurrent chain) that for all x

$$\sum_j a_x(j) = \mathsf{P}_x(\tau_\alpha < \infty) = 1. \tag{13.21}$$

The convergence in (13.19) will be shown to hold for all states of an aperiodic positive chain in the next section: we first motivate our need for it, and for related results in renewal theory, by developing the ergodic structure of chains with "ergodic atoms".

Ergodic atoms

If $\boldsymbol{\Phi}$ is positive Harris, an atom $\boldsymbol{\alpha} \in \mathcal{B}^+(\mathsf{X})$ is called *ergodic* if it satisfies

$$\lim_{n\to\infty} |P^n(\boldsymbol{\alpha},\boldsymbol{\alpha}) - \pi(\boldsymbol{\alpha})| = 0. \tag{13.22}$$

In the positive Harris case note that an atom can be ergodic only if the chain is aperiodic.

With this notation, and the prescription for analyzing ergodic behavior inherent in Proposition 13.1.1, we can prove surprisingly quickly the following solidarity result.

Theorem 13.1.2 *If $\boldsymbol{\Phi}$ is a positive Harris chain on a countable space, and if there exists an ergodic atom $\boldsymbol{\alpha}$, then for every initial state x*

$$\|P^n(x,\,\cdot\,) - \pi\| \to 0, \qquad n \to \infty. \tag{13.23}$$

Proof On a countable space the total variation norm is given simply by

$$\|P^n(x,\,\cdot\,) - \pi\| = \sum_y |P^n(x,y) - \pi(y)|$$

and so by (13.17) we have the total variation norm bounded by three terms:

$$\|P^n(x,\,\cdot\,) - \pi\| \le \sum_y {}_\alpha P^n(x,y) \;+ \sum_y |a_x * u - \pi(\boldsymbol{\alpha})| * t_y\,(n) \;+ \sum_y \pi(\boldsymbol{\alpha}) \sum_{j=n+1}^{\infty} t_y(j). \tag{13.24}$$

We need to show each of these goes to zero. From the representation (13.13) of π, and Harris positivity

$$\infty > \sum_y \pi(y) = \pi(\boldsymbol{\alpha}) \sum_{j=1}^{\infty} \sum_y t_y(j). \tag{13.25}$$

The third term in (13.24) is the tail sum in this representation and so we must have

$$\pi(\boldsymbol{\alpha}) \sum_{j=n+1}^{\infty} \sum_y t_y(j) \to 0, \qquad n \to \infty. \tag{13.26}$$

The first term in (13.24) also tends to zero, for we have the interpretation

$$\sum_y {}_\alpha P^n(x,y) = \mathsf{P}_x(\tau_\alpha \ge n) \tag{13.27}$$

and since $\boldsymbol{\Phi}$ is Harris recurrent $\mathsf{P}_x(\tau_\alpha \ge n) \to 0$ for every x.

Finally, the middle term in (13.24) tends to zero by a double application of Lemma D.7.1, first using the assumption that $\boldsymbol{\alpha}$ is ergodic so that (13.20) holds and, once we have this, using the finiteness of $\sum_{j=1}^{\infty}\sum_y t_y(j)$ given by (13.25). □

This approach may be extended to give the Ergodic Theorem for a general space chain when there is an ergodic atom in the state space. A first-entrance last-exit decomposition will again give us an elegant proof in this case, and we prove such a result in Section 13.2.3, from which basis we wish to prove the same type of ergodic result for any positive Harris chain. To do this, we must of course prove that the atom $\check{\boldsymbol{\alpha}}$ for the split skeleton chain $\check{\boldsymbol{\Phi}}^m$, which we always have available, is an ergodic atom.

To show that atoms for aperiodic positive chains are indeed ergodic, which is crucial to completing this argument, we need results from renewal theory. This is therefore necessarily the subject of the next section.

13.2 Renewal and regeneration

13.2.1 Coupling renewal processes

When $\boldsymbol{\alpha}$ is a recurrent atom in X, the sequence of return times given by $\tau_\alpha(1) = \tau_\alpha$ and for $n > 1$

$$\tau_\alpha(n) = \min(j > \tau_\alpha(n-1) : \Phi_j = \boldsymbol{\alpha})$$

is a specific example of a *renewal process*, as defined in Section 2.4.1.

The asymptotic structure of renewal processes has, deservedly, been the subject of a great deal of analysis: such processes have a central place in the asymptotic theory of many kinds of stochastic processes, but nowhere more than in the development of asymptotic properties of general ψ-irreducible Markov chains.

Our goal in this section is to provide essentially those results needed for proving the ergodic properties of Markov chains, and we shall do this through the use of the so-called "coupling approach". We will regrettably do far less than justice to the full power of renewal and regenerative processes, or to the coupling method itself: for more details on renewal and regeneration, the reader should consult Feller [76] or Kingman [136], whilst the more recent flowering of the coupling technique is well covered by the recent book by Lindvall [155].

As in Section 2.4.1 we let $p = \{p(j)\}$ denote the distribution of the increments in a renewal process, whilst $a = \{a(j)\}$ and $b = \{b(j)\}$ will denote possible delays in the first increment variable S_0. For $n = 1, 2, \ldots$ let S_n denote the time of the $(n+1)^{st}$ renewal, so that the distribution of S_n is given by $a * p^{n*}$ if S_0 has the delay distribution a.

Recall the standard notation

$$u(n) = \sum_{j=0}^{\infty} p^{j*}(n)$$

for the renewal function for $n \geq 0$. Since $p^{0*} = \delta_0$ we have $u(0) = 1$; by convention we will set $u(-1) = 0$.

If we let $Z(n)$ denote the indicator variables

$$Z(n) = \begin{cases} 1 & S_j = n, \text{ some } j \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

then we have

$$\mathsf{P}_a(Z(n) = 1) = a * u(n),$$

and thus the renewal function represents the probabilities of $\{Z(n) = 1\}$ when there is no delay, or equivalently when $a = \delta_0$.

The coupling approach involves the study of two linked renewal processes with the same increment distribution but different initial distributions, and, most critically, defined on the same probability space.

To describe this concept we define two sets of mutually independent random variables

$$\{S_0, S_1, S_2, \ldots\}, \qquad \{S'_0, S'_1, S'_2, \ldots\}$$

where each of the variables $\{S_1, S_2, \ldots\}$ and $\{S'_1, S'_2, \ldots\}$ are independent and identically distributed with distribution $\{p(j)\}$; but where the distributions of the independent variables S_0, S'_0 are a, b.

The *coupling time* of the two renewal processes is defined as

$$T_{ab} = \min\{j : Z_a(j) = Z_b(j) = 1\}$$

where Z_a, Z_b are the indicator sequences of each renewal process. The random time T_{ab} is the first time that a renewal takes place simultaneously in both sequences, and from that point onwards, because of the loss of memory at the renewal epoch, the renewal processes are identical in distribution.

The key requirement to use this method is that this coupling time be almost surely finite. In this section we will show that if we have an aperiodic *positive recurrent* renewal process with finite mean

$$m_p := \sum_{j=0}^{\infty} j p(j) < \infty \tag{13.28}$$

then such coupling times are always almost surely finite.

Proposition 13.2.1 *If the increment distribution has an aperiodic distribution p with $m_p < \infty$ then for any initial proper distributions a, b*

$$\mathsf{P}(T_{ab} < \infty) = 1. \tag{13.29}$$

Proof Consider the linked forward recurrence time chain $\mathbf{V}^*$ defined by (10.19), corresponding to the two independent renewal sequences $\{S_n, S'_n\}$.

Let $\tau_{1,1} = \min(n : V^*_n = (1,1))$. Since the first coupling takes place at $\tau_{1,1} + 1$,

$$T_{ab} = \tau_{1,1} + 1$$

and thus we have that

$$\mathsf{P}(T_{ab} > n) = \mathsf{P}_{a\times b}(\tau_{1,1} \geq n). \tag{13.30}$$

But we know from Section 10.3.1 that, under our assumptions of aperiodicity of p and finiteness of m_p, the chain $\mathbf{V}^*$ is $\delta_{1,1}$-irreducible and positive Harris recurrent. Thus for any initial measure μ we have *a fortiori*

$$\mathsf{P}_\mu(\tau_{1,1} < \infty) = 1;$$

and hence in particular for the initial measure $a \times b$, it follows that

$$\mathsf{P}_{a\times b}(\tau_{1,1} \geq n) \to 0, \qquad n \to \infty$$

as required. □

This gives a structure sufficient to prove

Theorem 13.2.2 *Suppose that a, b, p are proper distributions on $\mathbb{Z}_+$, and that u is the renewal function corresponding to p. Then provided p is aperiodic with mean $m_p < \infty$*

$$|a * u\,(n) - b * u\,(n)| \to 0, \quad n \to \infty. \tag{13.31}$$

Proof Let us define the random variables

$$Z_{ab}(n) = \begin{cases} Z_a(n) & n < T_{ab} \\ Z_b(n) & n \geq T_{ab} \end{cases}$$

so that for any n

$$\mathsf{P}(Z_{ab}(n) = 1) = \mathsf{P}(Z_a(n) = 1). \tag{13.32}$$

We have that

$$\begin{aligned} |a * u\,(n) - b * u\,(n)| &= |\mathsf{P}(Z_a(n) = 1) - \mathsf{P}(Z_b(n) = 1)| \\ &= |\mathsf{P}(Z_{ab}(n) = 1) - \mathsf{P}(Z_b(n) = 1)| \\ &= |\mathsf{P}(Z_a(n) = 1, T_{ab} > n) + \mathsf{P}(Z_b(n) = 1, T_{ab} \leq n) \\ &\quad - \mathsf{P}(Z_b(n) = 1, T_{ab} > n) - \mathsf{P}(Z_b(n) = 1, T_{ab} \leq n)| \\ &= |\mathsf{P}(Z_a(n) = 1, T_{ab} > n) - \mathsf{P}(Z_b(n) = 1, T_{ab} > n)| \\ &\leq \max\{\mathsf{P}(Z_a(n) = 1, T_{ab} > n), \mathsf{P}(Z_b(n) = 1, T_{ab} > n)\} \\ &\leq \mathsf{P}(T_{ab} > n). \end{aligned} \tag{13.33}$$

But from Proposition 13.2.1 we have that $\mathsf{P}(T_{ab} > n) \to 0$ as $n \to \infty$, and (13.31) follows. □

We will see in Section 18.1.1 that Theorem 13.2.2 holds even without the assumption that $m_p < \infty$. For the moment, however, we will concentrate on further aspects of coupling when we are in the positive recurrent case.

13.2.2 Convergence of the renewal function

Suppose that we have a positive recurrent renewal sequence with finite mean $m_p < \infty$. Then the proper probability distribution $e = e(n)$ defined by

$$e(n) := m_p^{-1} \sum_{j=n+1}^{\infty} p(j) = m_p^{-1}(1 - \sum_{j=0}^{n} p(j)) \tag{13.34}$$

has been shown in (10.17) to be the invariant probability measure for the forward recurrence time chain $\mathbf{V}^+$ associated with the renewal sequence $\{S_n\}$. It also follows that the delayed renewal distribution corresponding to the initial distribution e is given for every $n \geq 0$ by

$$\begin{aligned} \mathsf{P}_e(Z(n) = 1) &= e * u\,(n) \\ &= m_p^{-1}(1 - p * 1) * u\,(n) \\ &= m_p^{-1}(1 - p * 1) * (\sum_{j=0}^{\infty} p^{*j})\,(n) \\ &= m_p^{-1}(1 + 1 * (\sum_{j=1}^{\infty} p^{*j})(n) - p * 1 * (\sum_{j=0}^{\infty} p^{*j})\,(n)) \\ &= m_p^{-1}. \end{aligned} \tag{13.35}$$

For this reason the distribution e is also called the *equilibrium distribution* of the renewal process.

These considerations show that in the positive recurrent case, the key quantity we considered for Markov chains in (13.22) has the representation

$$|u(n) - m_p^{-1}| = |\mathsf{P}_{\delta_0}(Z(n) = 1) - \mathsf{P}_e(Z(n) = 1)| \tag{13.36}$$

and in order to prove an asymptotic limiting result for an expression of this kind, we must consider the probabilities that $Z(n) = 1$ from the initial distributions δ_0, e.

But we have essentially evaluated this already. We have

Theorem 13.2.3 *Suppose that a, p are proper distributions on $\mathbb{Z}_+$, and that u is the renewal function corresponding to p. Then provided p is aperiodic and has a finite mean m_p*

$$|a * u\,(n) - m_p^{-1}| \to 0, \quad n \to \infty. \tag{13.37}$$

PROOF The result follows from Theorem 13.2.2 by substituting the equilibrium distribution e for b and using (13.35). □

This has immediate application in the case where the renewal process is the return time process to an accessible atom for a Markov chain.

Proposition 13.2.4 **(i)** *If $\boldsymbol{\Phi}$ is a positive recurrent aperiodic Markov chain then any atom $\boldsymbol{\alpha}$ in $\mathcal{B}^+(\mathsf{X})$ is ergodic.*

(ii) *If $\boldsymbol{\Phi}$ is a positive recurrent aperiodic Markov chain on a countable space then for every initial state x*

$$\|P^n(x, \cdot\,) - \pi\| \to 0, \qquad n \to \infty. \tag{13.38}$$

PROOF We know from Proposition 10.2.2 that if $\boldsymbol{\Phi}$ is positive recurrent then the mean return time to any atom in $\mathcal{B}^+(\mathsf{X})$ is finite. If the chain is aperiodic then (i) follows directly from Theorem 13.2.3 and the definition (13.22).

The conclusion in (ii) then follows from (i) and Theorem 13.1.2. □

It is worth stressing explicitly that this result depends on the classification of positive chains in terms of finite mean return times to atoms: that is, in using renewal theory it is the equivalence of positivity and regularity of the chain that is utilized.

13.2.3 The regenerative decomposition for chains with atoms

We now consider general positive Harris chains and use the renewal theorems above to commence development of their ergodic properties.

In order to use the splitting technique for analysis of total variation norm convergence for general state space chains we must extend the first-entrance last-exit decomposition (13.9) to general spaces. For any sets $A, B \in \mathcal{B}(\mathsf{X})$ and $x \in \mathsf{X}$ we have, by decomposing the event $\{\Phi_n \in B\}$ over the times of the first and last entrances to A prior to n, that

$$P^n(x, B) = {}_AP^n(x, B) + \sum_{j=1}^{n-1} \int_A \Big[\sum_{k=1}^{j} \int_A {}_AP^k(x, dv) P^{j-k}(v, dw)\Big] {}_AP^{n-j}(w, B). \tag{13.39}$$

If we suppose that there is an atom $\boldsymbol{\alpha}$ and take $A = \boldsymbol{\alpha}$ then these forms are somewhat simplified: the decomposition (13.39) reduces to

$$P^n(x,B) = {}_\alpha P^n(x,B) + \sum_{j=1}^{n-1} \Big[\sum_{k=1}^{j} {}_\alpha P^k(x,\boldsymbol{\alpha}) P^{j-k}(\boldsymbol{\alpha},\boldsymbol{\alpha})\Big] {}_\alpha P^{n-j}(\boldsymbol{\alpha},B). \tag{13.40}$$

In the general state space case it is natural to consider convergence from an arbitrary initial distribution λ. It is equally natural to consider convergence of the integrals

$$\mathsf{E}_\lambda[f(\Phi_n)] = \int \lambda(dx) \int P^n(x,dy) f(w) \tag{13.41}$$

for arbitrary non-negative functions f. We will use either the probabilistic or the operator theoretic version of this quantity (as given by the two sides of (13.41)) interchangeably, as seems most transparent, in what follows.

We explore convergence of $\mathsf{E}_\lambda[f(\Phi_n)]$ for general (unbounded) f in detail in Chapter 14. Here we concentrate on bounded f, in view of the definition (13.7) of the total variation norm.

When $\boldsymbol{\alpha}$ is an atom in $\mathcal{B}^+(\mathsf{X})$, let us therefore extend the notation in (13.10)-(13.12) to the forms

$$a_\lambda(n) = \mathsf{P}_\lambda(\tau_\alpha = n) \tag{13.42}$$

$$t_f(n) = \int {}_\alpha P^n(\boldsymbol{\alpha},dy) f(y) = \mathsf{E}_\alpha[f(\Phi_n)\mathbb{1}\{\tau_\alpha \ge n\}] : \tag{13.43}$$

these are well-defined (although possibly infinite) for any non-negative function f on X and any probability measure λ on $\mathcal{B}(\mathsf{X})$.

As in (13.14) and (13.15) we can use this terminology to write the first entrance and last exit formulations as

$$\int \lambda(dx) P^n(x,\boldsymbol{\alpha}) = a_\lambda * u\,(n) \tag{13.44}$$

$$\int P^n(\boldsymbol{\alpha},dy) f(y) = u * t_f\,(n). \tag{13.45}$$

The first-entrance last-exit decomposition (13.40) can similarly be formulated, for any λ, f, as

$$\int \lambda(dx) \int P^n(x,dw) f(w) = \int \lambda(dx) \int {}_\alpha P^n(x,dw) f(w) + a_\lambda * u * t_f\,(n). \tag{13.46}$$

The general state space version of Proposition 13.1.1 provides the critical bounds needed for our approach to ergodic theorems. Using the notation of (13.41) we have two bounds which we shall refer to as *Regenerative Decompositions.*

Theorem 13.2.5 *Suppose that $\boldsymbol{\Phi}$ admits an accessible atom $\boldsymbol{\alpha}$ and is positive Harris recurrent with invariant probability measure π. Then for any probability measure λ and $f \ge 0$,*

$$\begin{aligned} |\,\mathsf{E}_\lambda[f(\Phi_n)] - \mathsf{E}_\alpha[f(\Phi_n)]\,| \quad &\le \quad \mathsf{E}_\lambda[f(\Phi_n)\mathbb{1}\{\tau_\alpha \ge n\}] \\ &\qquad + |a_\lambda * u - u| * t_f\,(n) \end{aligned} \tag{13.47}$$

$$\begin{aligned} |\,\mathsf{E}_\lambda[f(\Phi_n)] - \mathsf{E}_\pi[f(\Phi_n)]\,| \quad &\le \quad \mathsf{E}_\lambda[f(\Phi_n)\mathbb{1}\{\tau_\alpha \ge n\}] \\ &\qquad + |\,a_\lambda * u - \pi(\boldsymbol{\alpha})\,| * t_f\,(n) \\ &\qquad + \pi(\boldsymbol{\alpha}) \textstyle\sum_{j=n+1}^\infty t_f(j). \end{aligned} \tag{13.48}$$

Proof The first-entrance last-exit decomposition (13.46), in conjunction with the simple last exit decomposition in the form (13.45), gives the first bound on the distance between $\mathsf{E}_\lambda[f(\Phi_n)]$ and $\mathsf{E}_\alpha[f(\Phi_n)]$ in (13.47).

The decomposition (13.46) also gives

$$\begin{aligned} |\,\mathsf{E}_\lambda[f(\Phi_n)] - \mathsf{E}_\pi[f(\Phi_n)]\,| \quad \leq \quad & \mathsf{E}_\lambda[f(\Phi_n)\mathbb{1}\{\tau_\alpha \geq n\}] \\ & + \Big| a_\lambda * u * t_f\,(n) - \pi(\boldsymbol{\alpha}) \textstyle\sum_{j=1}^n t_f(j) \Big| \\ & + \Big| \pi(\boldsymbol{\alpha}) \textstyle\sum_{j=1}^n t_f(j) - \int \pi(dw) f(w) \Big| . \end{aligned} \tag{13.49}$$

Now in the general state space case we have the representation for π given from (10.32) by

$$\int \pi(dw) f(w) = \pi(\boldsymbol{\alpha}) \sum_1^\infty t_f(y); \tag{13.50}$$

and (13.48) now follows from (13.49). □

The Regenerative Decomposition (13.48) in Theorem 13.2.5 shows clearly what is needed to prove limiting results in the presence of an atom. Suppose that f is bounded. Then we must

(E1) control the third term in (13.48), which involves questions of the finiteness of π, but is independent of the initial measure λ: this finiteness is guaranteed for positive chains by definition;

(E2) control the first term in (13.48), which involves questions of the finiteness of the hitting time distribution of τ_α when the chain begins with distribution λ; this is automatically finite as required for a Harris recurrent chain, even without positive recurrence, although for chains which are only recurrent it clearly needs care;

(E3) control the middle term in (13.48), which again involves finiteness of π to bound its last element, but more crucially then involves only the ergodicity of the atom $\boldsymbol{\alpha}$, regardless of λ: for we know from Lemma D.7.1 that if the atom is ergodic so that (13.19) holds then also

$$\lim_{n\to\infty} a_\lambda * u\,(n) = \pi(\boldsymbol{\alpha}), \tag{13.51}$$

since for $\boldsymbol{\Phi}$ a Harris recurrent chain, any probability measure λ satisfies

$$\sum_n a_\lambda(n) = \mathsf{P}_\lambda(\tau_\alpha < \infty) = 1. \tag{13.52}$$

Thus recurrence, or rather Harris recurrence, will be used twice to give bounds: positive recurrence gives one bound; and, centrally, the equivalence of positivity and regularity ensures the atom is ergodic, exactly as in Theorem 13.2.3.

Bounded functions are the only ones relevant to total variation convergence. The Regenerative Decomposition is however valid for all $f \geq 0$. Bounds in this decomposition then involve integrability of f with respect to π, and a non-trivial extension of regularity to what will be called f-regularity. This will be held over to the next chapter, and here we formalize the above steps and incorporate them with the splitting technique, to prove the Aperiodic Ergodic Theorem.

13.3 Ergodicity of positive Harris chains

13.3.1 Strongly aperiodic chains

The prescription (E1)-(E3) above for ergodic behavior is followed in the proof of

Theorem 13.3.1 *If $\boldsymbol{\Phi}$ is a positive Harris recurrent and strongly aperiodic chain then for any initial measure λ*

$$\|\int \lambda(dx)P^n(x,\,\cdot\,) - \pi\| \to 0, \qquad n \to \infty. \tag{13.53}$$

Proof **(i)** Let us first assume that there is an accessible ergodic atom in the space. The proof is virtually identical to that in the countable case. We have

$$\|\int \lambda(dx)P^n(x,\,\cdot\,) - \pi\| = \sup_{|f|\le 1}\left|\int \lambda(dx)\int P^n(x,dw)f(w) - \int \pi(dw)f(w)\right| \tag{13.54}$$

and we use (13.48) to bound these terms uniformly for functions $f \le 1$.

Since $|f| \le 1$ the third term in (13.48) is bounded above by

$$\pi(\boldsymbol{\alpha})\sum_{n+1}^{\infty} t_1(j) \to 0, \qquad n \to \infty \tag{13.55}$$

since it is the tail sum in the representation (13.50) of $\pi(\mathsf{X})$.

The second term in (13.48) is bounded above by

$$|a_\lambda * u - \pi(\boldsymbol{\alpha})| * t_1(n) \to 0, \qquad n \to \infty, \tag{13.56}$$

by Lemma D.7.1; here we use the fact that $\boldsymbol{\alpha}$ is ergodic and, again, the representation that $\pi(\mathsf{X}) = \pi(\alpha)\sum_1^\infty t_1(j) < \infty$.

We must finally control the first term. To do this, we need only note that, again since $|f| \le 1$, we have

$$\mathsf{E}_\lambda[f(\Phi_n)\mathbb{1}\{\tau_\alpha \ge n\}] \le \mathsf{P}_\lambda(\tau_\alpha \ge n) \tag{13.57}$$

and this expression tends to zero by monotone convergence as $n \to \infty$, since $\boldsymbol{\alpha}$ is Harris recurrent and $\mathsf{P}_x(\tau_\alpha < \infty) = 1$ for every x.

Notice explicitly that in (13.55)-(13.57) the bounds which tend to zero are independent of the particular $|f| \le 1$, and so we have the required supremum norm convergence.

(ii) Now assume that $\boldsymbol{\Phi}$ is strongly aperiodic. Consider the split chain $\check{\boldsymbol{\Phi}}$: we know this is also strongly aperiodic from Proposition 5.5.6 (ii), and positive Harris from Proposition 10.4.2. Thus from Proposition 13.2.4 the atom $\check{\boldsymbol{\alpha}}$ is ergodic. Now our use of total variation norm convergence renders the transfer to the original chain easy. Using the fact that the original chain is the marginal chain of the split chain, and that π is the marginal measure of $\check{\pi}$, we have immediately

$$\begin{aligned}\|\int \lambda(dx)P^n(x,\,\cdot\,) - \pi\| &= 2\sup_{A\in\mathcal{B}(\mathsf{X})}|\int_{\mathsf{X}} \lambda(dx)P^n(x,A) - \pi(A)| \\ &= 2\sup_{A\in\mathcal{B}(\mathsf{X})}|\int_{\check{\mathsf{X}}} \lambda^*(dx_i)\check{P}^n(x_i,A) - \check{\pi}(A)|\end{aligned}$$

$$\begin{aligned} &\leq \; 2 \sup_{\check{B} \in \mathcal{B}(\check{\mathsf{X}})} \Big| \int_{\check{\mathsf{X}}} \lambda^*(dx_i) \check{P}^n(x_i, \check{B}) - \check{\pi}(\check{B}) \Big| \\ &= \; \Big\| \int \lambda^*(dx_i) \check{P}^n(x_i, \cdot\,) - \check{\pi} \Big\|, \qquad (13.58) \end{aligned}$$

where the inequality follows since the first supremum is over sets in $\mathcal{B}(\check{\mathsf{X}})$ of the form $A_0 \cup A_1$ and the second is over all sets in $\mathcal{B}(\check{\mathsf{X}})$.

Applying the result (i) for chains with accessible atoms shows that the total variation norm in (13.58) for the split chain tends to zero, so we are finished. □

13.3.2 The ergodic theorem for ψ-irreducible chains

We can now move from the strongly aperiodic chain result to arbitrary aperiodic Harris recurrent chains. This is made simpler as a result of another useful property of the total variation norm.

Proposition 13.3.2 *If π is invariant for P then the total variation norm*

$$\Big\| \int \lambda(dx) P^n(x, \cdot\,) - \pi \Big\|$$

is non-increasing in n.

PROOF We have from the definition of total variation and the invariance of π that

$$\begin{aligned} &\Big\| \int \lambda(dx) P^{n+1}(x, \cdot\,) - \pi \Big\| \\ &= \sup_{f: |f| \leq 1} \Big| \int \lambda(dx) P^{n+1}(x, dy) f(y) - \int \pi(dy) f(y) \Big| \\ &= \sup_{f: |f| \leq 1} \Big| \int \lambda(dx) P^n(x, dw) \Big[\int P(w, dy) f(y) \Big] - \int \pi(dw) \Big[\int P(w, dy) f(y) \Big] \Big| \\ &\leq \sup_{f: |f| \leq 1} \Big| \int \lambda(dx) P^n(x, dw) f(w) - \int \pi(dw) f(w) \Big| \qquad (13.59) \end{aligned}$$

since whenever $|f| \leq 1$ we also have $|Pf| \leq 1$. □

We can now prove the general state space result in the aperiodic case.

Theorem 13.3.3 *If $\boldsymbol{\Phi}$ is positive Harris and aperiodic then for every initial distribution λ*

$$\Big\| \int \lambda(dx) P^n(x, \cdot\,) - \pi \Big\| \to 0, \qquad n \to \infty. \qquad (13.60)$$

PROOF Since for some m the skeleton $\boldsymbol{\Phi}^m$ is strongly aperiodic, and also positive Harris by Theorem 10.4.5, we know that

$$\Big\| \int \lambda(dx) P^{nm}(x, \cdot\,) - \pi \Big\| \to 0, \qquad n \to \infty. \qquad (13.61)$$

The result for P^n then follows immediately from the monotonicity in (13.59). □

As we mentioned in the discussion of the periodic behavior of Markov chains, the results are not quite as simple to state in the periodic as in the aperiodic case; but they can be easily proved once the aperiodic case is understood.

The asymptotic behavior of positive recurrent chains which may not be Harris is also easy to state now that we have analyzed positive Harris chains.

The final formulation of these results for quite arbitrary positive recurrent chains is

Theorem 13.3.4 **(i)** *If Φ is positive Harris with period $d \geq 1$ then for every initial distribution λ*

$$\|d^{-1} \int \lambda(dx) \sum_{r=0}^{d-1} P^{nd+r}(x, \cdot\,) - \pi\| \to 0, \qquad n \to \infty. \tag{13.62}$$

(ii) *If Φ is positive recurrent with period $d \geq 1$ then there is a π-null set N such that for every initial distribution λ with $\lambda(N) = 0$*

$$\|d^{-1} \int \lambda(dx) \sum_{r=0}^{d-1} P^{nd+r}(x, \cdot\,) - \pi\| \to 0, \qquad n \to \infty. \tag{13.63}$$

PROOF The result (i) is straightforward to check from the existence of cycles in Section 5.4.3, together with the fact that the chain restricted to each cyclic set is aperiodic and positive Harris on the d-skeleton. We then have (ii) as a direct corollary of the decomposition of Theorem 9.1.5. □

Finally, let us complete the circle by showing the last step in the equivalences in Theorem 13.0.1. Notice that (13.64) is ensured by (13.1), using the Dominated Convergence Theorem, so that our next result is in fact marginally stronger than the corresponding statement of the Aperiodic Ergodic Theorem.

Theorem 13.3.5 *Let Φ be ψ-irreducible and aperiodic, and suppose that there exists some ν-small set $C \in \mathcal{B}^+(\mathsf{X})$ and some $P^\infty(C) > 0$ such that as $n \to \infty$*

$$\int_C \nu_C(dx)(P^n(x, C) - P^\infty(C)) \to 0 \tag{13.64}$$

where $\nu_C(\,\cdot\,) = \nu(\,\cdot\,)/\nu(C)$ is normalized to a probability on C. Then the chain is positive, and there exists a ψ-null set such that for every initial distribution λ with $\lambda(N) = 0$

$$\| \int \lambda(dx) P^n(x, \cdot\,) - \pi\| \to 0, \qquad n \to \infty. \tag{13.65}$$

PROOF Using the Nummelin splitting via the set C for the m-skeleton, we find that (13.64) taken through the sublattice nm is equivalent to

$$\delta^{-1}(\check{P}^n(\check{\boldsymbol{\alpha}}, \check{\boldsymbol{\alpha}}) - \delta P^\infty(C)) \to 0. \tag{13.66}$$

Thus the atom $\check{\boldsymbol{\alpha}}$ is ergodic and the results of Section 13.3 all hold, with $P^\infty(C) = \pi(C)$. □

13.4 Sums of transition probabilities

13.4.1 A stronger coupling theorem

In order to derive bounds such as those in (13.5) on the sums of n-step total variation differences from the invariant measure π, we need to bound sums of terms such as $|P^n(\boldsymbol{\alpha}, \boldsymbol{\alpha}) - \pi(\boldsymbol{\alpha})|$ rather than the individual terms. This again requires a renewal theory result, which we prove using the coupling method. We have

Proposition 13.4.1 *Suppose that a, b, p are proper distributions on $\mathbb{Z}_+$, and that u is the renewal function corresponding to p. Then provided p is aperiodic and has a finite mean m_p, and a, b also have finite means m_a, m_b, we have*

$$\sum_{n=0}^{\infty} |a * u\,(n) - b * u\,(n)| < \infty. \tag{13.67}$$

Proof We have from (13.33) that

$$\sum_{n=0}^{\infty} |a * u\,(n) - b * u\,(n)| \le \sum_{n=0}^{\infty} \mathsf{P}(T_{ab} > n) = \mathsf{E}[T_{ab}]. \tag{13.68}$$

Now we know from Section 10.3.1 that when p is aperiodic and $m_p < \infty$, the linked forward recurrence time chain $\mathbf{V}^*$ is positive recurrent with invariant probability

$$e^*(i,j) = e(i)e(j).$$

Hence from any state (i,j) with $e^*(i,j) > 0$ we have as in Proposition 11.1.1

$$\mathsf{E}_{i,j}[\tau_{1,1}] < \infty. \tag{13.69}$$

Let us consider specifically the initial distributions δ_0 and δ_1: these correspond to the undelayed renewal process and the process delayed by exactly one time unit respectively. For this choice of initial distribution we have for $n > 0$

$$\begin{aligned} \delta_0 * u\,(n) &= u(n) \\ \delta_1 * u\,(n) &= u(n-1) \end{aligned}$$

Now $\mathsf{E}[T_{01}] \le \mathsf{E}_{1,2}[\tau_{1,1}] + 1$ and it is certainly the case that $e^*(1,2) > 0$. So from (13.30), (13.68) and (13.69)

$$\mathrm{Var}\,(u) := \sum_{n=0}^{\infty} |u(n) - u(n-1)| \le \mathsf{E}_{1,2}[\tau_{1,1}] + 1 < \infty. \tag{13.70}$$

We now need to extend the result to more general initial distributions with finite mean. By the triangle inequality it suffices to consider only one arbitrary initial distribution a and to take the other as δ_0. To bound the resulting quantity $|a * u\,(n) - u(n)|$ we write the upper tails of a for $k \ge 0$ as

$$\overline{a}(k) := \sum_{j=k+1}^{\infty} a(j) = 1 - \sum_{j=0}^{k} a(j)$$

and put

$$w(k) = |u(k) - u(k-1)|.$$

We then have the relation

$$\begin{aligned}
\overline{a} * w\,(n) &= \sum_{j=0}^{n} \overline{a}(j) w(n-j) \\
&\geq |\sum_{j=0}^{n} [1 - \sum_{k=0}^{j} a(k)][u(n-j) - u(n-j-1)]| \\
&= |\sum_{j=0}^{n} [u(n-j) - u(n-j-1)] \\
&\qquad - \sum_{j=0}^{n} \sum_{k=0}^{j} a(k)[u(n-j) - u(n-j-1)]| \\
&= |u(n) - \sum_{k=0}^{n} a(k) \sum_{j=k}^{n} [u(n-j) - u(n-j-1)]| \\
&= |u(n) - \sum_{k=0}^{n} a(k) u(n-k)| \qquad (13.71)
\end{aligned}$$

so that

$$\sum_n |u(n) - a * u\,(n)| \leq \sum_n \overline{a} * w\,(n) = [\sum_n \overline{a}(n)][\sum_n w(n)]. \qquad (13.72)$$

But by assumption the mean $m_a = \sum \overline{a}(n)$ is finite, and (13.70) shows that the sequence $w(n)$ is also summable; and so we have

$$\sum_n |u(n) - a * u\,(n)| \leq m_a \mathrm{Var}\,(u) < \infty \qquad (13.73)$$

as required. □

It is obviously of considerable interest to know under what conditions we have

$$\sum_n |a * u\,(n) - m_p^{-1}| < \infty; \qquad (13.74)$$

that is, when this result holds with the equilibrium measure as one of the initial measures.

Using Proposition 13.4.1 we know that this will occur if the equilibrium distribution e has a finite mean; and since we know the exact structure of e it is obvious that $m_e < \infty$ if and only if

$$s_p := \sum_n n^2 p(n) < \infty.$$

In fact, using the exact form

$$m_e = [s_p - m_p]/[2m_p]$$

we have from Proposition 13.4.1 and in particular the bound (13.72) the following pleasing corollary:

Proposition 13.4.2 *If p is an aperiodic distribution with $s_p < \infty$ then*

$$\sum_n |u(n) - m_p^{-1}| \leq \mathrm{Var}\,(u)[s_p - m_p]/[2m_p] < \infty. \qquad (13.75)$$

□

13.4.2 General chains with atoms

We now refine the ergodic theorem Theorem 13.3.3 to give conditions under which sums such as

$$\sum_{n=1}^{\infty} \|P^n(x,\,\cdot\,) - P^n(y,\,\cdot\,)\|$$

are finite. A result such as this requires regularity of the initial states x, y: recall from Chapter 11 that a probability measure μ on $\mathcal{B}(\mathsf{X})$ is called regular, if

$$\mathsf{E}_\mu[\tau_B] < \infty, \qquad B \in \mathcal{B}^+(\mathsf{X}).$$

We will again follow the route of first considering chains with an atom, then translating the results to strongly aperiodic and thence to general chains.

Theorem 13.4.3 *Suppose $\boldsymbol{\Phi}$ is an aperiodic positive Harris chain and suppose that the chain admits an atom $\boldsymbol{\alpha} \in \mathcal{B}^+(\mathsf{X})$. Then for any regular initial distributions λ, μ,*

$$\sum_{n=1}^{\infty} \int\int \lambda(dx)\mu(dy)\|P^n(x,\,\cdot\,) - P^n(y,\,\cdot\,)\| < \infty; \tag{13.76}$$

and in particular, if $\boldsymbol{\Phi}$ is regular, then for every $x, y \in \mathsf{X}$

$$\sum_{n=1}^{\infty} \|P^n(x,\,\cdot\,) - P^n(y,\,\cdot\,)\| < \infty. \tag{13.77}$$

PROOF By the triangle inequality it will suffice to prove that

$$\sum_{n=1}^{\infty} \int \lambda(dx)\|P^n(x,\,\cdot\,) - P^n(\boldsymbol{\alpha},\,\cdot\,)\| < \infty; \tag{13.78}$$

that is, to assume that one of the initial distributions is δ_α.

If we sum the first Regenerative Decomposition (13.47) in Theorem 13.2.5 with $f \leq 1$ we find (13.78) is bounded by two sums: firstly,

$$\sum_{n=1}^{\infty} \int \lambda(dx)\,{}_\alpha P^n(x, \mathsf{X}) \quad = \quad \mathsf{E}_\lambda[\tau_\alpha] \tag{13.79}$$

which is finite since λ is regular; and secondly

$$\Big\{\sum_{n=1}^{\infty} \int \lambda(dx)|a_x * u\,(n) - u(n)|\Big\}\Big\{\sum_{n=1}^{\infty} {}_\alpha P^n(\boldsymbol{\alpha}, \mathsf{X})\Big\}. \tag{13.80}$$

To bound this term note that $\sum_{n=1}^{\infty} {}_\alpha P^n(\boldsymbol{\alpha}, \mathsf{X}) = \mathsf{E}_\alpha[\tau_\alpha] < \infty$ since every accessible atom is regular from Theorems 11.1.4 and 11.1.2; and so it remains only to prove that

$$\sum_{n=1}^{\infty} \int \lambda(dx)|a_x * u\,(n) - u(n)| < \infty. \tag{13.81}$$

From (13.72) we have

$$\begin{aligned}\sum_{n=1}^{\infty} |a_x * u\,(n) - u(n)| \quad &\leq \quad \Big(\sum_{n=1}^{\infty} a_x(n)\Big)\Big(\sum_{n=1}^{\infty} |u(n) - u(n-1)|\Big) \\ &= \quad \mathsf{E}_x[\tau_\alpha]\mathrm{Var}\,(u),\end{aligned}$$

and hence the sum (13.81) is bounded by $\mathsf{E}_\lambda[\tau_\alpha]\mathrm{Var}\,(u)$, which is again finite by Proposition 13.4.1 and regularity of λ. □

13.4.3 General aperiodic chains

The move from the atomic case is by now familiar.

Theorem 13.4.4 *Suppose $\boldsymbol{\Phi}$ is an aperiodic positive Harris chain. For any regular initial distributions λ, μ*

$$\sum_{n=1}^{\infty} \int\int \lambda(dx)\mu(dy)\|P^n(x,\,\cdot\,) - P^n(y,\,\cdot\,)\| < \infty. \tag{13.82}$$

Proof Consider the strongly aperiodic case. The theorem is valid for the split chain, since the split measures λ^*, μ^* are regular for $\check{\boldsymbol{\Phi}}$: this follows from the characterization in Theorem 11.3.12.

Since the result is a total variation result it remains valid when restricted to the original chain, as in (13.58).

In the arbitrary aperiodic case we can apply Proposition 13.3.2 to move to a skeleton chain, as in the proof of Theorem 13.2.5. □

The most interesting special case of this result is given in the following theorem.

Theorem 13.4.5 *Suppose $\boldsymbol{\Phi}$ is an aperiodic positive Harris chain and that $\boldsymbol{\alpha}$ is an accessible atom. If*

$$\mathsf{E}_\alpha[\tau_\alpha^2] < \infty \tag{13.83}$$

then for any regular initial distribution λ

$$\sum_{n=1}^{\infty} \|\lambda P^n - \pi\| < \infty. \tag{13.84}$$

□

Proof In the case where there is an atom $\boldsymbol{\alpha}$ in the space, we have as in Proposition 13.4.2 that π is a regular measure when the second-order moment (13.83) is finite, and the result is then a consequence of Theorem 13.4.4.

13.5 Commentary

It is hard to know where to start in describing contributions to these theorems. The countable chain case has an immaculate pedigree: Kolmogorov [139] first proved this result, and Feller [76] and Chung [49] give refined approaches to the single-state version (13.6), essentially through analytic proofs of the lattice renewal theorem.

The general state space results in the positive recurrent case are largely due to Harris [95] and to Orey [207]. Their results and related material, including a null recurrent version in Section 18.1 below are all discussed in a most readable way in Orey's monograph [208]. Prior to the development of the splitting technique, proofs utilized the concept of the tail σ-field of the chain, which we have not discussed so far, and will only touch on in Chapter 17.

The coupling proofs are much more recent, although they are usually dated to Doeblin [66]. Pitman [215] first exploited the positive recurrent coupling in the way we give it here, and his use of the result in Proposition 13.4.1 was even then new, as was Theorem 13.4.4.

Our presentation of this material has relied heavily on Nummelin [202], and further related results can be found in his Chapter 6. In particular, for results of this kind in a more general setting where the renewal sequence is allowed to vary from the probabilistic structure with $\sum_n p(n) = 1$ which we have used, the reader is referred to Chapters 4 and 6 of [202].

It is interesting to note that the first-entrance last-exit decomposition, which shows so clearly the role of the single ergodic atom, is a relative late-comer on the scene. Although probably used elsewhere, it surfaces in the form given here in Nummelin [200] and Nummelin and Tweedie [206], and appears to be less than well known even in the countable state space case. Certainly, the proof of ergodicity is much simplified by using the Regenerative Decomposition.

We should note, for the reader who is yet again trying to keep stability nomenclature straight, that even the "ergodicity" terminology we use here is not quite standard: for example, Chung [49] uses the word ergodic to describe certain ratio limit theorems rather than the simple limit theorem of (13.8). We do not treat ratio limit theorems in this book, except in passing in Chapter 17: it is a notable omission, but one dictated by the lack of interesting examples in our areas of application. Hence no confusion should arise, and our ergodic chains certainly coincide with those of Feller [76], Nummelin [202] and Revuz [223]. The latter two books also have excellent treatments of ratio limit theorems.

We have no examples in this chapter. This is deliberate. We have shown in Chapter 11 how to classify specific models as positive recurrent using drift conditions: we can say little else here other than that we now know that such models converge in the relatively strong total variation norm to their stationary distributions. Over the course of the next three chapters, we will however show that other much stronger ergodic properties hold under other more restrictive drift conditions; and most of the models in which we have been interested will fall into these more strongly stable categories.