The main reference is Resnick Secs. 2.12-2.15, which extends to countably infinite Markov chains as well as finite Markov chains.

Karlin & Taylor Secs. 3.1 & 3.2: use specifically finite-state arguments based on matrices and Perron-Frobenius Theorem.

Lawler Ch. 1 explains, without details, the intuition for the results for the finite-state case.

We will give details for the proof that an irreducible Markov chain will have a stationary distribution provided that there exists at least one state $j$ in the Markov chain such that:

\[ A \) \left( \mathbb{E} \left[ T_j (1) \mid X_0 = j \right] < \infty \right) \]

\[ \text{push of first return to state } j \]

A standard but tedious argument shows that assumption $A$ always holds true for a finite state Markov chain (argument is outlined in Problem 1.7 of Lawler). But we will formulate the arguments to apply to countable state Markov chains which satisfy assumption $A$.

The strategy of the proof will be to hypothesize and then verify that the following vector is indeed a stationary distribution for the Markov chain (where $j$ is a fixed reference state):

\[ \pi_k = \left( \frac{\mathbb{E} \left[ T_{j,k} \right] \mathbb{I} \left( X_0 = j \right)}{\mathbb{E} \left[ T_j (1) \mid X_0 = j \right]} \right) \]

Where

\[ T_{j,k} \leq \sum_{\eta=0}^{\infty} \mathbb{I} \left( X_\eta = k \right) \]

Is the number of visits that the Markov chain makes to state $k$
before visiting state $j$. 

Note that 

$$
\mathbb{P}_{j,j} = \frac{1}{\mathbb{E} \left[ T_{j,1} \mid X_0 = j \right]}
$$

This proof will actually yield the following useful corollary:

$$
\frac{\pi_k}{\pi_j} = \mathbb{E} \left[ T_{j,k} \mid X_0 = j \right]
$$

We need Assumption A for the proof primarily so that we know these quantities are finite.

(Note: $T_{j,k} \leq T_{j,1}$)

We will proceed with the proof by showing that our hypothesized stationary distribution actually satisfies the three properties of a stationary distribution:

\begin{enumerate}
\item[a)] $\pi_k \geq 0$ [✓]
\item[b)] $\sum_{k \in S} \pi_k = 1$ [✓]
\item[c)] $\mathbf{p} \cdot \mathbf{\pi} = \mathbf{\pi}$
\end{enumerate}

Proof of b):

Note that:

$$
\sum_{k \in S} \mathbb{E} T_{j,k} = \mathbb{E} T_{j,1}
$$
Now we finish the proof by verifying the third property \( c \) of the Markov chain.

Note that it is sufficient for this purpose to simply check that:

\[
\sum_{k \in S} \pi_k = \sum_{k \in S} \frac{\mathbb{E} [ T_{jk} | X_0 = j ]}{\mathbb{E} [ T_j(1) | X_0 = j ]}
\]

\[
= \sum_{k \in S} \frac{\mathbb{E} [ T_{jk} | X_0 = j ]}{\mathbb{E} [ T_j(1) | X_0 = j ]}
\]

\[
= \frac{\mathbb{E} [ \sum_{j \in S} T_{jk} | X_0 = j ]}{\mathbb{E} [ T_j(1) | X_0 = j ]}
\]

\[
= \frac{\mathbb{E} [ T_j(1) | X_0 = j ]}{\mathbb{E} [ T_j(1) | X_0 = j ]} = 1
\]

Now we finish the proof by verifying the third property \( c \) of the Markov chain.

Note that it is sufficient for this purpose to simply check that:

\[
\vec{v} \cdot \vec{p} = \vec{v} \neq \vec{0}
\]

\[
\vec{v}_k = \mathbb{E} [ T_{jk} | X_0 = j ]
\]

because \( \vec{\pi} = c \vec{v} \)

We note that this is equivalent to proving that \( \vec{v}_k \geq 0 \)

Is an invariant measure which has the properties: \( \vec{v}_k \geq 0 \)
To prove that \( \pi \) is an invariant measure, we will prepare by rewriting it more concretely in terms of the details of the trajectory of the Markov chain. This will allow us to show the desired equality by using the Markov property.

We begin by writing:

\[
\pi_j = \sum_{n=0}^{\infty} \mathbb{I}\{X_n = j, T_j(1) > n\}
\]

This is an important trick in moving a random variable from a limit of a sum (or integral) into the summand or integrand.

This will in particular allow the following commutation of expectation past the deterministic sum (can't do this for a sum with a random limit!!)

\[
\sum_{n=0}^{\infty} \mathbb{E}\left[ \mathbb{I}\{X_n = j, T_j(1) > n\} \mid X_0 = j \right]
\]

\[
= \sum_{n=0}^{\infty} \mathbb{E}\left[ \mathbb{I}\{X_n = j, T_j(1) > n\} \right] \mathbb{I}\{X_0 = j\}
\]

(interchange of infinite sum and expectation is justified because all terms are nonnegative, or by using dominated convergence based on the finite expected return time to state \( j \)).
Now we'll use a key relationship:

\[ \mathbb{E} I(B) = P(B) \]

\[ = \frac{1}{2} \left( P(B) + 0 + P(B) \right) \]

Similarly:

\[ \mathbb{E} \left[ I(B) \mid C \right] = P(B \mid C) \]

Applying this relationship we have:

\[ r_k = \sum_{n=0}^{\infty} P \left( X_n = k, T_j(1) > n \mid X_0 = j \right) \]

This is the form that it will be relatively convenient to verify the property.

Because it involves a conditional property and specific events involving where the Markov chain is (or is not) at certain epochs.

We are going to show, using our last representation of \( r_k \):

That:

\[ \sum_{k \in S} r_k P_{k,l} = r_l \quad \text{for all } l \in S \]

(equivalent to \( \delta_l \))

Let's examine the LHS:

\[ \sum_{k \in S} \sum_{n=0}^{\infty} P \left( X_n = k, T_j(1) > n \mid X_0 = j \right) P \left( X_n = l \mid X_{n+1} = k \right) \]
The reason for expressing $P_{kl}$ as we have is to try and combine this conditional probability with the other factor to get a single conditional probability that would look close to the definition of $\sqrt{\lambda}$.

The key tactical step will be to show that:

$$P(X_n = k, T_j (\Omega) \geq n | X_0 = j) P(X_{n+1} = l | X_n = k)$$

$$= P(A, B | C)$$

By establishing that:

i) When $k \neq j$: $P(X_{n+1} = l | X_n = k)$

$$= P(A | B, C)$$

And then the summand is:

$$P(B | C) P(A | B, C)$$

$$= P(A, B | C)$$

by defn cond prob

\[ \checkmark \]

\[ 3 \]

ii) When $k = j$: $P(B | C) = 0$

\[ 3 \]

So $P(A, B | C) = 0$ also

\[ \checkmark \]

Let's prove these two substeps:

i) $k \neq j$: $P(A | B, C) = P(X_{n+1} = l | X_n = k, T_j (\Omega) \geq n, X_0 = j)$

Now we apply the Markov property
Now we apply the Markov property
\[ p(x_{n+1} = l | x_n = k) \]

(i) \( k = j \):
\[ p(B_C) = p(\{x_k = k, T_j(1) > n | x_0 = j\}) \]

\[ x_n \neq j \]

\[ = 0 \quad \text{since } j = k \]

So by following our above chain of arguments:

\[ \text{LHS} = \sum_{k \in \mathcal{S}} \sum_{n=0}^{\infty} p(k \in \mathcal{S}) \]

Summing over \( A \) (partition of sample space)

\[ = \sum_{n=0}^{\infty} p(\{x_{n+1} = l, T_j(1) > n | x_0 = j\}) \]

\[ \text{RHS} = \sum_{n=0}^{\infty} p(\{x_n = l, T_j(1) > n | x_0 = j\}) \]

We need to show this is equal to

We will conduct this closing argument separately for the case where

\( l \neq j, \ l = j \)

When \( l \neq j \):

The \( n=0 \) term on the RHS is zero.

\[ \text{RHS} = \sum_{n=1}^{\infty} p(\{x_n = l, T_j(1) > n | x_0 = j\}) \]
\[ KHS = \sum_{n=1}^{\infty} P\left( X_n = \ell, j(1) > n \mid X_0 = j \right) \]

\[ = \sum_{n=0}^{\infty} P\left( X_{n+1} = \ell, j(1) > n+1 \mid X_0 = j \right) \]

Since \( X_{n+1} = \ell \) if \( j(1) \neq n+1 \)

\[ = \sum_{n=0}^{\infty} P\left( X_{n+1} = \ell, j(1) > n \mid X_0 = j \right) \]

\[ = \text{LHS} \checkmark \]

**Case \( \ell = j \):**

\[ LHS = \sum_{n=0}^{\infty} P\left( X_{n+1} = j, j(1) > n \mid X_0 = j \right) \]

\[ = \sum_{n=0}^{\infty} P\left( j(2) = n+1 \mid X_0 = j \right) \]
This completes the proof that our hypothesized stationary distribution does in fact satisfy the three properties of a stationary distribution, and therefore establishes its existence.

The proof of the uniqueness and limiting properties of stationary distributions can be found in Resnick Sec. 2.13.

His argument is based on an elegant technique that is used in modern research for establishing similar properties for very complex systems (of stochastic differential equations, etc.).

**Coupling argument:**

Suppose I have two Markov chains \( \{ X_n \} \) and \( \{ Y_n \} \) which have the same probability transition matrix but different initial probability distributions.

And suppose that I can argue that
Then by the Markov property, the probability distributions
\[ p(x_n = i, y_n = j) \]
Would eventually have to agree. Therefore, if we initialized the two Markov chains with different stationary distributions and they couple, then those stationary distributions must be the same (uniqueness). More generally, the Markov chain with arbitrary initial probability distribution would have to eventually converge to the unique stationary distribution if it couples to the Markov chain initialized with the stationary distribution (establishing the limit distribution property). Irreducibility and aperiodicity are needed to establish the coupling.