4

Irreducibility

This chapter is devoted to the fundamental concept of irreducibility: the idea that all parts of the space can be reached by a Markov chain, no matter what the starting point. Although the initial results are relatively simple, the impact of an appropriate irreducibility structure will have wide-ranging consequences, and it is therefore of critical importance that such structures be well understood.

The results summarized in Theorem 4.0.1 are the highlights of this chapter from a theoretical point of view. An equally important aspect of the chapter is, however, to show through the analysis of a number of models just what techniques are available in practice to ensure the initial condition of Theorem 4.0.1 ("φ-irreducibility") holds, and we believe that these will repay equally careful consideration.

Theorem 4.0.1 *If there exists an "irreducibility" measure φ on $\mathcal{B}(\mathsf{X})$ such that for every state x*

$$\varphi(A) > 0 \Rightarrow L(x, A) > 0 \tag{4.1}$$

then there exists an essentially unique "maximal" irreducibility measure ψ on $\mathcal{B}(\mathsf{X})$ such that

(i) *for every state x we have $L(x, A) > 0$ whenever $\psi(A) > 0$, and also*

(ii) *if $\psi(A) = 0$, then $\psi(\bar{A}) = 0$, where*

$$\bar{A} := \{y : L(y, A) > 0\} ;$$

(iii) *if $\psi(A^c) = 0$, then $A = A_0 \cup N$ where the set N is also ψ-null, and the set A_0 is absorbing in the sense that*

$$P(x, A_0) \equiv 1, \qquad x \in A_0.$$

PROOF The existence of a measure ψ satisfying the irreducibility conditions (i) and (ii) is shown in Proposition 4.2.2, and that (iii) holds is in Proposition 4.2.3. □

The term "maximal" is justified since we will see that φ is absolutely continuous with respect to ψ, written $\psi \succ \varphi$, for every φ satisfying (4.1); here the relation of absolute continuity of φ with respect to ψ means that $\psi(A) = 0$ implies $\varphi(A) = 0$.

Verifying (4.1) is often relatively painless. State space models on $\mathbb{R}^k$ for which the noise or disturbance distribution has a density with respect to Lebesgue measure will typically have such a property, with φ taken as Lebesgue measure restricted to an open set (see Section 4.4, or in more detail, Chapter 7); chains with a regeneration point $\boldsymbol{\alpha}$ reached from everywhere will satisfy (4.1) with the trivial choice of $\varphi = \delta_\alpha$ (see Section 4.3).

The extra benefit of defining much more accurately the sets which are avoided by "most" points, as in Theorem 4.0.1 (ii), or of knowing that one can omit ψ-null sets and restrict oneself to an absorbing set of "good" points as in Theorem 4.0.1 (iii), is then of surprising value, and we use these properties again and again. These are however far from the most significant consequences of the seemingly innocuous assumption (4.1): far more will flow in Chapter 5, and thereafter.

The most basic structural results for Markov chains, which lead to this formalization of the concept of irreducibility, involve the analysis of communicating states and sets. If one can tell which sets can be reached with positive probability from particular starting points $x \in \mathsf{X}$, then one can begin to have an idea of how the chain behaves in the longer term, and then give a more detailed description of that longer term behavior.

Our approach therefore commences with a description of communication between sets and states which precedes the development of irreducibility.

4.1 Communication and Irreducibility: Countable Spaces

When X is general, it is not always easy to describe the specific points or even sets which can be reached from different starting points $x \in \mathsf{X}$. To guide our development, therefore, we will first consider the simpler and more easily understood situation when the space X is countable; and to fix some of these ideas we will initially analyze briefly the communication behavior of the random walk on a half line defined by (RWHL1), in the case where the increment variable takes on integer values.

4.1.1 Communication: random walk on a half line

Recall that the random walk on a half line $\boldsymbol{\Phi}$ is constructed from a sequence of i.i.d. random variables $\{W_i\}$ taking values in $\mathbb{Z} = (\ldots, -2, -1, 0, 1, 2, \ldots)$, by setting

$$\Phi_n = [\Phi_{n-1} + W_n]^+. \tag{4.2}$$

We know from Section 3.3.2 that this construction gives, for $y \in \mathbb{Z}_+$,

$$\begin{aligned} P(x,y) &= \mathsf{P}(W_1 = y - x), \\ P(x,0) &= \mathsf{P}(W_1 \le -x). \end{aligned} \tag{4.3}$$

In this example, we might single out the set $\{0\}$ and ask: can the chain ever reach the state $\{0\}$?

It is transparent from the definition of $P(x,0)$ that $\{0\}$ can be reached with positive probability, and in one step, provided the distribution Γ of the increment $\{W_n\}$ has an infinite negative tail. But suppose we have, not such a long tail, but only $\mathsf{P}(W_n < 0) > 0$, with, say,

$$\Gamma(w) = \delta > 0 \tag{4.4}$$

for some $w < 0$. Then we have for any x that after $n = [x/w]$ steps,

$$\mathsf{P}_x(\Phi_n = 0) \geq \mathsf{P}(W_1 = w, W_2 = w, \dots, W_n = w) = \delta^n > 0$$

so that $\{0\}$ is always reached with positive probability.

On the other hand, if $\mathsf{P}(W_n < 0) = 0$ then it is equally clear that $\{0\}$ cannot be reached with positive probability from any starting point other than 0. Hence $L(x, 0) > 0$ for all states x or for none, depending on whether (4.4) holds or not.

But we might also focus on points other than $\{0\}$, and it is then possible that a number of different sorts of behavior may occur, depending on the distribution of W. If we have $\mathsf{P}(W = y) > 0$ for all $y \in \mathbb{Z}$ then from any state there is positive probability of $\boldsymbol{\Phi}$ reaching any other state at the next step. But suppose we have the distribution of the increments $\{W_n\}$ concentrated on the even integers, with

$$\mathsf{P}(W = 2y) > 0, \qquad \mathsf{P}(W = 2y + 1) = 0, \qquad y \in \mathbb{Z},$$

and consider any odd valued state, say w. In this case w cannot be reached from any even valued state, even though from w itself it is possible to reach every state with positive probability, via transitions of the chain through $\{0\}$.

Thus for this rather trivial example, we already see X breaking into two subsets with substantially different behavior: writing $\mathbb{Z}_+^0 = \{2y, y \in \mathbb{Z}_+\}$ and $\mathbb{Z}_+^1 = \{2y + 1, y \in \mathbb{Z}_+\}$ for the set of non-negative even and odd integers respectively, we have

$$\mathbb{Z}_+ = \mathbb{Z}_+^0 \cup \mathbb{Z}_+^1,$$

and from $y \in \mathbb{Z}_+^1$, every state may be reached, whilst for $y \in \mathbb{Z}_+^0$, only states in $\mathbb{Z}_+^0$ may be reached with positive probability.

Why are these questions of importance?

As we have already seen, the random walk on a half line above is one with many applications: recall that the transition matrices of $\mathbf{N} = \{N_n\}$ and $\mathbf{N}^* = \{N_n^*\}$, the chains introduced in Section 2.4.2 to describe the number of customers in GI/M/1 and M/G/1 queues, have exactly the structure described by (4.3).

The question of reaching $\{0\}$ is then clearly one of considerable interest, since it represents exactly the question of whether the queue will empty with positive probability. Equally, the fact that when $\{W_n\}$ is concentrated on the even integers (representing some degenerate form of batch arrival process) we will always have an even number of customers has design implications for number of servers (do we always want to have two?), waiting rooms and the like.

But our efforts should and will go into finding conditions to preclude such oddities, and we turn to these in the next section, where we develop the concepts of communication and irreducibility in the countable space context.

4.1.2 Communicating classes and irreducibility

The idea of a Markov chain $\boldsymbol{\Phi}$ reaching sets or points is much simplified when X is countable and the behavior of the chain is governed by a transition probability matrix $P = P(x, y)$, $x, y \in \mathsf{X}$. There are then a number of essentially equivalent ways of defining the operation of communication between states.

The simplest is to say that state x *leads to state* y, which we write as $x \to y$, if $L(x, y) > 0$, and that two distinct states x and y in X *communicate*, written $x \leftrightarrow y$, when $L(x, y) > 0$ and $L(y, x) > 0$. By convention we also define $x \to x$.

The relation $x \leftrightarrow y$ is often defined equivalently by requiring that there exists $n(x,y) \geq 0$ and $m(y,x) \geq 0$ such that $P^n(x,y) > 0$ and $P^m(y,x) > 0$; that is, $\sum_{n=0}^{\infty} P^n(x,y) > 0$ and $\sum_{n=0}^{\infty} P^n(y,x) > 0$.

Proposition 4.1.1 *The relation "$\leftrightarrow$" is an equivalence relation, and so the equivalence classes $C(x) = \{y : x \leftrightarrow y\}$ cover X, with $x \in C(x)$.*

PROOF By convention $x \leftrightarrow x$ for all x. By the symmetry of the definition, $x \leftrightarrow y$ if and only if $y \leftrightarrow x$.

Moreover, from the Chapman-Kolmogorov relationships (3.24) we have that if $x \leftrightarrow y$ and $y \leftrightarrow z$ then $x \leftrightarrow z$. For suppose that $x \to y$ and $y \to z$, and choose $n(x,y)$ and $m(y,z)$ such that $P^n(x,y) > 0$ and $P^m(y,z) > 0$. Then we have from (3.24)

$$P^{n+m}(x,z) \geq P^n(x,y)P^m(y,z) > 0$$

so that $x \to z$: the reverse direction is identical. □

Chains for which all states communicate form the basis for future analysis.

Irreducible Spaces and Absorbing Sets

If $C(x) = \mathsf{X}$ for some x, then we say that X (or the chain $\{X_n\}$) is *irreducible*.

We say $C(x)$ is *absorbing* if $P(y, C(x)) = 1$ for all $y \in C(x)$.

When states do not all communicate, then although each state in $C(x)$ communicates with every other state in $C(x)$, it is possible that there are states $y \in [C(x)]^c$ such that $x \to y$. This happens, of course, if and only if $C(x)$ is not absorbing.

Suppose that X is not irreducible for $\boldsymbol{\Phi}$. If we reorder the states according to the equivalence classes defined by the communication operation, and if we further order the classes with absorbing classes coming first, then we have a decomposition of P such as that depicted in Figure 4.1.

Here, for example, the blocks $C(1)$, $C(2)$ and $C(3)$ correspond to absorbing classes, and block D contains those states which are not contained in an absorbing class. In the extreme case, a state in D may communicate only with itself, although it must lead to some other state from which it does not return. We can write this decomposition as

$$\mathsf{X} = \left(\sum_{x \in I} C(x) \right) \cup D \tag{4.5}$$

where the sum is of disjoint sets.

This structure allows chains to be analyzed, at least partially, through their constituent irreducible classes. We have

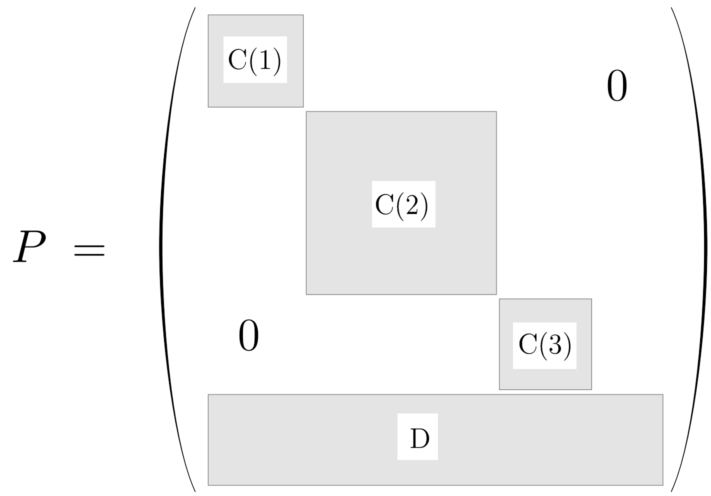

Fig. 4.1. Block decomposition of P into communicating classes

Proposition 4.1.2 *Suppose that $C := C(x)$ is an absorbing communicating class for some $x \in \mathsf{X}$. Let P_C denote the matrix P restricted to the states in C. Then there exists an irreducible Markov chain $\boldsymbol{\Phi}_C$ whose state space is restricted to C and whose transition matrix is given by P_C.*

Proof We merely need to note that the elements of P_C are positive, and

$$\sum_{y \in C} P(x, y) \equiv 1, \qquad x \in C$$

because C is absorbing: the existence of $\boldsymbol{\Phi}_C$ then follows from Theorem 3.2.1, and irreducibility of $\boldsymbol{\Phi}_C$ is an obvious consequence of the communicating class structure of C. □

Thus for non-irreducible chains, we can analyze at least the absorbing subsets in the decomposition (4.5) as separate chains.

The virtue of the block decomposition described above lies largely in this assurance that any chain on a countable space can be studied assuming irreducibility. The "irreducible absorbing" pieces $C(x)$ can then be put together to deduce most of the properties of a reducible chain.

Only the behavior of the remaining states in D must be studied separately, and in analyzing stability D may often be ignored. For let J denote the indices of the states for which the communicating classes are not absorbing. If the chain starts in $D = \bigcup_{y \in J} C(y)$, then one of two things happens: either it reaches one of the absorbing sets $C(x), x \in \mathsf{X} \backslash J$, in which case it gets absorbed: or, as the only other alternative, the chain leaves every finite subset of D and "heads to infinity".

To see why this might hold, observe that, for any fixed $y \in J$, there is some state $z \in C(y)$ with $P(z, [C(y)]^c) = \delta > 0$ (since $C(y)$ is not an absorbing class), and $P^m(y, z) = \beta > 0$ for some $m > 0$ (since $C(y)$ is a communicating class). Suppose that in fact the chain returns a number of times to y: then, on each of these returns, one has a probability greater than $\beta\delta$ of leaving $C(y)$ exactly $m + 1$ steps later, and this probability is independent of the past due to the Markov property.

Now, as is well known, if one tosses a coin with probability of a head given by $\beta\delta$ infinitely often, then one eventually actually gets a head: similarly, one eventually leaves the class $C(y)$, and because of the nature of the relation $x \leftrightarrow y$, one never returns.

Repeating this argument for any finite set of states in D indicates that the chain leaves such a finite set with probability one.

There are a number of things that need to be made more rigorous in order for this argument to be valid: the forgetfulness of the chain at the random time of returning to y, giving the independence of the trials, is a form of the Strong Markov Property in Proposition 3.4.6, and the so-called "geometric trials argument" must be formalized, as we will do in Proposition 8.3.1 (iii).

Basically, however, this heuristic sketch is sound, and shows the directions in which we need to go: we find absorbing irreducible sets, and then restrict our attention to them, with the knowledge that the remainder of the states lead to clearly understood and (at least from a stability perspective) somewhat irrelevant behavior.

4.1.3 Irreducible models on a countable space

Some specific models will illustrate the concepts of irreducibility. It is valuable to notice that, although in principle irreducibility involves P^n for all n, in practice we usually find conditions only on P itself that ensure the chain is irreducible.

The forward recurrence time model Let p be the increment distribution of a renewal process on $\mathbb{Z}_+$, and write

$$r = \sup(n : p(n) > 0). \tag{4.6}$$

Then from the definition of the forward recurrence chain it is immediate that the set $A = \{1, 2, \dots, r\}$ is absorbing, and the forward recurrence chain restricted to A is irreducible: for if $x, y \in A$, with $x > y$ then $P^{x-y}(x, y) = 1$ whilst

$$P^{y+r-x}(y, x) > P^{y-1}(y, 1)p(r)P^{r-x}(r, x) = p(r) > 0. \tag{4.7}$$

Queueing models Consider the number of customers $\mathbf{N}$ in the GI/M/1 queue. As shown in Proposition 3.3.1, we have $P(x, x+1) = p_0 > 0$, and so the structure of $\mathbf{N}$ ensures that by iteration, for any $x > 0$

$$P^x(0, x) > P(0, 1)P(1, 2)\dots P(x-1, x) = [p_0]^x > 0.$$

But we also have $P(x, 0) > 0$ for any $x \geq 0$: hence we conclude that for any pair $x, y \in \mathsf{X}$, we have

$$P^{y+1}(x, y) > P(x, 0)P^y(0, y) > 0.$$

Thus the chain $\mathbf{N}$ is irreducible no matter what the distribution of the interarrival times.

A similar approach shows that the embedded chain $\mathbf{N}^*$ of the M/G/1 queue is always irreducible.

Unrestricted random walk Let d be the greatest common divisor of $\{n : \Gamma(n) > 0\}$. If we have a random walk on $\mathbb{Z}$ with increment distribution Γ, each of the sets $D_r = \{md + r, m \in \mathbb{Z}\}$ for each $r = 0, 1, \ldots, d-1$ is absorbing, so that the chain is not irreducible.

However, provided $\Gamma(-\infty, 0) > 0$ and $\Gamma(0, \infty) > 0$ the chain is irreducible when restricted to any one D_r. To see this we can use Lemma D.7.4: since $\Gamma(md) > 0$ for all $m > m_0$ we only need to move m_0 steps to the left and then we can reach all states in D_r above our starting point in one more step. Hence this chain admits a finite number of irreducible absorbing classes.

For a different type of behavior, let us suppose we have an increment distribution on the integers, $\mathsf{P}(W_n = x) > 0$, $x \in \mathbb{Z}$, so that $d = 1$; but assume the chain itself is defined on the whole set of rationals $\mathbb{Q}$.

If we start at a value $q \in \mathbb{Q}$ then $\boldsymbol{\Phi}$ "lives" on the set $C(q) = \{n + q, n \in \mathbb{Z}\}$, which is both absorbing and irreducible: that is, we have $P(q, C(q)) = 1, q \in \mathbb{Q}$, and for any $r \in C(q)$, $P(r, q) > 0$ also.

Thus this chain admits a countably infinite number of absorbing irreducible classes, in contrast to the behavior of the chain on the integers.

4.2 ψ-Irreducibility

4.2.1 The concept of φ-irreducibility

We now wish to develop similar concepts of irreducibility on a general space X. The obvious problem with extending the ideas of Section 4.1.2 is that we cannot define an analogue of "$\leftrightarrow$", since, although we can look at $L(x, A)$ to decide whether a set A is reached from a point x with positive probability, we cannot say in general that we return to single states x.

This is particularly the case for models such as the linear models for which the n-step transition laws typically have densities; and even for some of the models such as storage models where there is a distinguished reachable point, there are usually no other states to which the chain returns with positive probability.

This means that we cannot develop a decomposition such as (4.5) based on a countable equivalence class structure: and indeed the question of existence of a so-called "Doeblin decomposition"

$$\mathsf{X} = \left(\sum_{x \in I} C(x) \right) \cup D, \tag{4.8}$$

with the sets $C(x)$ being a countable collection of absorbing sets in $\mathcal{B}(\mathsf{X})$ and the "remainder" D being a set which is in some sense ephemeral, is a non-trivial one. We shall not discuss such reducible decompositions in this book although, remarkably, under a variety of reasonable conditions such a countable decomposition does hold for chains on quite general state spaces.

Rather than developing this type of decomposition structure, it is much more fruitful to concentrate on irreducibility analogues. The one which forms the basis for much modern general state space analysis is φ-irreducibility.

φ-Irreducibility for general space chains

We call $\boldsymbol{\Phi} = \{\Phi_n\}$ *φ-irreducible* if there exists a measure φ on $\mathcal{B}(\mathsf{X})$ such that, whenever $\varphi(A) > 0$, we have $L(x, A) > 0$ for all $x \in \mathsf{X}$.

There are a number of alternative formulations of φ-irreducibility. Define the transition kernel

$$K_{a_{\frac{1}{2}}}(x, A) := \sum_{n=0}^{\infty} P^n(x, A) 2^{-(n+1)}, \qquad x \in \mathsf{X},\ A \in \mathcal{B}(\mathsf{X}); \tag{4.9}$$

this is a special case of the resolvent of $\boldsymbol{\Phi}$ introduced in Section 3.4.2, and which we consider in Section 5.5.1 in more detail. The kernel $K_{a_{\frac{1}{2}}}$ defines for each x a probability measure equivalent to $I(x, A) + U(x, A) = \sum_{n=0}^{\infty} P^n(x, A)$, which may be infinite for many sets A.

Proposition 4.2.1 *The following are equivalent formulations of φ-irreducibility:*

(i) *for all $x \in \mathsf{X}$, whenever $\varphi(A) > 0$, $U(x, A) > 0$;*

(ii) *for all $x \in \mathsf{X}$, whenever $\varphi(A) > 0$, there exists some $n > 0$, possibly depending on both A and x, such that $P^n(x, A) > 0$;*

(iii) *for all $x \in \mathsf{X}$, whenever $\varphi(A) > 0$ then $K_{a_{\frac{1}{2}}}(x, A) > 0$.*

PROOF The only point that needs to be proved is that if $L(x, A) > 0$ for all $x \in A^c$ then, since $L(x, A) = P(x, A) + \int_{A^c} P(x, dy) L(y, A)$, we have $L(x, A) > 0$ for all $x \in \mathsf{X}$: thus the inclusion of the zero-time term in $K_{a_{\frac{1}{2}}}$ does not affect the irreducibility. □

We will use these different expressions of φ-irreducibility at different times without further comment.

4.2.2 Maximal irreducibility measures

Although seemingly relatively weak, the assumption of φ-irreducibility precludes several obvious forms of "reducible" behavior. The definition guarantees that "big" sets (as measured by φ) are always reached by the chain with some positive probability, no matter what the starting point: consequently, the chain cannot break up into separate "reduced" pieces.

For many purposes, however, we need to know the reverse implication: that "negligible" sets B, in the sense that $\varphi(B) = 0$, are avoided with probability one from most starting points. This is by no means the case in general: any non-trivial restriction of an irreducibility measure is obviously still an irreducibility measure, and such restrictions can be chosen to give zero weight to virtually any selected part of the space.

For example, on a countable space if we only know that $x \to x^*$ for every x and some specific state $x^* \in \mathsf{X}$, then the chain is δ_{x^*}-irreducible.

This is clearly rather weaker than normal irreducibility on countable spaces, which demands two-way communication. Thus we now look to measures which are extensions, not restrictions, of irreducibility measures, and show that the φ-irreducibility condition extends in such a way that, if we do have an irreducible chain in the sense of Section 4.1, then the natural irreducibility measure (namely counting measure) is generated as a "maximal" irreducibility measure.

The maximal irreducibility measure will be seen to define the range of the chain much more completely than some of the other more arbitrary (or pragmatic) irreducibility measures one may construct initially.

Proposition 4.2.2 *If $\boldsymbol{\Phi}$ is φ-irreducible for some measure φ, then there exists a probability measure ψ on $\mathcal{B}(\mathsf{X})$ such that*

(i) *$\boldsymbol{\Phi}$ is ψ-irreducible;*

(ii) *for any other measure φ', the chain $\boldsymbol{\Phi}$ is φ'-irreducible if and only if $\psi \succ \varphi'$;*

(iii) *if $\psi(A) = 0$, then $\psi\{y : L(y, A) > 0\} = 0$;*

(iv) *the probability measure ψ is equivalent to*

$$\psi'(A) := \int_{\mathsf{X}} \varphi'(dy) K_{a_{\frac{1}{2}}}(y, A),$$

for any finite irreducibility measure φ'.

Proof Since any probability measure which is equivalent to the irreducibility measure φ is also an irreducibility measure, we can assume without loss of generality that $\varphi(\mathsf{X}) = 1$. Consider the measure ψ constructed as

$$\psi(A) := \int_{\mathsf{X}} \varphi(dy) K_{\frac{1}{2}}(y, A). \tag{4.10}$$

It is obvious that ψ is also a probability measure on $\mathcal{B}(\mathsf{X})$. To prove that ψ has all the required properties, we use the sets

$$\bar{A}(k) = \left\{ y : \sum_{n=1}^{k} P^n(y, A) > k^{-1} \right\}.$$

The stated properties now involve repeated use of the Chapman-Kolmogorov equations. To see (i), observe that when $\psi(A) > 0$, then from (4.10), there exists some k such that $\varphi(\bar{A}(k)) > 0$, since $\bar{A}(k) \uparrow \left\{ y : \sum_{n \geq 1} P^n(y, A) > 0 \right\} = \mathsf{X}$. For any fixed x, by φ-irreducibility there is thus some m such that $P^m(x, \bar{A}(k)) > 0$. Then we have

$$\sum_{n=1}^{k} P^{m+n}(x, A) = \int_{\mathsf{X}} P^m(x, dy) \Big(\sum_{n=1}^{k} P^n(y, A) \Big) \geq k^{-1} P^m(x, \bar{A}(k)) > 0,$$

which establishes ψ-irreducibility.

Next let φ' be such that $\boldsymbol{\Phi}$ is φ'-irreducible. If $\varphi'(A) > 0$, we have $\sum_n P^n(y, A) > 0$ for all y, and by its definition $\psi(A) > 0$, whence $\psi \succ \varphi'$. Conversely, suppose that the chain is ψ-irreducible and that $\psi \succ \varphi'$. If $\varphi'\{A\} > 0$ then $\psi\{A\} > 0$ also, and by ψ-irreducibility it follows that $K_{a_{\frac{1}{2}}}(x, A) > 0$ for any $x \in \mathsf{X}$. Hence $\boldsymbol{\Phi}$ is φ'-irreducible, as required in (ii).

Result (iv) follows from the construction (4.10) and the fact that any two maximal irreducibility measures are equivalent, which is a consequence of (ii).

Finally, we have that

$$\int_{\mathsf{X}} \psi(dy) P^m(y, A) 2^{-m} = \int_{\mathsf{X}} \varphi(dy) \sum_n P^{m+n}(y, A) 2^{-(n+m+1)} \leq \psi(A)$$

from which the property (iii) follows immediately. □

Although there are other approaches to irreducibility, we will generally restrict ourselves, in the general space case, to the concept of φ-irreducibility; or rather, we will seek conditions under which it holds. We will consistently use ψ to denote an arbitrary maximal irreducibility measure for $\boldsymbol{\Phi}$.

ψ-Irreducibility Notation

(i) The Markov chain is called *ψ-irreducible* if it is φ-irreducible for some φ and the measure ψ is a *maximal irreducibility* measure satisfying the conditions of Proposition 4.2.2.

(ii) We write

$$\mathcal{B}^+(\mathsf{X}) := \{A \in \mathcal{B}(\mathsf{X}) : \psi(A) > 0\}$$

for the sets of positive ψ-measure; the equivalence of maximal irreducibility measures means that $\mathcal{B}^+(\mathsf{X})$ is uniquely defined.

(iii) We call a set $A \in \mathcal{B}(\mathsf{X})$ *full* if $\psi(A^c) = 0$.

(iv) We call a set $A \in \mathcal{B}(\mathsf{X})$ *absorbing* if $P(x, A) = 1$ for $x \in A$.

The following result indicates the links between absorbing and full sets. This result seems somewhat academic, but we will see that it is often the key to showing that very many properties hold for ψ-almost all states.

Proposition 4.2.3 *Suppose that $\boldsymbol{\Phi}$ is ψ-irreducible. Then*

(i) *every absorbing set is full,*

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

PROOF If A is absorbing, then were $\psi(A^c) > 0$, it would contradict the definition of ψ as an irreducibility measure: hence A is full.

Suppose now that A is full, and set

$$B = \{y \in \mathsf{X} : \sum_{n=0}^{\infty} P^n(y, A^c) = 0\}.$$

We have the inclusion $B \subseteq A$ since $P^0(y, A^c) = 1$ for $y \in A^c$. Since $\psi(A^c) = 0$, from Proposition 4.2.2 (iii) we know $\psi(B) > 0$, so in particular B is non-empty. By the Chapman-Kolmogorov relationship, if $P(y, B^c) > 0$ for some $y \in B$, then we would have

$$\sum_{n=0}^{\infty} P^{n+1}(y, A^c) \geq \int_{B^c} P(y, dz) \Big\{ \sum_{n=0}^{\infty} P^n(z, A^c) \Big\}$$

which is positive: but this is impossible, and thus B is the required absorbing set. □

If a set C is absorbing and if there is a measure ψ for which

$$\psi(B) > 0 \Rightarrow L(x, B) > 0, \qquad x \in C$$

then we will call C an absorbing ψ-irreducible set.

Absorbing sets on a general space have exactly the properties of those on a countable space given in Proposition 4.1.2.

Proposition 4.2.4 *Suppose that A is an absorbing set. Let P_A denote the kernel P restricted to the states in A. Then there exists a Markov chain $\boldsymbol{\Phi}_A$ whose state space is A and whose transition matrix is given by P_A. Moreover, if $\boldsymbol{\Phi}$ is ψ-irreducible then $\boldsymbol{\Phi}_A$ is ψ-irreducible.*

PROOF The existence of $\boldsymbol{\Phi}_A$ is guaranteed by Theorem 3.4.1 since $P_A(x, A) \equiv 1, x \in A$. If $\boldsymbol{\Phi}$ is ψ-irreducible then A is full and the result is immediate by Proposition 4.2.3. □

The effect of these two propositions is to guarantee the effective analysis of restrictions of chains to full sets, and we shall see that this is indeed a fruitful avenue of approach.

4.2.3 Uniform accessibility of sets

Although the relation $x \leftrightarrow y$ is not a generally useful one when X is uncountable, since $P^n(x, y) = 0$ in many cases, we now introduce the concepts of "accessibility" and, more usefully, "uniform accessibility" which strengthens the notion of communication on which ψ-irreducibility is based.

We will use uniform accessibility for chains on general and topological state spaces to develop solidarity results which are almost as strong as those based on the equivalence relation $x \leftrightarrow y$ for countable spaces.

Accessibility

We say that a set $B \in \mathcal{B}(\mathsf{X})$ is *accessible* from another set $A \in \mathcal{B}(\mathsf{X})$ if $L(x, B) > 0$ for every $x \in A$;

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

$$\inf_{x \in A} L(x, B) \geq \delta; \tag{4.11}$$

and when (4.11) holds we write $A \rightsquigarrow B$.

The critical aspect of the relation "$A \rightsquigarrow B$" is that it holds uniformly for $x \in A$. In general the relation "$\rightsquigarrow$" is non-reflexive although clearly there may be sets A, B such that A is uniformly accessible from B and B is uniformly accessible from A.

Importantly, though, the relationship is transitive. In proving this we use the notation

$$U_A(x, B) = \sum_{n=1}^{\infty} {}_A P^n(x, B), \qquad x \in \mathsf{X},\ A, B \in \mathcal{B}(\mathsf{X});$$

introduced in (3.34).

Lemma 4.2.5 *If $A \rightsquigarrow B$ and $B \rightsquigarrow C$ then $A \rightsquigarrow C$.*

PROOF Since the probability of ever reaching C is greater than the probability of ever reaching C after the first visit to B, we have

$$\inf_{x \in A} U_C(x, C) \geq \inf_{x \in A} \int_B U_B(x, dy) U_C(y, C) \geq \inf_{x \in A} U_B(y, B) \inf_{x \in B} U_C(y, C) > 0$$

as required. □

We shall use the following notation to describe the communication structure of the chain.

Communicating sets

The set $\bar{A} := \{x \in \mathsf{X} : L(x, A) > 0\}$ is the set of points from which A is accessible.

The set $\bar{A}(m) := \{x \in \mathsf{X} : \sum_{n=1}^{m} P^n(x, A) \geq m^{-1}\}$.

The set $A^0 := \{x \in \mathsf{X} : L(x, A) = 0\} = [\bar{A}]^c$ is the set of points from which A is not accessible.

Lemma 4.2.6 *The set $\bar{A} = \cup_m \bar{A}(m)$, and for each m we have $\bar{A}(m) \rightsquigarrow A$.*

PROOF The first statement is obvious, whilst the second follows by noting that for all $x \in \bar{A}(m)$ we have

$$L(x, A) \geq \mathsf{P}_x(\tau_A \leq m) \geq m^{-2}.$$

□

It follows that if the chain is ψ-irreducible, then we can find a countable cover of X with sets from which any other given set A in $\mathcal{B}^+(\mathsf{X})$ is uniformly accessible, since $\bar{A} = \mathsf{X}$ in this case.

4.3 ψ-Irreducibility For Random Walk Models

One of the main virtues of ψ-irreducibility is that it is even easier to check than the standard definition of irreducibility introduced for countable chains. We first illustrate this using a number of models related to random walk.

4.3.1 Random walk on a half line

Let $\boldsymbol{\Phi}$ be a random walk on the half line $[0, \infty)$, with transition law as in Section 3.5. The communication structure of this chain is made particularly easy because of the "atom" at $\{0\}$.

Proposition 4.3.1 *The random walk on a half line $\boldsymbol{\Phi} = \{\Phi_n\}$ with increment variable W is φ-irreducible, with $\varphi(0, \infty) = 0$, $\varphi(\{0\}) = 1$, if and only if*

$$\mathsf{P}(W < 0) = \Gamma(-\infty, 0) > 0; \tag{4.12}$$

and in this case if C is compact then $C \rightsquigarrow \{0\}$.

PROOF The necessity of (4.12) is trivial. Conversely, suppose for some δ, $\varepsilon > 0$, $\Gamma(-\infty, -\varepsilon) > \delta$. Then for any n, if $x/\varepsilon < n$,

$$P^n(x, \{0\}) \geq \delta^n > 0.$$

If $C = [0, c]$ for some c, then this implies for all $x \in C$ that

$$\mathsf{P}_x(\tau_0 \leq c/\varepsilon) \geq \delta^{1+c/\varepsilon}$$

so that $C \rightsquigarrow \{0\}$ as in Lemma 4.2.6. □

It is often as simple as this to establish φ-irreducibility: it is not a difficult condition to confirm, or rather, it is often easy to set up "grossly sufficient" conditions such as (4.12) for φ-irreducibility.

Such a construction guarantees φ-irreducibility, but it does not tell us very much about the motion of the chain. There are clearly many sets other than $\{0\}$ which the chain will reach from any starting point. To describe them in this model we can easily construct the maximal irreducibility measure. By considering the motion of the chain after it reaches $\{0\}$ we see that $\boldsymbol{\Phi}$ is also ψ-irreducible, where

$$\psi(A) = \sum_n P^n(0, A) 2^{-n};$$

we have that ψ is maximal from Proposition 4.2.2.

4.3.2 Storage models

If we apply the result of Proposition 4.3.1 to the simple storage model defined by (SSM1) and (SSM2), we will establish ψ-irreducibility provided we have

$$\mathsf{P}(S_n - J_n < 0) > 0.$$

Provided there is some probability that no input takes place over a period long enough to ensure that the effect of the increment S_n is eroded, we will achieve δ_0-irreducibility in one step. This amounts to saying that we can "turn off" the input for a period longer than s whenever the last input amount was s, or that we need a positive probability of the input remaining turned off for longer than s/r. One sufficient condition for this is obviously that the distribution H have infinite tails.

Such a construction may fail without the type of conditions imposed here. If, for example, the input times are deterministic, occurring at every integer time point, and if the input amounts are always greater than unity, then we will not have an irreducible system: in fact we will have, in the terms of Chapter 9 below, an evanescent system which always avoids compact sets below the initial state.

An underlying structure as pathological as this seems intuitively implausible, of course, and is in any case easily analyzed. But in the case of content-dependent release rules, it is not so obvious that the chain is always φ-irreducible. If we assume $R(x) = \int_0^x [r(y)]^{-1} dy < \infty$ as in (2.33), then again if we can "turn off" the input process for longer than $R(x)$ we will hit $\{0\}$; so if we have

$$\mathsf{P}(T_i > R(x)) > 0$$

for all x we have a δ_0-irreducible model. But if we allow $R(x) = \infty$ as we may wish to do for some release rules where $r(x) \to 0$ slowly as $x \to 0$, which is not unrealistic, then even if the inter-input times T_i have infinite tails, this simple construction will fail. The empty state will never be reached, and some other approach is needed if we are to establish φ-irreducibility.

In such a situation, we will still get μ^{Leb}-irreducibility, where μ^{Leb} is Lebesgue measure, if the inter-input times T_i have a density with respect to μ^{Leb}: this can be determined by modifying the "turning off" construction above. Exact conditions for φ-irreducibility in the completely general case appear to be unknown to date.

4.3.3 Unrestricted random walk

The random walk on a half line, and the various applications of it in storage and queueing, have a single state reached from all initial points, which forms a natural candidate to generate an irreducibility measure. The unrestricted random walk requires more analysis, and is an example where the irreducibility measure is not formed by a simple regenerative structure.

For unrestricted random walk $\boldsymbol{\Phi}$ given by

$$\Phi_{k+1} = \Phi_k + W_{k+1},$$

and satisfying the assumption (RW1), let us suppose the increment distribution Γ of $\{W_n\}$ has an absolutely continuous part with respect to Lebesgue measure μ^{Leb} on $\mathbb{R}$, with a density γ which is positive and bounded from zero at the origin; that is, for some $\beta > 0, \delta > 0$,

$$\mathsf{P}(W_n \in A) \geq \int_A \gamma(x)\, dx,$$

and

$$\gamma(x) \geq \delta > 0, \qquad |x| < \beta.$$

Set $C = \{x : |x| \leq \beta/2\}$: if $B \subseteq C$, and $x \in C$ then

$$\begin{aligned} P(x, B) &= P(W_1 \in B - x) \\ &\geq \int_{B-x} \gamma(y)\, dy \\ &\geq \delta \mu^{\mathrm{Leb}}(B). \end{aligned}$$

But now, exactly as in the previous example, from any x we can reach C in at most $n = 2|x|/\beta$ steps with positive probability, so that μ^{Leb} restricted to C forms an irreducibility measure for the unrestricted random walk.

Such behavior might not hold without a density. Suppose we take Γ concentrated on the rationals $\mathbb{Q}$, with $\Gamma(r) > 0$, $r \in \mathbb{Q}$. After starting at a value $r \in \mathbb{Q}$ the chain $\boldsymbol{\Phi}$ "lives" on the set $\{r + q, q \in \mathbb{Q}\} = \mathbb{Q}$ so that $\mathbb{Q}$ is absorbing. But for any $x \in \mathbb{R}$ the set $\{x + q, q \in \mathbb{Q}\} = x + \mathbb{Q}$ is also absorbing, and thus we can produce, for this random walk on $\mathbb{R}$, an uncountably infinite number of absorbing irreducible sets.

It is precisely this type of behavior we seek to exclude for chains on a general space, by introducing the concepts of ψ-irreducibility above.

4.4 ψ-Irreducible Linear Models

4.4.1 Scalar models

Let us consider the scalar autoregressive AR(k) model

$$Y_n = \alpha_1 Y_{n-1} + \alpha_2 Y_{n-2} + \ldots + \alpha_k Y_{n-k} + W_n,$$

where $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, as defined in (AR1). If we assume the Markovian representation in (2.1), then we can determine conditions for ψ-irreducibility very much as for random walk.

In practice the condition most likely to be adopted is that the innovation process W has a distribution Γ with an everywhere positive density. If the innovation process is Gaussian, for example, then clearly this condition is satisfied. We will see below, in the more general Proposition 4.4.3, that the chain is then μ^{Leb}-irreducible regardless of the values of $\alpha_1, \ldots, \alpha_k$.

It is however not always sufficient for φ-irreducibility to have a density only positive in a neighborhood of zero. For suppose that $\mathbf{W}$ is uniform on $[-1, 1]$, and that $k = 1$ so we have a first order autoregression. If $|\alpha_1| \leq 1$ the chain will be $\mu^{\mathrm{Leb}}_{[-1,1]}$-irreducible under such a density condition: the argument is the same as for the random walk. But if $|\alpha_1| > 1$, then once we have an initial state larger than $(|\alpha_1| - 1)^{-1}$, the chain will monotonically "explode" towards infinity and will not be irreducible.

This same argument applies to the general model (2.1) if the zeros of the polynomial $A(z) = 1 - \alpha_1 z^1 - \cdots - \alpha_k z^k$ lie outside of the closed unit disk in the complex plane $\mathbb{C}$. In this case $Y_n \to 0$ as $n \to \infty$ when W_n is set equal to zero, and from this observation it follows that it is possible for the chain to reach $[-1, 1]$ at some time in the future from every initial condition. If some root of $A(z)$ lies within the open unit disk in $\mathbb{C}$ then again "explosion" will occur and the chain will not be irreducible.

Our argument here is rather like that in the dam model, where we considered deterministic behavior with the input "turned off". We need to be able to drive the chain deterministically towards a center of the space, and then to be able to ensure that the random mechanism ensures that the behavior of the chain from initial conditions in that center are comparable.

We formalize this for multidimensional linear models in the rest of this section.

4.4.2 Communication for linear control models

Recall that the linear control model LCM(F,G) defined in (LCM1) by $x_{k+1} = Fx_k + Gu_{k+1}$ is called *controllable* if for each pair of states $x_0, x^\star \in \mathsf{X}$, there exists $m \in \mathbb{Z}_+$ and a sequence of control variables $(u_1^\star, \ldots u_m^\star) \in \mathbb{R}^p$ such that $x_m = x^\star$ when $(u_1, \ldots u_m) = (u_1^\star, \ldots u_m^\star)$, and the initial condition is equal to x_0.

This is obviously a concept of communication between states for the deterministic model: we can choose the inputs u_k in such a way that all states can be reached from any starting point. We first analyze this concept for the deterministic control model then move on to the associated linear state space model LSS(F,G), where we see that controllability of LCM(F,G) translates into ψ-irreducibility of LSS(F,G) under appropriate conditions on the noise sequence.

For the LCM(F,G) model it is possible to decide explicitly using a finite procedure when such control can be exerted. We use the following rank condition for the pair of matrices (F, G):

Controllability for the Linear Control Model

Suppose that the matrices F and G have dimensions $n \times n$ and $n \times p$, respectively.

(LCM3) The matrix

$$C_n := [F^{n-1}G \mid \cdots \mid FG \mid G] \tag{4.13}$$

is called the *controllability matrix*, and the pair of matrices (F, G) is called *controllable* if the controllability matrix C_n has rank n.

It is a consequence of the Cayley Hamilton Theorem, which states that any power F^k is equal to a linear combination of $\{I, F, \dots, F^{n-1}\}$, where n is equal to the dimension of F (see [39] for details), that (F, G) is controllable if and only if

$$[F^{k-1}G \mid \cdots \mid FG \mid G]$$

has rank n for some $k \in \mathbb{Z}_+$.

Proposition 4.4.1 *The linear control model LCM(F,G) is controllable if the pair (F, G) satisfy the rank condition (LCM3).*

PROOF When this rank condition holds it is straightforward that in the LCM(F,G) model any state can be reached from any initial condition in k steps using some control sequence $(u_1, \dots, u_k)$, for we have by

$$x_k = F^k x_0 + [F^{k-1}G \mid \cdots \mid FG \mid G] \begin{pmatrix} u_1 \\ \vdots \\ u_k \end{pmatrix} \tag{4.14}$$

and the rank condition implies that the range space of the matrix $[F^{k-1}G \mid \cdots \mid FG \mid G]$ is equal to $\mathbb{R}^n$. □

This gives us as an immediate application

Proposition 4.4.2 *The autoregressive AR(k) model may be described by a linear control model (LCM1), which can always be constructed so that it is controllable.*

PROOF For the linear control model associated with the autoregressive model described by (2.1), the state process $\mathbf{x}$ is defined inductively by

$$x_n = \begin{bmatrix} \alpha_1 & \cdots & \cdots & \alpha_k \\ 1 & & & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \end{bmatrix} x_{n-1} + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u_n,$$

and we can compute the controllability matrix C_n of (LCM3) explicitly:

$$C_n = [F^{n-1}G \mid \cdots \mid FG \mid G] = \begin{bmatrix} \eta_{k-1} & \cdots & \eta_2 & \eta_1 & 1 \\ \vdots & \cdot & & 1 & 0 \\ \eta_2 & & \cdot & & \vdots \\ \eta_1 & 1 & & & \vdots \\ 1 & 0 & \cdots & \cdots & 0 \end{bmatrix}$$

where we define $\eta_0 = 1$, $\eta_i = 0$ for $i < 0$, and for $j \geq 2$,

$$\eta_j = \sum_{i=1}^{k} \alpha_i \eta_{j-i}.$$

The triangular structure of the controllability matrix now implies that the linear control system associated with the AR(k) model is controllable. □

4.4.3 Gaussian linear models

For the LSS(F,G) model

$$X_{k+1} = FX_k + GW_{k+1}$$

described by (LSS1) and (LSS2) to be ψ-irreducible, we now show that it is sufficient that the associated LCM(F,G) model be controllable and the noise sequence $\mathbf{W}$ have a distribution that in effect allows a full cross-section of the possible controls to be chosen. We return to the general form of this in Section 6.3.2 but address a specific case of importance immediately. The Gaussian linear state space model is described by (LSS1) and (LSS2) with the additional hypothesis

Disturbance for the Gaussian state space model

(LSS3) The noise variable W has a Gaussian distribution on $\mathbb{R}^p$ with zero mean and unit variance: that is, $W \sim N(0, I)$, where I is the $p \times p$ identity matrix.

If the dimension p of the noise were the same as the dimension n of the space, and if the matrix G were full rank, then the argument for scalar models in Section 4.4 would immediately imply that the chain is μ^{Leb}-irreducible. In more general situations we use controllability to ensure that the chain is μ^{Leb}-irreducible.

Proposition 4.4.3 *Suppose that the LSS(F,G) model is Gaussian and the associated control model is controllable.*

Then the LSS(F,G) model is φ-irreducible for any non-trivial measure φ which possesses a density on $\mathbb{R}^n$, Lebesgue measure is a maximal irreducibility measure, and for any compact set A and any set B with positive Lebesgue measure we have $A \rightsquigarrow B$.

PROOF If we can prove that the distribution $P^k(x, \cdot)$ is absolutely continuous with respect to Lebesgue measure, and has a density which is everywhere positive on $\mathbb{R}^n$, it will follow that for any φ which is non-trivial and also possesses a density, $P^k(x, \cdot) \succ \varphi$ for all $x \in \mathbb{R}^n$: for any such φ the chain is then φ-irreducible. This argument also shows that Lebesgue measure is a maximal irreducibility measure for the chain.

Under condition (LSS3), for each deterministic initial condition $x_0 \in \mathsf{X} = \mathbb{R}^n$, the distribution of X_k is also Gaussian for each $k \in \mathbb{Z}_+$ by linearity, and so we need only to prove that $P^k(x, \cdot)$ is not concentrated on some lower dimensional subspace of $\mathbb{R}^n$. This will happen if and only if the variance of the distribution $P^k(x, \cdot)$ is of full rank for each x.

We can compute the mean and variance of X_k to obtain conditions under which this occurs. Using (4.14) and (LSS3), for each initial condition $x_0 \in \mathsf{X}$ the conditional mean of X_k is easily computed as

$$\mu_k(x_0) := \mathsf{E}_{x_0}[X_k] = F^k x_0 \tag{4.15}$$

and the conditional variance of X_k is given independently of x_0 by

$$\Sigma_k := \mathsf{E}_{x_0}[(X_k - \mu_k(x_0))(X_k - \mu_k(x_0))^\top] = \sum_{i=0}^{k-1} F^i G G^\top F^{i\top}. \tag{4.16}$$

Using (4.16), the variance of X_k has full rank n for some k if and only if the *controllability grammian*, defined as

$$\sum_{i=0}^{\infty} F^i G G^\top F^{i\top}, \tag{4.17}$$

has rank n. From the Cayley Hamilton Theorem again, the conditional variance of X_k has rank n for some k if and only if the pair (F, G) is controllable and, if this is the case, then one can take $k = n$.

Under (LSS1)-(LSS3), it thus follows that the k-step transition function possesses a smooth density; we have $P^k(x, dy) = p_k(x, y)dy$ where

$$p_k(x, y) = (2\pi|\Sigma_k|)^{-k/2} \exp\{-\tfrac{1}{2}(y - F^k x)^\top \Sigma_k^{-1}(y - F^k x)\} \tag{4.18}$$

and $|\Sigma_k|$ denotes the determinant of the matrix Σ_k. Hence $P^k(x, \cdot)$ has a density which is everywhere positive, as required, and this implies finally that for any compact set A and any set B with positive Lebesgue measure we have $A \rightsquigarrow B$. □

Assuming, as we do in the result above, that W has a density which is everywhere positive is clearly something of a sledge hammer approach to obtaining ψ-irreducibility, even though it may be widely satisfied. We will introduce more delicate methods in Chapter 7 which will allow us to relax the conditions of Proposition 4.4.3.

Even if (F, G) is not controllable then we can obtain an irreducible process, by appropriate restriction of the space on which the chain evolves, under the Gaussian assumption. To define this formally, we let $\mathsf{X}_0 \subset \mathsf{X}$ denote the *range space* of the controllability matrix:

$$\begin{aligned}\mathsf{X}_0 &= \mathcal{R}([F^{n-1}G \mid \cdots \mid FG \mid G]) \\ &= \Big\{\sum_{i=0}^{n-1} F^i G w_i : w_i \in \mathbb{R}^p\Big\},\end{aligned}$$

which is also the range space of the controllability grammian. If $x_0 \in \mathsf{X}_0$ then so is $Fx_0 + Gw_1$ for any $w_1 \in \mathbb{R}^p$. This shows that the set X_0 is absorbing, and hence the LSS(F,G) model may be restricted to X_0.

The restricted process is then described by a linear state space model, similar to (LSS1), but evolving on the space X_0 whose dimension is strictly less than n. The matrices (F_0, G_0) which define the dynamics of the restricted process are a controllable pair, so that by Proposition 4.4.3, the restricted process is μ^{Leb}-irreducible.

4.5 Commentary

The communicating class concept was introduced in the initial development of countable chains by Kolmogorov [140] and used systematically by Feller [76] and Chung [49] in developing solidarity properties of states in such a class.

The use of ψ-irreducibility as a basic tool for general chains was essentially developed by Doeblin [65, 67], and followed up by many authors, including Doob [68], Harris [95], Chung [48], Orey [207]. Much of their analysis is considered in greater detail in later chapters. The maximal irreducibility measure was introduced by Tweedie [272], and the result on full sets is given in the form we use by Nummelin [202]. Although relatively simple they have wide-ranging implications.

Other notions of irreducibility exist for general state space Markov chains. One can, for example, require that the transition probabilities

$$K_{\frac{1}{2}}(x, \cdot) = \sum_{n=0}^{\infty} P^n(x, \cdot) 2^{-(n+1)}$$

all have the same null sets. In this case the maximal measure ψ will be equivalent to $K_{\frac{1}{2}}(x, \cdot)$ for every x. This was used by Nelson [192] and Šidák [238] to derive solidarity properties for general state space chains similar to those we will consider in Part II. This condition, though, is hard to check, since one needs to know the structure of $P^n(x, \cdot)$ in some detail; and it appears too restrictive for the minor gains it leads to.

In the other direction, one might weaken φ-irreducibility by requiring only that, whenever $\varphi(A) > 0$, we have $\sum_n P^n(x, A) > 0$ only for φ-almost all $x \in \mathsf{X}$. Whilst this expands the class of "irreducible" models, it does not appear to be noticeably more useful in practice, and has the drawback that many results are much harder to prove as one tracks the uncountably many null sets which may appear. Revuz [223] Chapter 3 has a discussion of some of the results of using this weakened form.

The existence of a block decomposition of the form

$$\mathsf{X} = \Big(\sum_{x \in I} C(x)\Big) \cup D$$

such as that for countable chains, where the sum is of disjoint irreducible sets and D is in some sense ephemeral, has been widely studied. A recent overview is in Meyn and Tweedie [182], and the original ideas go back, as so often, to Doeblin [67], after whom such decompositions are named. Orey [208], Chapter 9, gives a very accessible account of the measure-theoretic approach to the Doeblin decomposition.

Application of results for ψ-irreducible chains has become more widespread recently, but the actual usage has suffered a little because of the somewhat inadequate available discussion in the literature of practical methods of verifying ψ-irreducibility. Typically the assumptions are far too restrictive, as is the case in assuming that innovation processes have everywhere positive densities or that accessible regenerative atoms exist (see for example Laslett et al [153] for simple operations research models, or Tong [267] in time series analysis).

The detailed analysis of the linear model begun here illustrates one of the recurring themes of this book: the derivation of stability properties for stochastic models by consideration of the properties of analogous controlled deterministic systems. The methods described here have surprisingly complete generalizations to nonlinear models. We will come back to this in Chapter 7 when we characterize irreducibility for the NSS(F) model using ideas from nonlinear control theory.

Irreducibility, whilst it is a cornerstone of the theory and practice to come, is nonetheless rather a mundane aspect of the behavior of a Markov chain. We now explore some far more interesting consequences of the conditions developed in this chapter.