A Simple Introduction to Finite Discrete Time Markov Chains

Serge Haddad

1 LMF, ENS Paris-Saclay & CNRS & INRIA, Université Paris-Saclay, France
serge.haddad@lsv.fr

Abstract
The aim of these lecture notes is to provide to the reader the main results about finite discrete time Markov chains with a minimal background in mathematics.

Digital Object Identifier 10.4230/LIPIcs...

1 Definitions and a key property

Consider a vector \( \mathbf{v} \) over a set of indices \( S \) and \( S' \subseteq S \). Then \( \mathbf{v}[S'] \) denotes \( \sum_{i \in S'} v[i] \).

Similarly let \( \mathbf{M} \) be a matrix whose set of column indices is \( S \) and \( S' \subseteq S \). Then \( \mathbf{M}[i,S'] \) denotes \( \sum_{j \in S'} M[i,j] \).

Definition 1. A (finite) discrete time Markov chain (DTMC) \( \mathcal{C} = (S, \mathbf{P}) \) is defined by:
- A finite set of states \( S = \{1, \ldots, s\} \);
- A stochastic matrix \( \mathbf{P} \) from \( S \times S \) to \( \mathbb{R}^+ \), i.e., such that for all \( i \), \( \mathbf{P}[i,S] = 1 \).

An initialized DTMC \( (\mathcal{C}, \pi_0) \) is equipped with a initial distribution \( \pi_0 \) from \( S \) to \( \mathbb{R}^+ \), i.e., such that \( \pi_0[S] = 1 \).

The behaviour of an initialized DTMC is given by its distribution at time \( n \) (also called transient distribution), \( \pi_n \overset{\text{def}}{=} \pi_0 \mathbf{P}^n \). We say that \( \mathcal{C} \) starts from \( i \in S \) if \( \pi_0[i] = 1 \).

The graph \( G_{\mathcal{C}} \) is an abstraction of \( \mathcal{C} \) that can be seen as describing the qualitative behaviour of the DTMC.

Definition 2. Let \( \mathcal{C} \) be a DTMC. Then the directed graph \( G_{\mathcal{C}} \) is defined by:
- \( S \), its set of vertices;
- For all \( i, j \), there is an edge from \( i \) to \( j \) if \( \mathbf{P}[i,j] > 0 \).

One can recover the DTMC from its graph by labelling every edge \( (i,j) \) by \( \mathbf{P}[i,j] \) as in Figure 1. The period of a strongly connected graph is the greatest common divisor (gcd) of the lengths of its circuits. The graph is aperiodic if its period is 1. The period of a strongly connected \( G_{\mathcal{C}} \) is called the periodicity of \( \mathcal{C} \).

Definition 3. Let \( \mathcal{C} \) be a DTMC. Then \( \mathcal{C} \) is irreducible if \( G_{\mathcal{C}} \) is strongly connected.

One can also focus on the status of states.

Definition 4. Let \( \mathcal{C} \) be a DTMC and \( i \in S \).
- \( i \) is transient if there is a positive probability to never return in \( i \) when starting from \( i \). Otherwise \( i \) is recurrent.
- When \( i \) is recurrent, \( i \) is null recurrent if the mean return time to \( i \) is infinite. Otherwise \( i \) is positive recurrent.
- \( i \) is aperiodic if the gcd of the possible return times to \( i \) is 1. \( i \) is ergodic if it is positive recurrent and aperiodic.

© Serge Haddad; licensed under Creative Commons License CC-BY

Leibniz International Proceedings in Informatics

LIPIcs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany
\( \mathcal{C} \) is \textit{ergodic} (resp. transient, recurrent, null recurrent, positive recurrent, aperiodic) if all its states are ergodic (resp. transient, recurrent, null recurrent, positive recurrent, aperiodic).

\textbf{Example 5.} Figure 1 depicts the (labelled) graph of a DTMC. It has three strongly connected components (SCC), thus it is not irreducible. By definition, the restrictions of this DTMC on its two bottom strongly connected components (BSCC) \{4, 5\} and \{6, 7, 8\} are irreducible DTMC. The restriction on \{4, 5\} has period 2 while the restriction on \{6, 7, 8\} is aperiodic.

Let us recall that using Tarjan’s algorithm computing the strongly connected components of a graph, one can decide in linear time whether \( \mathcal{C} \) is irreducible. Later we will see that, when \( \mathcal{C} \) is irreducible, one can compute its periodicity in linear time (assuming a computation of gcd in constant time).

Let \((u_n)_{n \in \mathbb{N}}\) be a sequence. Then \((u_n)_{n \in \mathbb{N}}\) is Cesaro convergent to \(\ell\) if:

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{0 \leq m < n} u_m = \ell
\]

The matrix \(P^*\) of the next proposition has a natural interpretation: \(P^*[i, j]\) is the long run average probability to be in \(j\) when starting from \(i\). We will call the \textit{long run average distribution} starting from \(i\), the \(i^{th}\) row of \(P^*\).

\textbf{Proposition 6.} Let \(P\) be a stochastic matrix. Then \((P^n)_{n \in \mathbb{N}}\) is Cesaro convergent to a stochastic matrix, denoted \(P^*\) and one has:

\[
P^*P = PP^* = P^*P^* = P^*
\]

\textbf{Proof.} Let \(\tilde{P}_n \overset{\text{def}}{=} \frac{1}{n} \sum_{0 \leq m < n} P^m\) for \(n > 0\).

\(\tilde{P}_n\) is a stochastic matrix thus the sequence \((\tilde{P}_n)_{n \geq 0}\) is bounded. We are going to prove that this sequence has a single accumulation point and thus is convergent.

Pick a sequence of indices \(n_0 < n_1 < \cdots\) such that \(L \overset{\text{def}}{=} \lim_{k \to \infty} \tilde{P}_{n_k}\) exists.

\[
\tilde{P}_nP = P\tilde{P}_n = \tilde{P}_n + \frac{1}{n}(P^n - \text{Id})
\]
Applying these equalities to \( n_k \) and letting \( k \) go to \( \infty \) yields: \( LP = PL = L \).
Let \( L' \) be another limit of a subsequence of \( (\tilde{P}_n)_{n>0} \). Then: \( PL' = L'P = L' \).
By iteration, \( P^nL' = L'P^n = L' \) for all \( n \).
By linear combination, \( \tilde{P}_nL' = L'\tilde{P}_n = L' \) for all \( n \).

Applying this equality for \( n_k \) and letting \( k \) go to \( \infty \) yields \( LL' = L'L = L' \).
Swapping \( L \) and \( L' \) yields \( L'L = LL' = L \). Thus \( L' = L \).
So \( (\tilde{P}_n)_{n>0} \) is convergent and a limit of stochastic matrices is stochastic.

\section{Irreducible Aperiodic DTMC}

The analysis of irreducible aperiodic DTMC is at the core of study of DTMC. To summarize in words the main results of irreducible aperiodic DTMC: (1) the transient distribution converges toward a steady-state distribution, (2) the convergence is linear, (3) this steady-state distribution is independent of the initial distribution, (4) the \( i \)th item of this distribution is the inverse of the mean return time to \( i \), and (5) it is the single solution of a linear equation system.

In the next proposition, we use the folk arithmetical result that if \( \gcd(a_1, \ldots, a_m) = 1 \) then there exists \( n_0 \) such that for all \( n \geq n_0 \), there exist nonnegative numbers \( \alpha_1, \ldots, \alpha_m \) with \( n = \sum_{r \leq m} \alpha_r a_r \). A matrix is positive if all its items are positive.

\begin{proposition}
Let \( P \) be the stochastic matrix of an irreducible aperiodic DTMC \( C \). Then there exists \( z \in \mathbb{N} \) such that \( P^z \) is a positive matrix.
\end{proposition}

\begin{proof}
Let \( a_1, \ldots, a_m \) be the lengths of the (elementary) circuits of \( G_C \). Let \( i, j \leq s \). Since \( G_C \) is strongly connected, there exists a path \( \rho \) of length \( \ell \leq s(s-1) \) from \( i \) to \( j \) that goes through all states of \( S \).
Let \( n_0 \) be such that for all \( n \geq n_0 \), there exist \( \alpha_1, \ldots, \alpha_m \) with \( n = \sum_{r \leq m} \alpha_r a_r \). By inserting \( \rho \) for \( r \) copies of the circuit of length \( a_i \), one gets a path \( \rho' \) from \( i \) to \( j \) of length \( \ell + n \).
Thus \( z \defeq s(s-1) + n_0 \) ensures that \( P^z \) is positive.
\end{proof}

\begin{example}
Let us look at the irreducible aperiodic DTMC of the BSCC \{6,7,8\} in Figure 1. \( P \) is not positive while \( P^2 \) is positive.
\[
\begin{pmatrix}
0.09 & 0.21 & 0.7 \\
0.2 & 0.6 & 0.2 \\
0.1 & 0.26 & 0.64
\end{pmatrix} \cdot \begin{pmatrix}
0.3 & 0.7 & 0 \\
0 & 0 & 1 \\
0.2 & 0.6 & 0.2
\end{pmatrix} = \begin{pmatrix}
0.26 & 0.64 & 0.14 \\
0.2 & 0.6 & 0.2 \\
0.2 & 0.6 & 0.2
\end{pmatrix}
\]
We are now in position to characterize the status of the states of an irreducible aperiodic DTMC \( C \).
\end{example}

\begin{corollary}
All states of an irreducible aperiodic DTMC \( C \) are aperiodic.
\end{corollary}

\begin{proof}
The assertion holds due to the fact that \( P^z[i, i] > 0 \) and \( P^{z+1}[i, i] > 0 \).
\end{proof}

The two next propositions only assume irreducibility of DTMC.

\begin{proposition}
All states of an irreducible DTMC \( C \) are positive recurrent.
\end{proposition}

\begin{proof}
\end{proof}
A Simple Introduction to Finite Discrete Time Markov Chains

**Proof.** Let \( i \in S \). For all \( j \), let \( q^n_j \) be the probability that starting from \( j \), \( i \) has not been visited up to time \( n \). Then \( (q_j^n)_{n \geq 0} \) is decreasing and its limit is the probability to never return in \( i \).

For all \( j \neq i \), there is a path \( \rho_j \) from \( j \) to \( i \) with length at most \( s-1 \) and there is a circuit \( \rho_i \) of length at most \( s \) including \( i \). Let \( \varepsilon = \min_{j \in S} \Pr(\rho_j) > 0 \). We show by induction that for all \( j \neq i \), \( q^n_j \leq (1 - \varepsilon)^n \). This is true for \( n = 0 \). Assume the assertion is true up to \( n \). The probability to avoid \( i \) up to time \( (n+1)s \) is the probability to avoid \( i \) up to time \( s \) and from some arbitrary \( j \neq i \) during \( ns \) remaining time instants to avoid \( i \). Thus \( q^{(n+1)s}_j \leq (1 - \varepsilon) \max_{j^* \neq i} (q^{ns}_{j^*}) \leq (1 - \varepsilon)^{n+1} \). So \( \lim_{n \to \infty} q^n_i = 0 \) entailing that \( i \) is recurrent. Let \( V \) be the random variable corresponding to the first return to \( i \) starting from \( i \). Recall that the mean value of \( V \), \( \mathbb{E}(V) \) is equal to \( \sum_{n \geq 0} \Pr(V \geq n) \). Grouping the terms \( s \) by \( s \), one gets \( \mathbb{E}(V) \leq s \sum_{n \in \mathbb{N}} (1 - \varepsilon)^n < \infty \). Thus \( i \) is positive recurrent.

**Corollary 11.** An irreducible aperiodic DTMC \( C \) is ergodic.

**Proposition 12.** Let \( i \) be a state of an irreducible DTMC \( C \) and \( \mu \) be the mean return time to \( i \). Then \( P^*[i, i] = \frac{1}{\mu} \).

**Proof.** Let \( T \) be the random return time to \( i \).

For all \( n \in \mathbb{N} \), let \( g_n = \Pr(T > n) \). Observe that \( \mu = \sum_{n \in \mathbb{N}} g_n \).

Assume that \( C \) starts in \( i \) and define for all \( m \in \mathbb{N} \) and \( k \leq m \), the event \( Ev_{k, m} \) as: "\( k \) is the last time that \( C \) visits \( i \) in the interval \([0, m]\). \( \{Ev_{k, m}\}_{k \leq m} \) constitutes a partition of the samplings and \( \Pr(Ev_{k, m}) = \pi_k[i]g_{m-k} \).

Thus: \( \sum_{k \leq m} \pi_k[i]g_{m-k} = 1 \).

Let \( s \) be such that \( \sum_{n < s} g_n > \mu - \varepsilon \) and \( n > n_\varepsilon \).

Then from Equation 1, one gets: \( \sum_{n < s} \pi_k[i] \geq 1 \).

Letting \( n \) go to \( \infty \) one gets: \( \mu \geq 1 \).

**Proposition 13.** Let \( P \) be the stochastic matrix of an irreducible aperiodic DTMC. Then all rows of \( P^* \) are identical and \( P^* \) is a positive matrix.

**Proof.** Let \( j \leq s \) and \( i_{\text{max}} \in \operatorname{arg\,max} \{ P^*[i, j] \mid i \leq s \} \). Let \( z \) be such that \( P^z \) is a positive matrix. Recall that:

\[
P^* = P^z P^*
\]

So \( P^*[i_{\text{max}}, j] = \sum_{i \leq s} P^z[i_{\text{max}}, i]P^*[i, j] \). Since all \( P^z[i_{\text{max}}, i] \) are positive and \( P^z[i_{\text{max}}, S] = 1 \) this implies that for all \( i \), \( P^*[i_{\text{max}}, j] = P^*[i, j] \).

Due to the previous proposition for all \( j \), \( P^*[j, j] > 0 \).

Let \( i, j \in S \). \( P^*[i, j] = P^*[j, j] > 0 \). So \( P^* \) is positive.
In the sequel one denotes $\pi^*$ any row of $P^*$. The next proposition shows that $\pi^*$ can be computed in polynomial time.

**Proposition 14.** Let $C$ be an irreducible aperiodic DTMC. Then $\pi^*$ is the single solution of the linear equation system:

$$XP = X \land X \cdot 1 = 1 \land X \in \mathbb{R}^s$$ (2)

**Proof.** Since $P^*P = P^*$, $\pi^*P = \pi^*$ which establishes that $\pi^*$ is a solution of system (2).

Let $\pi$ be a solution of system (2). Then for all $n \in \mathbb{N}$, $\pi P^n = \pi$ and consequently $\pi P_n = \pi$.

Letting $n$ go to $\infty$, one gets $\pi P^* = \pi$. On the other hand for all $\pi$ such that $\pi \cdot 1 = 1$, $\pi P^* = \pi^*$ implying that $\pi = \pi^*$.

**Example 15.** Let us look at the irreducible aperiodic DTMC of the BSCC $\{6, 7, 8\}$ in Figure 1. The single solution of the following equation system is $(\begin{pmatrix} 10 \over 33 \ 28 \ 35 \over 33 \end{pmatrix})$.

$$\begin{pmatrix} 0.3 & 0.7 & 0 \\ 0 & 0 & 1 \\ 0.2 & 0.6 & 0.2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

The next proposition is the main result about irreducible aperiodic initialized DTMC and simultaneously establishes that: (1) the long run average probability to be in some state $j$ is also the limit of the probability to be in $j$ when time goes to infinity whatever the initial distribution and (2) the convergence is linear.

**Proposition 16.** Let $(C, \pi_0)$ be an irreducible aperiodic initialized DTMC. Then there exists some $0 < \lambda < 1$ such that:

$$||\pi^* - \pi_n|| = O(\lambda^n)$$

**Proof.** Since all rows of $P^*$ are identical and equal to $\pi^*$, for any stochastic matrix $P'$, $P'P^* = P^*$ and for any distribution $\pi'$, $\pi'P^* = \pi^*$.

Let $z$ be such that $P^z$ is positive. There is some $0 < \delta < 1$ such that $P^z \geq \delta P^*$. Let $\theta \overset{\text{def}}{=} 1 - \delta$ and $Q \overset{\text{def}}{=} \frac{1}{\theta}P^z - \frac{1 - \delta}{\pi - \pi^*}$. $Q$ is a stochastic matrix and fulfills:

$$P^z = \theta Q + (1 - \theta)P^*$$

Let us prove that for all $n$:

$$P^zn = \theta^n Q^n + (1 - \theta^n)P^*(\iff P^zn - P^* = \theta^n(Q^n - P^*))$$

Observe that:

$$P^*Q = \frac{1}{\theta}P^z - \frac{1 - \theta}{\theta}P^*P^* = \frac{1}{\theta}P^* - \frac{1 - \theta}{\theta}P^* = P^*$$

We proceed by recurrence:

$$P^{zn+z} = (\theta^n Q^n + (1 - \theta^n)P^*)((\theta Q + (1 - \theta)P^*))$$

$$= \theta^{n+1}Q^{n+1} + ((1 - \theta^n)\theta + (1 - \theta)\theta^n + (1 - \theta^n)(1 - \theta))P^* = \theta^{n+1}Q^{n+1} + (1 - \theta^{n+1})P^*$$

Multiplying by $P^r$ with $0 \leq r < z$, one gets that for all $n$ and $r < z$:

$$P^{zn+r} - P^* = \theta^n(Q^nP^r - P^*)$$
Multiplying by $\pi_0$, one gets that for all $n$ and $r < z$:

$$\pi_{zn+r} - \pi^* = \theta^n(\pi_0Q^nP^r - \pi^*)$$

Since $|\pi_0Q^nP^r - \pi^*|_\infty \leq 1$, the proposition is established with $\lambda = \theta^{\frac{1}{z}}$.

We call $\pi^*$ the steady-state distribution of $C$.

### 3 Irreducible Discrete Time Markov Chains

The analysis of irreducible DTMC reduces in several ways to the analysis of irreducible aperiodic DTMC based on a cyclic decomposition of the Markov chain into $p$ components where $p$ is its periodicity. To summarize in words the main results of irreducible DTMC: (1) by considering steps of $p$ transitions, each component of the DTMC behaves as an ergodic DTMC, (2) the long run average distribution is independent of the initial distribution, and (3) it is the single solution of a linear equation system.

**Proposition 17.** Let $C$ be an irreducible DTMC with periodicity $p$. Then $S = S_0 \cup S_1 \cup \ldots \cup S_{p-1}$ with:

$$\forall x < p \forall i \in S_x \forall j \in S P_{i,j} > 0 \Rightarrow j \in S_{(x+1) \mod p}$$

Furthermore $p$ is the greatest integer fulfilling this property.

**Proof.** Let $i$ be a state. For all $0 \leq x < p$, let:

$$S_x \overset{\text{def}}{=} \{ j \mid \text{there is a path in } G_C \text{ from } i \text{ to } j \text{ with length equal to } \ell p + x \text{ for some } \ell \}$$

Irreducibility of $C$ implies $S = S_0 \cup S_1 \cup \ldots \cup S_{p-1}$.

By definition, $\forall j \in S_x \forall j' \in S P_{j,j'} > 0 \Rightarrow j' \in S_{(x+1) \mod p}$.

Assume there exists $j \in S_x \cap S_{x'}$ with $x \neq x'$ and associated paths $\rho$ and $\rho'$. Since there is a path $\rho''$ from $j$ to $i$, at least one of the circuits $\rho\rho''$ and $\rho'\rho''$ has a length which is not a multiple of $p$, leading to a contradiction.

Assume there is some $p' > p$ fulfilling this property. Then the periodicity of $C$ is a multiple of $p'$, leading to another contradiction.

**Example 18.** The irreducible DTMC of Figure 2 has period 2. Its two components are drawn in red $\{1, 3, 4, 8\}$ and green $\{2, 5, 6, 7\}$.

**Proposition 19.** Let $C$ be an irreducible DTMC. Then Algorithm 1 applied to $G_C$ computes in linear time its periodicity of $C$ and its decomposition.

**Proof.** Algorithm 1 is a modification of the covering tree building by a breadth first exploration starting from vertex 1 as root. The modification consists in (1) memorizing the height of vertices in array $Height$ and (2) labelling every edge $(i, j)$ out of the tree by the integer $Height[i] - Height[j] + 1$. The value $p$ returned is the gcd of these labels (in fact one computes the gcd on the fly). At the end the algorithm computes the remainder of the heights of vertices of the division by $p$ and $u \in S_x$ iff $HeightRemainder[u] = x$.

Let $p_G$ denote the period of $G$. Given two paths with same source and destination, the difference between the lengths of these paths must be a multiple of $p_G$, using the same argument applied in the proof of Proposition 17.
Algorithm 1: Computing the periodicity and the decomposition

\textbf{Period}(G): an integer and an integer array indexed by the vertices
\textbf{Input}: $G = (S, E)$, an oriented graph with $S = \{1, \ldots, s\}$
\textbf{Output}: $p$, the period of $G$, $Height$, $HeightRemainder$ arrays of size $s$
\textbf{Data}: $i, j$ integers, $Q$ a queue

\begin{algorithmic}
\FOR{$i$ from 1 to $s$} \text{Height}[$i$] $\leftarrow \infty$
\end{algorithmic}
\begin{algorithmic}
\STATE $p \leftarrow 0$; \text{Height}[1] $\leftarrow 0$; \text{InsertQueue}(Q, 1)
\STATE // \text{perc}[1] $\leftarrow 1$
\WHILE{not \text{EmptyQueue}(Q)}
\STATE $i \leftarrow \text{ExtractQueue}(Q)$
\FOR{$(i, j) \in E$}
\IF{\text{Height}[$j$] $= \infty$}
\STATE // \text{perc}[$j$] $\leftarrow i$
\STATE \text{Height}[$j$] $\leftarrow \text{Height}[$i$] + 1$
\STATE \text{InsertQueue}(Q, j);
\ELSE $p \leftarrow \gcd(p, \text{Height}[$i$] - \text{Height}[$j$] + 1)$
\ENDIF
\ENDFOR
\FOR{$i$ from 1 to $s$} \text{HeightRemainder}[$i$] $\leftarrow \text{Height}[$i$] \mod p$
\ENDFOR
\RETURN $p, \text{HeightRemainder}$
\end{algorithmic}

\textbf{Figure 2} An irreducible DTMC with period 2.
Let \((i, j)\) be an edge out of the tree. Let us denote \(\sigma_i\), the path from 1 to \(i\) along the tree and \(\sigma_j\) the path from 1 to \(j\) along the tree. The length of \(\sigma_i\) is \(\text{Height}[i]\), the one of \(\sigma_j\) is \(\text{Height}[j]\). With the edge \((i, j)\), one obtains another path \(\sigma_i(i, j)\) from 1 to \(j\). The difference between the lengths of the two paths is: 
\[
\text{Height}[i] - \text{Height}[j] + 1
\]
which must be a multiple of \(p_G\). Since \((i, j)\) is arbitrary, one deduces that \(p_G | p\).

By construction, an edge of the tree joins a vertex of \(S_x\) to a vertex of \(S_{x+1} \mod p\). An edge \((i, j)\) out of the tree joins \(i \in S_{\text{Height}[i] \mod p}\) to \(j \in S_{\text{Height}[j] \mod p}\) but \(\text{Height}[i] - \text{Height}[j] + 1 \mod p = 0\). So, \(S_{\text{Height}[i] \mod p} = S_{\text{Height}[i+1] \mod p}\). By Proposition 17, \(p \leq p_G\). Since \(p_G | p\), one obtains \(p = p_G\).

Example 20. Figure 3 depicts the execution of Algorithm 1 on DTMC of Figure 2. There are three edges out of the tree with labels 0, 2 and 4. Thus the period of the graph is 2.

Let \(M\) be a matrix over \(S \times S\) and \(S' \subseteq S\). Then \(M_{S'}\) is the restriction of \(M\) on \(S' \times S'\). Let \(C\) be an irreducible DTMC with periodicity \(p\) and associated decomposition \((S_x)_{x<p}\). Then for all \(x<p\), \(P_{x,x+1 \mod p}\) is the restriction of \(P\) on \(S_x \times S_{x+1} \mod p\). Observe that all other items of \(P\) are null (see below).

\[
P = \begin{pmatrix}
0 & P_{0,1} & \cdots & \cdots & 0 \\
0 & 0 & P_{1,2} & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \cdots & P_{p-2,p-1} \\
P_{p-1,0} & 0 & \cdots & \cdots & 0
\end{pmatrix}
\]

Proposition 21. Let \(C\) be an irreducible DTMC with periodicity \(p\) and associated decomposition \((S_x)_{x<p}\). For all \(x<P\), let \(C_x = (S_x, P^{p}_{S_x})\). Then \(C_x\) is irreducible and aperiodic. Furthermore for all \(x<p\), let \(\pi^x\) the steady-state distribution of \(C_x\). Then for all \(x<p\), \(\pi^{(x+1 \mod p)*} = \pi^x P_{x,x+1 \mod p}\).

Proof. Since every path of length \(p\) in \(G_C\) starting from a state of \(S_x\) leads to a state of \(S_x\), \(P^{p}_{S_x}\) is a stochastic matrix.

Let \(i,j \in S_x\). Consider a path from \(i\) to \(j\) in \(G_C\). This path has length \(rp\) for some \(r\). Thus \(C_x\) is irreducible.
Let \( p_x \) be the periodicity of \( C_x \). Consider \( \rho \) a circuit of \( G_C \). Due to the decomposition of \( C \) w.l.o.g. this circuit starts from some \( s_0 \in S_i \) and every \( p \) steps visits a state \( s_j \in S_x \) until some \( s_j = s_0 \). Thus \( s_0 s_1 \ldots s_f = s_0 \) is a circuit of \( G_{C_x} \) whose length is a multiple of \( p_x \). Thus the length of \( \rho \) is a multiple of \( p_x p \). Since the gcd of the length of circuits of \( G_C \) is \( p \), this implies \( p_x = 1 \).

Observe that for all \( x < p \), \( \pi^{x \cdot s} = \pi^{x \cdot s} P^{s_{p_x}}_{S_x} = \pi^{x \cdot s} P_{x \cdot x + 1 \mod p \cdot x} \).

Thus \( \pi^{x \cdot s} P_{x \cdot x + 1 \mod p \cdot x} \cdot P_{x \cdot x + 1 \mod p \cdot x} = \pi^{x \cdot s} P_{x \cdot x + 1 \mod p \cdot x} P_{x \cdot x + 1 \mod p \cdot x} \).

By Proposition 14, this implies that \( \pi^{(x+1 \mod p) \cdot s} = \pi^{x \cdot s} P_{x \cdot x + 1 \mod p} \).

The next corollary is similar to Proposition 16 for irreducible aperiodic initialized DTMC and describes the asymptotic behaviour of an irreducible initialized DTMC.

**Corollary 22.** Let \((C, \pi_0)\) be an irreducible initialized DTMC with periodicity \( p \) and associated decomposition \((S_x)_{x < p}\). For all \( x < p \), let \( \pi^{x \cdot s} \) the steady-state distribution of \( C_x = (S_x, P^{s_{p_x}}_{S_x}) \).

Pick some \( r < p \). Define \( \pi^{*} \) by: For all \( x < p \) and all \( s \in S_x \), \( \pi^{*}[s] = \pi^{x \cdot s}[s] \pi_0[S_{x-r \mod p}] \).

Then \( \lim_{n \to \infty} \pi^{p_{r+n}} = \pi^{*} \).

Let \( C \) be an irreducible DTMC with periodicity \( p \) and associated decomposition \((S_x)_{x < p}\). For all \( 0 < x < p \), let \( \pi^{x \cdot s} \) be the steady-state distribution of \( C_x \). Let \( \pi^{*} \) the distribution over \( S \) defined by: for all \( x \) and \( s \in S_x \), \( \pi^{*}[s] = \frac{1}{p} \pi^{x \cdot s}[s] \). The next proposition generalizes Proposition 14.

**Proposition 23.** Let \( C \) be an irreducible DTMC with periodicity \( p \) and associated decomposition \((S_x)_{x < p}\). Then \( \pi^{*} \) is the single solution of the linear equation system:

\[
X P = X \wedge X \cdot 1 = 1 \wedge X \in \mathbb{R}^p
\]

**Proof.** Due to the last assertion of Proposition 21, \( \pi^{*} \) is a solution of Equation (3).

Let \( \pi \) be a solution of Equation (3). Then \( \pi P^p = \pi \). For all \( x < p \), let us denote \( \pi_x \) the projection of \( \pi \) on \( S_x \); this implies that \( \pi_x P^p_{S_x} = \pi_x \) and thus \( \pi_x = \lambda_x \pi^{x \cdot s} \) for some \( \lambda_x \). Thus the projection of \( \pi P \) on \( S_x \) is equal to \( \lambda_x \pi_x \mod p \pi^{x \cdot s} \) which implies that all \( \lambda_x \)'s are equal. Since \( \sum_{s \in S_x} \pi_x[s] = 1 \), this entails that for all \( x \), \( \lambda_x = \frac{1}{p} \).

**Example 24.** The long run average distribution of the component \( \{4, 5\} \) of DTMC in Figure 1 is \( \left(\frac{1}{2}, \frac{1}{2}\right) \).

The next corollary generalizes Proposition 13.

**Corollary 25.** Let \( P \) be the stochastic matrix of an irreducible DTMC. Then all rows of \( P^* \) are identical.

**Proof.** Let us recall that \( P^* P = P^* \). This means that every row of \( P^* \) satisfies Equation (3) and due to Proposition 23 is equal to \( \pi^{*} \).

**4 General Discrete Time Markov Chains**

As already said every BSCC of a DTMC \( C \) is an irreducible DTMC and can be analyzed independently of the other parts of \( C \). So here we focus on states of \( S \) that do not belong to
a BSCC (shown to be transient). To summarize in words the main results about transient states: (1) there is an analytical formula for the probability to reach any BSCC and based on it, (2) one can compute in polynomial time the long run average distribution starting from any transient state, and (3) when all BSCCs are ergodic, there is a steady-state distribution (that depends on the initial distribution) also computable in polynomial time.

**Definition 26.** Let \( C \) be a DTMC. Then \( T \subset S \) is the set of states defined as the complementary set of the BSCCs of \( G_C \).

**Proposition 27.** Let \( C \) be a DTMC. Then \( T \) is the set of transient states

**Proof.** We already know that the states of a BSCC are positive recurrent. Let \( i \in T \). Then in \( G_C \) there is a path from \( i \) to some \( j \) belonging to a BSCC, so with no path from \( j \) to \( i \). Thus the probability to never return in \( i \) is positive.

**Corollary 28.** Let \( C \) be a DTMC. Then \( C \) is ergodic if and only if it is a disjoint union of irreducible aperiodic DTMCs.

In the sequel, we consider the partition of \( S = T \cup \bigcup_{1 \leq x \leq b} B_x \) where every \( B_x \) is a BSCC. For sake of simplicity, \( P_{T,x} \) denotes the restriction of \( P \) on \( T \times B_x \).

**Example 29.** In the DTMC of Figure 1, \( T = \{1, 2, 3\} \), \( B_1 = \{4, 5\} \) and \( B_2 = \{6, 7, 8\} \).

\[
P_{T,1} = \begin{pmatrix} 0 & 0.3 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad P_{T,2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0 \end{pmatrix}
\]

**Proposition 30.** Let \( C \) be a DTMC. There exists \( 0 < \theta < 1 \) such that for all \( i \in S \) and all \( n \in \mathbb{N} \), \( P^n[i, T] \leq \theta^{n-s} \) and thus \( \lim_{n \to \infty} P^n[i, T] = 0 \).

**Proof.** When \( i \in B_x \) for all \( n \in \mathbb{N} \), \( P^n[i, T] = 0 \). For all \( i \in T \), there is an elementary path \( \rho_i \) of length at most \( s-1 \) starting from \( i \) and ending in some BSCC. Let \( \varepsilon = \min_{i \in T} \Pr(\rho_i) > 0 \). We show by induction that for all \( i \in T \) and \( P^{(s-1)}[i, T] \leq (1-\varepsilon)^n \). This is true for \( n = 0 \). Assume the assertion is true up to \( n \).

\[
P^{(n+1)(s-1)}[i, T] = \sum_{j \in T} P^{(s-1)}[i, j] P^{(n-1)}[j, T] \leq (1-\varepsilon)^{n} \sum_{j \in T} P^{(s-1)}[i, j] = (1-\varepsilon)^n P^{(s-1)}[i, T]
\]

Since for all \( i \), \( P^{(s-1)}[i, T] \leq 1-\varepsilon \), the assertion is proved. Defining \( \theta = (1-\varepsilon)^{\frac{1}{s-1}} \) and observing that \( P^n[i, T] \) is decreasing w.r.t. \( n \) achieves the proof.

Let \( \text{Id}_T \) be the identity matrix over \( T \).

**Proposition 31.** Let \( C \) be a DTMC. Then \( \text{Id}_T - P_T \) is invertible. Let \( i \in T \) and \( x \leq b \). Then: \( (\text{Id}_T - P_T)^{-1}P_{T,x} [i, B_x] \) is the probability to reach \( B_x \) starting from \( i \).

**Proof.** Consider the following equation:

\[
(Id_T - P_T) \sum_{n \leq m} P^n_T = Id_T - P_T^{m+1}
\]

Due to the previous proposition \( ||P^n_T|| \leq \theta^{n-s} \) for some \( 0 < \theta < 1 \) (implying the convergence of \( \sum_{n \in \mathbb{N}} P^n_T \)) and thus \( \lim_{n \to \infty} P^n_T = 0 \), letting \( m \) goes to infinity yields:

\[
(Id_T - P_T) \sum_{n \in \mathbb{N}} P^n_T = Id_T
\]
Thus \( \sum_{n \in \mathbb{N}} P^n_T = (\text{Id}_T - P_T)^{-1} \).
One can partition the reachability event of \( B_x \) starting from \( i \) depending on the time for reaching it. The probability to reach \( B_x \) starting from \( i \) at time \( n + 1 \) is the probability to stay in \( T \) until time \( n \) and then going to \( B_x \) thus equal to: \( (P^n_T P_T) [i, B_x] \).
Summing over \( n \), one gets:
\[
\left( \sum_{n \in \mathbb{N}} P^n_T \right) P_T [i, B_x]
\]
which concludes the proof.

Let \( \pi^{\ast} \) be the long run average distribution of the BSCC \( B_x \). The next propositions show that the asymptotic behaviour when starting from a transient state depends on these distributions and on the probabilities to reach the BSCCs.

**Proposition 32.** Let \( C \) be a DTMC. Let \( i \in T \). Define \( \pi^\ast \) as follows.
For all \( j \in T \), \( \pi^\ast [j] = 0 \) and for all \( x \leq b \) and all \( j \in B_x \):
\[
\pi^\ast [j] = ((\text{Id}_T - P_T)^{-1} P_T) [i, B_x] \pi^\ast [j]
\]
Then \( \pi^\ast \) is the long run average distribution starting from \( i \).

**Proof.** \( \lim_{n \to \infty} \frac{1}{n} \sum_{m < n} P^m_T = 0 \). Thus for all \( j \in T \), \( \pi^\ast [j] = 0 \) is the long run average probability to be in \( j \) when starting from \( i \).
Let \( x \leq b \) and \( j \in B_x \). Pick some \( \varepsilon > 0 \). There exists \( n_0 \) such that \( \| \sum_{m \geq n_0} P^m_T \| < \varepsilon \).
Pick some \( n_1 \) such for all \( n \geq n_1 \),
\[
\max(|\pi^{\ast}[j]| - \frac{1}{n + n_0} \sum_{m < n} P^m[k, j]|, |\pi^{\ast}[j]| - \frac{1}{n} \sum_{m < n} P^m[k, j]| | k \in B_x \| < \varepsilon
\]
Then for all \( n \geq n_0 + n_1 \):
\[
\sum_{m < n_0} \sum_{k \in B_x} (P^m_T P_T) [i, k] \frac{1}{n} \sum_{r \leq n - n_0} P^r[k, j] \leq \frac{1}{n} \sum_{m \in \mathbb{N}} P^m[i, j] \leq \sum_{m \in \mathbb{N}} (P^m_T P_T) [i, k] \frac{1}{n} \sum_{r \leq n} P^r[k, j]
\]
\[
\sum_{m \in \mathbb{N}} (P^m_T P_T) [i, B_x] \pi^{\ast}[j] - \varepsilon \leq \frac{1}{n} \sum_{m \in \mathbb{N}} P^m[i, j] \leq \sum_{m \in \mathbb{N}} (P^m_T P_T) [i, B_x] \pi^{\ast}[j] + \varepsilon
\]
\[
\sum_{m \in \mathbb{N}} (P^m_T P_T) [i, B_x] \pi^{\ast}[j] - 2 \varepsilon \leq \frac{1}{n} \sum_{m \in \mathbb{N}} P^m[i, j] \leq \sum_{m \in \mathbb{N}} (P^m_T P_T) [i, B_x] \pi^{\ast}[j] + \varepsilon
\]
These inequalities establish the result.

**Proposition 33.** Let \( C \) be a DTMC with all its BSCCs aperiodic. Let \( i \in T \) and \( \pi_0(i) = 1 \).
Then \( \lim_{n \to \infty} \pi_n \) exists and is defined by: For all \( j \in T \), \( \pi^\ast [j] = 0 \) and for all \( x \leq b \) and all \( j \in B_x \):
\[
\pi^\ast [j] = ((\text{Id}_T - P_T)^{-1} P_T) [i, B_x] \pi^\ast [j]
\]

**Proof.** The proof has already been established for \( j \in T \) by Proposition 30.
Let \( x \leq b \) and \( j \in B_x \). Pick some \( \varepsilon > 0 \). There exists \( n_0 \) such that \( \| \sum_{m \geq n_0} P^m_T \| < \varepsilon \).
Pick some \( n_1 \) such for all \( n \geq n_1 \),
\[
\max(|\pi^{\ast}[j] - P^n[k, j]| | k \in B_x \| < \varepsilon
\]
Then for all $n \geq n_0 + n_1$:

$$
\sum_{m < n_0} \sum_{k \in B_x} (P^m T P_{T,x})_{i,k} P^{n-n_0}[k,j] \leq P^n[i,j] \leq \sum_{m < n_0} \sum_{k \in B_x} (P^m T P_{T,x})_{i,k} P^{n-n_0}[k,j] + \sum_{m \geq n_0} (P^m T P_{T,x})_{i,B_x} P^n[i,j] - \epsilon \leq P^n[i,j] \leq \sum_{m \geq n_0} (P^m T P_{T,x})_{i,B_x} P^{n-n_0}[k,j] + 2\epsilon
$$

These inequalities establish the result.

\section{A Brief Overview of Infinite Discrete Time Markov Chains}

In infinite DTMC, the set $S$ may be countable. Some of the results presented above remain true but require different proofs that can be found at the following url \url{http://www.lsv.fr/~haddad/coursproba12.pdf}.

\begin{proposition}
Let $C$ be an infinite irreducible DTMC. Then all states of $C$ have the same status w.r.t (positive) recurrence and periodicity.
\end{proposition}

\begin{proposition}
Let $C$ be an infinite irreducible aperiodic DTMC. Then $C$ is ergodic if and only if there exists some distribution $\pi^*$ such that $\pi^* P = \pi^*$.
Furthermore when $C$ is ergodic, $XP = X$ admits a single distribution $\pi^*$ solution with $\pi^*[i]$ being the inverse of the mean return time to $i$.
Last $\pi^*$ is a steady state distribution whatever the initial distribution.
\end{proposition}

\begin{proposition}
Let $C$ be an infinite irreducible DTMC. Then $C$ is positive recurrent if and only if there exists some distribution $\pi^*$ such that $\pi^* P = \pi^*$.
Furthermore when $C$ is positive recurrent, $XP = X$ admits a single distribution $\pi^*$ solution.
\end{proposition}

\begin{proposition}
Let $C$ be an infinite irreducible DTMC. Then $C$ is recurrent if and only if there exists some positive vector $v$ such that $vP = v$.
Furthermore when $C$ is recurrent, $v$ is unique up to a scalar factor.
\end{proposition}

\begin{example}
We illustrate these propositions with a classical example of random walk. The DTMC $C$ of Figure 4 is irreducible and all states have period 2.
\begin{itemize}
\item $C$ is transient if $p > \frac{1}{2}$;
\item $C$ is null recurrent if $p = \frac{1}{2}$;
\item $C$ is positive recurrent if $p < \frac{1}{2}$.
\end{itemize}
\end{example}