doi: 10.2478/v10062-009-0009-7

ANNALES
UNIVERSITATIS MARIAE CURIE-SKŁODOWSKA
LUBLIN – POLONIA

VOL. LXIII, 2009 SECTIO A 91–107

MARIUSZ GÓRAJSKI

Reduction of absorbing Markov chain

Abstract. In this paper we consider an absorbing Markov chain with finite number of states. We focus especially on random walk on transient states. We present *a graph reduction method* and prove its validity. Using this method we build algorithms which allow us to determine the distribution of time to absorption, in particular we compute its moments and the probability of absorption. The main idea used in the proofs consists in observing a non-decreasing sequence of stopping times. Random walk on the initial Markov chain observed exclusively in the stopping times $\tau_1, \tau_2, \dots$ is equivalent to some new Markov chain.

1. Introduction and notation. A comprehensive study of Markov chain can be found in many monographs on the foundation of probability theory e.g. [1], [5]. Some methods of computing the probability of absorption and the moments of time to absorption are known in the literature. For example the mean value rules (see [10], [3], or [7]) or Engel's probabilistic abacus (see [3], [4] or [11]) can be used to compute expected time to absorption and probabilities of absorption. To obtain the moments of the time to absorption one can use the method based on algebraic properties of fundamental matrix for an absorbing Markov chain (see [7] Theorem 3.2, [8]). Nevertheless we present different probabilistic technique which allows us to determine not only moments of the time to absorption but also its distribution. We use a "graph reduction method". Up to now only some specific examples of

2000 *Mathematics Subject Classification.* 60J10, 60J20.
Key words and phrases. Absorbing Markov chain, distribution of time to absorption.
I wish to thank Prof. Adam Paszkiewicz for helpful discussion and his invaluable suggestions.

the graph reduction method were described (cf. [11]) and probabilities of absorption were computed for them.

In our opinion in publications cited above described techniques and algorithms demand more rigorous arguments. In this paper we present uniform proofs utilizing a technique, which can be called roughly speaking, a graph reduction. It consists in observing the Markov chain in time moments being an increasing sequence of stopping times. Some special class of an increasing sequence of stopping times (so-called *strategies*) will be crucial. The fundamental Theorem 2.1 describes the strategies and shows the character of our algorithms.

Recent research into absorbing Markov chains is focused on their applications inter alia in biology (see for instance [6] and the references given there) and in industrial engineering (see [2]). In [9] some properties of fundamental matrix for an absorbing Markov chain are used to solve Possion's equation with Dirichlet boundary condition. In [6] a Lyapunov-type sufficient condition for absorbing Markov chains on a countable state space to almost surely reach the absorbing set is given. Some generalizations of the player ruin problem are set up as a multivariate absorbing Markov chain and solved in [12].

In this paper we identify a Markov chain with its state space, initial state and transition matrix. We will consider a sequence of absorbing Markov chains starting in a one fixed state, with values in a finite state spaces $W_0 \supset W_1 \supset W_2 \ldots$ and with the transition matrices $g_0, g_1, g_2, \ldots$ respectively. Recall that g is a transition matrix if $g : W \times W \mapsto [0,1]$ satisfy $\sum_{a_2 \in W} g(a_1, a_2) = 1$ for all $a_1 \in W$. We set notation:

$$\boldsymbol{x} = (x_0, x_1, x_2, \ldots) \in W_0^{\mathbb{N}}, \quad \boldsymbol{y} = (y_0, y_1, y_2, \ldots) \in W_1^{\mathbb{N}}$$

for the trajectories of first two Markov chains in the mentioned sequence. Denote by $x_n : W_0^{\mathbb{N}} \mapsto W_0$, $y_n : W_1^{\mathbb{N}} \mapsto W_1$ the projections: $x_n(\boldsymbol{x}) := x_n$, $\boldsymbol{x} \in W_0^{\mathbb{N}}$ and $y_n(\boldsymbol{y}) := y_n$, $\boldsymbol{y} \in W_1^{\mathbb{N}}$. All considered Markov chains start in the fixed initial state $e \in W_i$, $i = 0, 1, 2, \ldots$, so we can denote by $\mathfrak{X}_i$ the sets of those $(z_0, z_1, \ldots) \in W_i^{\mathbb{N}}$ for which $z_0 = e$. Let $\boldsymbol{a} = (a_0, a_1, \ldots, a_n) \in W_0^{n+1}$ and $\boldsymbol{b} = (b_0, b_1, \ldots, b_n) \in W_1^{n+1}$ be some paths of length n. If $a_0 = e$ we denote by $C_{\boldsymbol{a}}^0$ the cylinder sets in $\mathfrak{X}_0$, determined by a, i.e.:

$$C_{\boldsymbol{a}}^0 = \{\boldsymbol{x} \in \mathfrak{X}_0 : (x_0, x_1, \ldots, x_n) = (a_0, a_1, \ldots, a_n)\}. \tag{1.1}$$

If $b_0 = e$, in a similar way we define cylinder set $C_{\boldsymbol{b}}^1$ in $\mathfrak{X}_1$. The cylinder sets generate natural filtrations $\{\mathcal{C}_n\}_{n \in \mathbb{N}}$, $\{\mathcal{C}_n^1\}_{n \in \mathbb{N}}$ in $\mathfrak{X}_0$, $\mathfrak{X}_1$ for the observed Markov chains i.e.

$$\mathcal{C}_n = \left\{C_{\boldsymbol{a}}^0 : \boldsymbol{a} \in W_0^{n+1},\, a_0 = e\right\}, \ \mathcal{C}_n^1 = \left\{C_{\boldsymbol{b}}^1 : \boldsymbol{b} \in W_1^{n+1},\, b_0 = e\right\}. \tag{1.2}$$

In $\left(\mathfrak{X}_0, \sigma\left(\{\mathcal{C}_n\}_{n\in\mathbb{N}}\right)\right)$ and $\left(\mathfrak{X}_1, \sigma\left(\{\mathcal{C}^1_n\}_{n\in\mathbb{N}}\right)\right)$ we define probability measures P and P^1 by

$$(1.3)\qquad \begin{aligned} P\left(C^0_{\boldsymbol{a}}\right) &= g_0\left(e, a_1\right) g_0\left(a_1, a_2\right) \dots g_0\left(a_{n-1}, a_n\right), \quad C^0_{\boldsymbol{a}} \in \mathcal{C}_n, \\ P^1\left(C^1_{\boldsymbol{b}}\right) &= g_1\left(e, b_1\right) g_1\left(b_1, b_2\right) \dots g_1\left(b_{n-1}, b_n\right), \quad C^1_{\boldsymbol{b}} \in \mathcal{C}^1_n. \end{aligned}$$

We also use the following short notation for the conditional probabilities

$$(1.4)\qquad P_{\boldsymbol{a}}\left(\cdot\right) = P\left(\cdot | C^0_{\boldsymbol{a}}\right) \text{ and } P^1_{\boldsymbol{b}}\left(\cdot\right) = P^1\left(\cdot | C^1_{\boldsymbol{b}}\right).$$

On the grounds of methods which we focus on and the tradition of Płocki's book [11] we use the terminology.

Definition 1.1. The pair $\left(\left(W_0, g_0\right); e\right)$ is called *stochastic graph (or an absorbing Markov chain)* if

(1) $g_0 : W_0 \times W_0 \mapsto [0,1]$ is stochastic matrix, $e \in W_0$ (thus $\left(\left(W_0, g_0\right); e\right)$ is a Markov chain);
(2) there exists the set $S \subset W_0$ of absorbing states i.e. $s \in S$ if $g_0\left(s,s\right) = 1$;
(3) $\forall_{(a\in W_0\setminus S)}\, \exists_{(n\in\mathbb{N},\, \boldsymbol{a}\in W_0^n)}\ \ g_0\left(a, a_1\right) g_0\left(a_1, a_2\right) \dots g_0\left(a_{n-1}, a_n\right) g_0\left(a_n, s\right) > 0$ for some $s \in S$;
(4) $\forall_{(s\in S)} \exists_{(n\in\mathbb{N},\ \boldsymbol{a}\in W_0^n)}\ g_0\left(e, a_1\right) g_0\left(a_1, a_2\right) \dots g_0\left(a_{n-1}, a_n\right) g_0\left(a_n, s\right) > 0.$

From now on, we consider a fixed stochastic graph $\left(\left(W_0, g_0\right); e\right)$. Let $p\left(s\right)$ denote the probability of absorption in the state $s \in S$ i.e. (cf. (1.3))

$$(1.5)\qquad p\left(s\right) = P\left(\bigcup_{n\in\mathbb{N}} \left\{\boldsymbol{x} \in \mathfrak{X}_0 : x_n\left(\boldsymbol{x}\right) = s\right\}\right).$$

By the definition of stochastic graph it is easy to see that for the probabilities $p\left(s\right)$ we have $\sum_{s\in S} p\left(s\right) = 1$.

Recall that $\tau : \mathfrak{X}_0 \mapsto \mathbb{N}$ is a stopping time for the filtration $\{\mathcal{C}_n\}_{n\in\mathbb{N}}$ (cf. (1.2)) if $\{\boldsymbol{x} : \tau\left(\boldsymbol{x}\right) = m\} \in \mathcal{C}_m$ for all $m \in \mathbb{N}$. Denote by

$$\mathcal{F}_\tau = \{A \subset \mathfrak{X}_0 : \forall_{m\in\mathbb{N}}\ A \cap \{\tau = m\} \in \mathcal{C}_m\}$$

the σ-algebra generated by τ. We define $x_\tau : \mathfrak{X}_0 \mapsto W_0$ by $x_\tau(\boldsymbol{x}) \coloneqq x_{\tau(\boldsymbol{x})}(\boldsymbol{x})$.

The paper is organized as follows. In the next section we introduce the notions of *strategy* and *reduced graph*. The strategy is a special sequence of stopping times $\tau_0 < \tau_1 < \tau_2 \dots$ and the reduced graph is roughly speaking a pair $\left(W_1, g_1\right)$ obtained by observing the initial Markov chain only in the time moments $\tau_0, \tau_1, \tau_2, \dots$. In Section 3 some special reduced graphs are presented and used to compute the probabilities to absorption. The last section introduces methods of computing the distribution of time to absorption. Then we use our method to solve a classical problem in game theory.

2. Strategy. In this section we make specific the above-mentioned concepts of strategy and reduced graph. Next, we formulate and prove Theorem 2.1, which states that a reduced graph obtained from a reduced graph is still a reduced graph. Theorem 2.1 is not trivial and is a basis for convenient algorithms proposed in Section 3 for efficient calculations (which can also be done numerically) of probability to absorption.

More precisely, for the aims being realized in Section 4 we formulate a theorem about the graph reduction in a stronger form, after definition of a sequence $\tau_0 < \tau_1 < \tau_2 \dots$ being a *strong strategy* and the definition of a *strong reduced graph* (Definitions 2.1, 2.2).

We believe that the role of assertions of the type above-mentioned have been underestimated.

Definition 2.1. An increasing sequence $\tau_0, \tau_1, \tau_2, \dots$ of stopping times for the filtration $\{\mathcal{C}_n\}_{n\in\mathbb{N}}$ is called (W_1, g_1)-*strategy* if (cf. (1.1)–(1.4)):

(1) $x_{\tau_n} : \mathfrak{X}_0 \mapsto W_1$ for $n \geq 0$.

(2) For all $m \geq n \geq 0$ and for $\boldsymbol{a} \in W_0^{m+1}$ satisfying $a_0 = e$, $P\left(C^0_{\boldsymbol{a}}\right) > 0$ and $\tau_n = m$ on whole cylinder $C^0_{\boldsymbol{a}}$ we have, for $a_m = b_1$

$$P_{\boldsymbol{a}}\left(x_{\tau_n} = b_1, x_{\tau_{n+1}} = b_2\right) = g_1\left(b_1, b_2\right) \text{ for all } b_2 \in W_1. \tag{2.1}$$

Lemma 2.1. *Let τ_0, τ_1, τ_2, ... be a (W_1, g_1)-strategy for the filtration $\{\mathcal{C}_n\}_{n\in\mathbb{N}}$. For all $k \geq 1$, $m \geq n \geq 0$ and for $\boldsymbol{a} \in W_0^{m+1}$ satisfying $a_0 = e$, $P\left(C^0_{\boldsymbol{a}}\right) > 0$ and $\tau_n = m$ on $C^0_{\boldsymbol{a}}$ we have, for $a_m = b_1 \in W_1$*

$$\begin{aligned} &P_{\boldsymbol{a}}\left(x_{\tau_n} = b_1, x_{\tau_{n+1}} = b_2, \dots, x_{\tau_{n+k}} = b_{k+1}\right) \\ &\quad = g_1\left(b_1, b_2\right) g_1\left(b_2, b_3\right) \dots g_1\left(b_k, b_{k+1}\right) \end{aligned} \tag{2.2}$$

for all $b_2, \dots, b_{k+1} \in W_1$.

Proof. Suppose that (2.2) is true for some $k > 1$, then for $k+1$ we have

$$\begin{aligned} &P_{\boldsymbol{a}}\left(x_{\tau_n} = b_1, x_{\tau_{n+1}} = b_2, \dots, x_{\tau_{n+k+1}} = b_{k+2}\right) \\ &\quad = P_{\boldsymbol{a}}\left(x_{\tau_{n+k+1}} = b_{k+2} | x_{\tau_n} = b_1, x_{\tau_{n+1}} = b_2, \dots, x_{\tau_{n+k}} = b_{k+1}\right) \\ &\quad\quad \times g_1\left(b_1, b_2\right) \dots g_1\left(b_k, b_{k+1}\right). \end{aligned}$$

It is sufficient to show that

$$\begin{aligned} &P_{\boldsymbol{a}}\left(x_{\tau_{n+k+1}} = b_{k+2} | x_{\tau_n} = b_1, x_{\tau_{n+1}} = b_2, \dots, x_{\tau_{n+k}} = b_{k+1}\right) \\ &\quad = g_1\left(b_{k+1}, b_{k+2}\right). \end{aligned} \tag{2.3}$$

Notice that $\left\{x_{\tau_n} = b_1, x_{\tau_{n+1}} = b_2, \dots, x_{\tau_{n+k}} = b_{k+1}\right\}$ is a sum of some cylinder sets $C^0_{\boldsymbol{d}}$, $d \in W_0^l$ such that

$$\tau_{n+k} = l \text{ on } C_{\boldsymbol{d}} \text{ and } d_l = b_{k+1},$$

then from the definition of (W_1, g_1)-strategy we obtain (2.3). □

Definition 2.2. A pair $((W_1, g_1), e)$ is called a *reduced graph* obtained from $((W_0, g_0), e)$ if there exists a (W_1, g_1)-strategy on $\mathfrak{X}_0$.

In the case described by Definition 2.2 we also say that a reduced graph $((W_1, g_1), e)$ is determined by the (W_1, g_1)-strategy $\tau_0, \tau_1, \tau_2, \ldots$ on $\mathfrak{X}_0$ or shortly that $((W_1, g_1), e)$ is obtained from $((W_0, g_0); e)$ by $\tau_0, \tau_1, \tau_2, \ldots$. From the definition of (W_1, g_1)-strategy $\tau_0, \tau_1, \tau_2, \ldots$ it is easy to see that $((W_1, g_1), e)$ is a stochastic graph on $\left(\mathfrak{X}_1, \sigma\left(\{\mathcal{C}_n^1\}_{n\in\mathbb{N}}\right), P^1\right)$ (cf. (1.3)). Notice that S is again the set of absorbing states in $((W_1, g_1), e)$ since $\tau_n \mapsto \infty$ (on whole $\mathfrak{X}_0$). We also obtain surjective measurable transformation $\boldsymbol{Y}$: $\mathfrak{X}_0 \mapsto \mathfrak{X}_1$ such that

$$\boldsymbol{Y}((x_0, x_1, x_2, \ldots)) := (x_{\tau_0}, x_{\tau_1}, x_{\tau_2},) \tag{2.4}$$

and, by Lemma 2.1, cf. (2.2),

$$P^1 = P \circ \boldsymbol{Y}^{-1}. \tag{2.5}$$

We can identify $\boldsymbol{y} \in \mathfrak{X}_1$ with the set $\{\boldsymbol{x} \in \mathfrak{X}_0 : \boldsymbol{Y}(\boldsymbol{x}) = \boldsymbol{y}\} \subset \mathfrak{X}_0$. Notice that every sequence of stopping times $\eta_0, \eta_1, \eta_2, \ldots$ for the filtration $\{\mathcal{C}_n^1\}_{n\in\mathbb{N}}$ on $\mathfrak{X}_1$ (cf. (1.2)) determines a sequence of stopping times $(\tilde{\eta}_0, \tilde{\eta}_1, \tilde{\eta}_2, \ldots)$ for the filtration $\{\sigma(x_{\tau_0}, x_{\tau_1}, \ldots, x_{\tau_n})\}_{n\in\mathbb{N}}$ on $\left(\mathfrak{X}_0, \sigma\left(\{\mathcal{C}_n\}_{n\in\mathbb{N}}\right), P\right)$ in the following way

$$\tilde{\eta}_i(\boldsymbol{x}) := \eta_i(\boldsymbol{Y}(\boldsymbol{x})), \ i = 0, 1, 2, \ldots \tag{2.6}$$

Denote by $p_1(s) = P^1\left(\bigcup_{n\in\mathbb{N}} \{y_n = s\}\right)$ the probability of absorption in the stochastic graph $((W_1, g_1), e)$.

Lemma 2.2. *For all $s \in S$ the following holds $p(s) = p_1(s)$.*

Proof. Fix $s \in S$. From (2.5) we have $p_1(s) = P\left(\bigcup_{n\in\mathbb{N}} \{x_{\tau_n} = s\}\right)$. Hence it is enough to prove that $\bigcup_{n\in N} \{x_n = s\} \subset \bigcup_{n\in N} \{x_{\tau_n} = s\}$. Indeed, if $\boldsymbol{x} \in \bigcup_{n\in N} \{x_n = s\}$, then there exist $n \geq 1$ and $\boldsymbol{a} \in W_0^n$, $a_0 = e$ such that $\boldsymbol{x} \in C_{\boldsymbol{a}}^0 \cap \{x_n = s\}$. Since $\tau_k \mapsto \infty$ a.s. and for every $k \in N$, $\tau_k < \infty$ a.s., then there exist $m \geq 0$ and $k \geq 1$ such that $\tau_k(\boldsymbol{x}) = n + m$. Therefore $x_{\tau_k}(\boldsymbol{x}) = s$. ☐

As we said at the beginning of this section we need to consider strategies with some additional property, namely times between visiting state $b_2 \in W_1$ from being in the state $b_1 \in W_1$ are conditionally independent:

Definition 2.3. A (W_1, g_1)-strategy $(\tau_0, \tau_1, \tau_2, \ldots)$ for the filtration $\{\mathcal{C}_n\}_{n\in\mathbb{N}}$ is called a (W_1, g_1)-*strong strategy*, if for any $b_1, b_2 \in W_1$ there exists a function $\mathbb{N} \ni k \mapsto g_1^{(k)}(b_1, b_2) \in [0, 1]$ defined by

$$g_1^{(k)}(b_1, b_2) := P_{\boldsymbol{a}}\left(\tau_{n+1} - \tau_n = k, x_{\tau_{n+1}} = b_2, x_{\tau_n} = b_1\right), \ k \geq 1,$$

for any $n \geq m \geq 0$ and for any $\boldsymbol{a} \in W_0^{m+1}$ such that $a_0 = e$, $P\left(C_{\boldsymbol{a}}^0\right) > 0$, $a_m = b_1$ and $\tau_n = m$ on $C_{\boldsymbol{a}}^0$. In this case we call the pair $((W_1, g_1), e)$ *a strong reduced graph.*

We denote by

$$p_1^{(\cdot)}(b_1,b_2) = P_{\boldsymbol{a}}\left(\tau_{n+1}-\tau_n=\cdot|x_{\tau_n}=b_1,x_{\tau_{n+1}}=b_2\right) = \begin{cases} \frac{g_1^{(\cdot)}(b_1,b_2)}{g_1(b_1,b_2)} & \text{if } g_1(b_1,b_2)>0,\\ 0 & \text{if } g_1(b_1,b_2)=0,\end{cases} \tag{2.7}$$

the probability distribution of time of transition in "one step" between the states b_1 and b_2 in the strong reduced graph $((W_1,g_1),e)$.

Let $m_1^n(a,b)=\sum_{i=1}^{\infty} i^n p_1^{(i)}(a,b)$ denote the nth moment of distribution $p_1^{(\cdot)}(a,b)$. With the strong reduced graph $((W_1,g_1),e)$ we connect the matrix $\mathbb{P}_1=\left[p_1^{(\cdot)}(a,b)\right]_{a,b\in W_1}$ of distributions of time of transition in "one step" between the states W_1 and the sequence of matrices of its moments $\{M_1^n\}_{n\in\mathbb{N}}$, where $M_1^n=[m_1^n(a,b)]_{a,b\in W_1}$, $n\in\mathbb{N}$. We shortly call M_1^n the matrix of nth moments for the strong reduced graph $((W_1,g_1),e)$.

For a fixed $\boldsymbol{b}=(b_1,b_2,\ldots,b_m)\in W_1^m$ denote by $p_1^{(\cdot)}\boldsymbol{b}$ the convolution of distributions $p_1^{(\cdot)}(b_1,b_2), p_1^{(\cdot)}(b_2,b_3),\ldots,p_1^{(\cdot)}(b_{m-1},b_m)$ i.e.

$$p_1^{(k)}\boldsymbol{b}=\sum_{i_1=1}^{k-1}\sum_{i_2=1}^{i_1}\cdots\sum_{i_{m-2}=1}^{i_{m-3}} p_1^{(k-i_1)}(b_1,b_2)\,p_1^{(i_1-i_2)}(b_2,b_3)\ldots p_1^{(i_{m-2})}(b_{m-1},b_m)$$

and by $m_1^n\boldsymbol{b}$ the nth moment of distribution $p_1^{(\cdot)}\boldsymbol{b}$.

Lemma 2.3. *For any* $m,l\in\mathbb{N}$ *and for any* $\boldsymbol{b}\in W_1^{m+1}$, $\boldsymbol{a}\in W_0^{l+1}$ *satisfying* $a_0=e$, $a_l=b_0$, $P_{\boldsymbol{a}}\left(\left(x_{\tau_n},x_{\tau_{n+1}},\ldots,x_{\tau_{n+m}}\right)=\boldsymbol{b}\right)>0$ *and* $\tau_n=l$ *on* $C_{\boldsymbol{a}}^0$ *we have*

$$P_{\boldsymbol{a}}\left(\tau_{n+m}-\tau_n=\cdot|\left(x_{\tau_n},x_{\tau_{n+1}},\ldots,x_{\tau_{n+m}}\right)=\boldsymbol{b}\right)=p_1^{(\cdot)}\boldsymbol{b}. \tag{2.8}$$

Proof. The proof is done by induction with respect to m. For $m=1$ (2.8) is satisfied by the definition of strong strategy $\tau_0,\tau_1,\ldots$. Suppose that (2.8) holds for $m>1$. We show that (2.8) is true for $m+1$. Indeed, let $\boldsymbol{b}\in W_1^{m+2}$, $\boldsymbol{a}\in W_0^{l+1}$ satisfy $a_0=e$, $a_l=b_0$, $P_{\boldsymbol{a}}\left(\left(x_{\tau_n},x_{\tau_{n+1}},\ldots,x_{\tau_{n+m+1}}\right)=\boldsymbol{b}\right)>0$ and $\tau_n=l$ on $C_{\boldsymbol{a}}^0$, then

$$\begin{aligned} &P_{\boldsymbol{a}}\left(\tau_{n+m+1}-\tau_n=k|\left(x_{\tau_n},\ldots,x_{\tau_{n+m+1}}\right)=\boldsymbol{b}\right)\\ &=\sum_i\Big[P_{\boldsymbol{a}}\left(\tau_{n+m+1}-\tau_n=k|\left(x_{\tau_n},\ldots,x_{\tau_{n+m+1}}\right)=\boldsymbol{b},\tau_{n+1}-\tau_n=i\right)\\ &\qquad\times P_{\boldsymbol{a}}\left(\tau_{n+1}-\tau_n=i|\left(x_{\tau_n},\ldots,x_{\tau_{n+m+1}}\right)=\boldsymbol{b}\right)\Big]. \end{aligned} \tag{2.9}$$

Notice that from the definition of $((W_1,g_1),e)$-strong strategy $\tau_0,\tau_1,\ldots$ we have

$$\begin{aligned} &P_{\boldsymbol{a}}\left(\tau_{n+1}-\tau_n=i|\left(x_{\tau_n},\ldots,x_{\tau_{n+m+1}}\right)=\boldsymbol{b}\right)\\ &\quad=P_{\boldsymbol{a}}\left(\tau_{n+1}-\tau_n=i|x_{\tau_n}=b_0,x_{\tau_{n+1}}=b_1\right)=p_1^{(i)}(b_0,b_1) \end{aligned} \tag{2.10}$$

(cf. (2.7)) and from the inductive assumption we get

$$\tag{2.11} \begin{aligned} &P_{\boldsymbol{a}}\left(\tau_{n+m+1}-\tau_n=k|x_{\tau_n}=b_0,\ldots,x_{\tau_{n+m}}=b_m,\tau_{n+1}-\tau_n=i\right)\\ &=P_{\boldsymbol{a}}\big(\tau_{n+m+1}-\tau_{n+1}=k-i|\\ &\qquad\qquad \tau_{n+1}=i+l,x_{\tau_{n+1}}=b_1,\ldots,x_{\tau_{n+m}}=b_m\big)\\ &=p_1^{(k-i)}\left(b_1,\ldots,b_{m+1}\right). \end{aligned}$$

By combining (2.10) and (2.11) with (2.9) we obtain the desired formula

$$P_{\boldsymbol{a}}\left(\tau_{n+m+1}-\tau_n=k|\left(x_{\tau_n},x_{\tau_{n+1}},\ldots,x_{\tau_{n+m+1}}\right)=\boldsymbol{b}\right)=p_1^{(k)}\boldsymbol{b}. \qquad \square$$

Now we formulate the main result on reduced graph and strong reduced graph.

Theorem 2.1. *If $((W_1,g_1),e)$ is a reduced (strong reduced) graph obtained from $((W_0,g_0),e)$ and $((W_2,g_2),e)$ is a reduced (strong reduced) graph obtained from $((W_1,g_1),e)$, then $((W_2,g_2),e)$ is the reduced (respectively strong reduced) graph obtained from $((W_0,g_0),e)$.*

In the proof of Theorem 2.1 there is among other things the following subtlety: a (W_1,g_1)-strategy $\tau_0,\tau_1,\tau_2,\ldots$ is defined on $\big(\mathfrak{X}_0,\sigma\big(\{\mathcal{C}_n\}_{n\in\mathbb{N}}\big),P\big)$ while a (W_2,g_2)-strategy $\eta_0,\eta_1,\eta_2,\ldots$ is connected with different Markov chain $((W_1,g_1),e)$ and defined on $\big(\mathfrak{X}_1,\sigma\big(\{\mathcal{C}_n^1\}_{n\in\mathbb{N}}\big),P^1\big)$ being the canonical space for this Markov chain. Obviously, since $((W_1,g_1),e)$ is a reduced graph obtained from $((W_0,g_0),e)$, then $\eta_0,\eta_1,\eta_2,\ldots$ generate a sequence of random variables on $\big(\mathfrak{X}_0,\sigma\big(\{\mathcal{C}_n\}_{n\in\mathbb{N}}\big),P\big)$ given by (2.6).

Proof. Since the theorem covers two cases, we divide the proof into two parts. First we prove the version for a reduced graph, after that we show the assertion of the theorem for a strong reduced graph.

Let $\tau_0,\tau_1,\tau_2,\ldots$ be a (W_1,g_1)-strategy for the filtration $\{\mathcal{C}_n\}_{n\in\mathbb{N}}$ and η_0, η_1, $\eta_2,\ldots$ be a (W_2,g_2)-strategy for the filtration $\big\{\mathcal{C}_n^1\big\}_{n\in\mathbb{N}}$. From (2.6) we get a sequence of random variables $\tilde\eta_0,\tilde\eta_1,\tilde\eta_2,\ldots$. We show that $(\tau_{\tilde\eta_0},\tau_{\tilde\eta_1},\tau_{\tilde\eta_2},\ldots)$ is (W_2,g_2) -strategy for the filtration $\{\mathcal{C}_n\}_{n\in\mathbb{N}}$.

Obviously $\tau_{\tilde\eta_0}<\tau_{\tilde\eta_1}<\tau_{\tilde\eta_2},\ldots$ and for any $n,m\in\mathbb{N}$ we have

$$\{\tau_{\tilde\eta_n}=m\}=\bigcup_{k\in\mathbb{N}}\{\tau_k=m,\tilde\eta_n=k\}\in\mathcal{C}_m$$

by $\{\tilde\eta_n=k\}\in\sigma\left(x_{\tau_0},x_{\tau_1},\ldots,x_{\tau_k}\right)\subset\mathcal{F}_{\tau_k}$. By the definition of $(\eta_0,\eta_1,\eta_2,\ldots)$ we see that for any $n\in\mathbb{N}$ measurable function $x_{\tau_{\tilde\eta_n}}$ on $\mathfrak{X}_0$ takes value in W_2.

Let $l,n\in\mathbb{N}$ and $\boldsymbol{a}\in W_0^{l+1}$ be such that $a_0=e$, $P\left(C_{\boldsymbol{a}}^0\right)>0$, $a_l=a$ and $\tau_{\tilde\eta_n}=l$ on $C_{\boldsymbol{a}}^0$. Then there exist $m\le l$ and $\boldsymbol{b}\in W_1^{m+1}$, $b_0=e$, $b_m=a$ such that

$$\tau_m\left(\boldsymbol{x}\right)=l,\ \tilde\eta_n\left(\boldsymbol{x}\right)=m\ \text{ and }\ \left(x_{\tau_0},x_{\tau_1},\ldots,x_{\tau_m}\right)\left(\boldsymbol{x}\right)=\boldsymbol{b}$$

for all $\boldsymbol{x} \in C^0_{\boldsymbol{a}}$ and

$$\{(x_{\tau_0}, x_{\tau_1}, \dots, x_{\tau_m}) = \boldsymbol{b}\} \subset \{\tilde{\eta}_n = m\}.$$

From the last inclusion and the fact that $\boldsymbol{Y}$ is surjective we get

$$C^1_{\boldsymbol{b}} = \boldsymbol{Y}\left(\boldsymbol{Y}^{-1}\left(C^1_{\boldsymbol{b}}\right)\right) = \boldsymbol{Y}\left(\{(x_{\tau_0}, x_{\tau_1}, \dots, x_{\tau_m}) = \boldsymbol{b}\}\right) \subset \{\eta_n = m\}.$$

We show that for any $c \in W_2$ we have

$$P\left(x_{\tau_{\tilde{\eta}_n}} = a, x_{\tau_{\tilde{\eta}_{n+1}}} = c | C^0_{\boldsymbol{a}}\right) = g_2(a, c). \tag{2.12}$$

On $C^1_{\boldsymbol{b}} \subset \mathfrak{X}_1$ we have a probability measure $P^1_{\boldsymbol{b}}$ for which $P_{\boldsymbol{a}}$ generates $P^1_{\boldsymbol{b}}$ i.e.

$$P^1_{\boldsymbol{b}} = P_{\boldsymbol{a}} \circ \boldsymbol{Y}^{-1}. \tag{2.13}$$

Indeed, notice that from the definition of $\boldsymbol{Y}$ we have

$$\begin{aligned}\boldsymbol{Y}^{-1}&\left(C^1_{(e,b_1,\dots,b_{m+k})}\right)\\ &= \left\{\boldsymbol{x} \in \mathfrak{X}_0 : \left(x_{\tau_0}, x_{\tau_1}, \dots, x_{\tau_{m+k}}\right)(\boldsymbol{x}) = (e, b_1, \dots, b_{m+k})\right\},\end{aligned}$$

for any $b_{m+1}, \dots, b_{m+k} \in W_1$. Hence from Lemma 2.1 we get

$$P_{\boldsymbol{a}}\left(\boldsymbol{Y}^{-1}\left(C^1_{\boldsymbol{b}}\right)\right) = g_1(b_m, b_{m+1}) \dots g_1(b_{m+k-1}, b_{m+k}).$$

Thus (2.13) have been proved.

From the definition of $(\eta_0, \eta_1, \eta_2, \dots)$ and (2.13) for $m > n$ and $C^1_{\boldsymbol{b}}$ we can write

$$g_2(a, c) = P^1_{\boldsymbol{b}}\left(y_{\eta_n} = a, y_{\eta_{n+1}} = c\right) = P_{\boldsymbol{a}} \circ \boldsymbol{Y}^{-1}\left(y_{\eta_n} = a, y_{\eta_{n+1}} = c\right), \tag{2.14}$$

for any $c \in W_2$. To finish the proof notice that

$$\begin{aligned}\boldsymbol{Y}^{-1}\left(y_{\eta_n} = a, y_{\eta_{n+1}} = c\right) &= \left\{\boldsymbol{x} \in \mathfrak{X}_0 : \boldsymbol{Y}(\boldsymbol{x}) \in \left\{y_{\eta_n} = a, y_{\eta_{n+1}} = c\right\}\right\}\\ &= \left\{\boldsymbol{x} \in \mathfrak{X}_0 :\ x_{\tau_{\tilde{\eta}_n}} = a, x_{\tau_{\tilde{\eta}_{n+1}}} = c\right\}.\end{aligned}$$

Then from (2.14) we finally obtain (2.12), which implies that $(\tau_{\tilde{\eta}_0}, \tau_{\tilde{\eta}_1}, \tau_{\tilde{\eta}_2}, \dots)$ is a (W_2, g_2)-strategy for the filtration $\{\mathcal{C}_n\}_{n\in\mathbb{N}}$.

Now assume additionally that $(\tau_0, \tau_1, \tau_2, \dots)$ is a (W_1, g_1)-strong strategy for the filtration $\{\mathcal{C}_n\}_{n\in\mathbb{N}}$ and $(\eta_0, \eta_1, \eta_2, \dots)$ is a (W_2, g_2)-strong strategy for the filtration $\left\{\mathcal{C}^1_n\right\}_{n\in\mathbb{N}}$. We show that $(\tau_{\tilde{\eta}_0}, \tau_{\tilde{\eta}_1}, \tau_{\tilde{\eta}_2}, \dots)$ is a (W_2, g_2)-strong strategy for the filtration $\{\mathcal{C}_n\}_{n\in\mathbb{N}}$.

Indeed, for any $k \in \mathbb{N}$

$$
\begin{aligned}
&P_{\boldsymbol{a}}\left(\tau_{\tilde{\eta}_{n+1}} - \tau_{\tilde{\eta}_n} = k, x_{\tau_{\tilde{\eta}_{n+1}}} = c, x_{\tau_{\tilde{\eta}_n}} = a\right)\\
&\quad = \sum_{1\le i\le k} P_{\boldsymbol{a}}\left(\tau_{\tilde{\eta}_{n+1}} - \tau_{\tilde{\eta}_n} = k, x_{\tau_{\tilde{\eta}_{n+1}}} = c, x_{\tau_{\tilde{\eta}_n}} = a, \tilde{\eta}_{n+1} - \tilde{\eta}_n = i\right)\\
&\quad = \sum_{1\le i\le k} \left[P_{\boldsymbol{a}}\left(\tau_{m+i} - \tau_m = k | x_{\tau_m} = a, x_{\tau_{m+i}} = c, \tilde{\eta}_{n+1} - \tilde{\eta}_n = i\right)\right.\\
&\qquad\qquad \left.\times P_{\boldsymbol{a}}\left(\tilde{\eta}_{n+1} - \tilde{\eta}_n = i, x_{\tau_m} = a, x_{\tau_{m+i}} = c\right)\right].
\end{aligned}
\tag{2.15}
$$

Since $C^0_{\boldsymbol{a}} \subset \{\tau_m = l, \tilde{\eta}_n = m\}$ and by (2.13) we have

$$
\begin{aligned}
&P_{\boldsymbol{a}}\left(x_{\tau_m} = a, x_{\tau_{m+i}} = c, \tilde{\eta}_{n+1} - \tilde{\eta}_n = i\right)\\
&\quad = P^1_{\boldsymbol{b}}\left(y_{\eta_{n+1}} = a, y_{\eta_n} = c, \eta_{n+1} - \eta_n = i\right) = g_2^{(i)}(a, c).
\end{aligned}
$$

Then the equality

$$
\begin{aligned}
&\left\{x_{\tau_m} = a, x_{\tau_{m+i}} = c, \tilde{\eta}_{n+1} - \tilde{\eta}_n = i\right\} \cap C^0_{\boldsymbol{a}}\\
&\quad = \bigcup_{\boldsymbol{c}\in I(i,a,c)} C^0_{\boldsymbol{a}} \cap \left\{\left(x_{\tau_m}, x_{\tau_{m+1}}, \ldots, x_{\tau_{m+i-1}}, x_{\tau_{m+i}}\right) = \boldsymbol{c}\right\}
\end{aligned}
$$

holds for some set $I(i, a, c) \subset W_1^{i+1}$ depending on some i, a and c, such that $c_i = c$, $c_0 = a$. Hence, and from (2.15) we get

$$
\begin{aligned}
&P_{\boldsymbol{a}}\left(\tau_{\tilde{\eta}_{n+1}} - \tau_{\tilde{\eta}_n} = k, x_{\tau_{\tilde{\eta}_{n+1}}} = c, x_{\tau_{\tilde{\eta}_n}} = a\right)\\
&= \sum_{1\le i\le k} P_{\boldsymbol{a}}\left(\tau_{m+i} - \tau_m = k | \bigcup_{\boldsymbol{c}\in I(i,a,c)} \left\{\left(x_{\tau_m}, \ldots, x_{\tau_{m+i}}\right) = \boldsymbol{c}\right\}\right) g_2^{(i)}(a, c)\\
&= \sum_{1\le i\le k} \frac{\sum_{\boldsymbol{c}\in I(i,a,c)} P_{\boldsymbol{a}}\left(\tau_{m+i} - \tau_m = k | \left(x_{\tau_m}, \ldots, x_{\tau_{m+i}}\right) = \boldsymbol{c}\right) g_1(a, c_1)\ldots g_1(c_{i-1}, c)}{\sum_{\boldsymbol{c}\in I(i,a,c)} g_1(a, c_1)\ldots g_1(c_{i-1}, c)}\\
&\qquad \times g_2^{(i)}(a, c).
\end{aligned}
$$

Using Lemma 2.3 we get

$$
\begin{aligned}
&P_{\boldsymbol{a}}\left(\tau_{\tilde{\eta}_{n+1}} - \tau_{\tilde{\eta}_n} = k, x_{\tau_{\tilde{\eta}_{n+1}}} = c, x_{\tau_{\tilde{\eta}_n}} = a\right)\\
&\quad = \sum_{1\le i\le k} \frac{\sum_{\boldsymbol{c}\in I(i,a,c)} p_1^{(k)}\boldsymbol{c}\; g_1(a, c_1)\ldots g_1(c_{i-1}, c)}{\sum_{\boldsymbol{c}\in I(i,a,c)} g_1(a, c_1)\ldots g_1(c_{i-1}, c)} g_2^{(i)}(a, c)
\end{aligned}
\tag{2.16}
$$

and we can define function $\mathbb{N} \ni k \mapsto \tilde{g}_2^{(k)}(a, c) \in [0, 1]$ by

$$
\tilde{g}_2^{(k)}(a, c) := P_{\boldsymbol{a}}\left(\tau_{\tilde{\eta}_{n+1}} - \tau_{\tilde{\eta}_n} = k, x_{\tau_{\tilde{\eta}_{n+1}}} = c, x_{\tau_{\tilde{\eta}_n}} = a\right). \tag{2.17}
$$

□

Now, we see that every reduced graph has less states than its initial one. At the same time each reduced graph has the initial state e and the set S of absorbing states. Moreover, we have shown in Lemma 2.2 that for each absorbing state $s \in S$ the probabilities of absorption in state s are the same for the initial stochastic graph and for the reduced graph. With each reduced graph we combined the matrix of distributions of time of transition in "one step" between its states (cf. (2.7)) and the sequence of matrices of moments of those distributions. Now we describe algorithms which allow us to obtain the final reduced graph, i.e. the reduced graph with the state space consisting only of the initial state e and the absorbing states S. Then from its transition matrix, matrix of distributions of time of transition in one step between its states and the moment matrices we obtain the probabilities to absorption, distributions of time to absorption and moments of time to absorption respectively.

3. Graph reduction. In this section we describe the algorithms of creating some special reduced graphs. The first algorithm of graph reduction we call a loop reduction.

Definition 3.1. A pair $((W_1, g_1), e)$ is formed from the stochastic graph $((W_0, g_0), e)$ by *the loop (a, a) reduction* if:

(1) $W_1 = W_0$ and $g_1(b, c) = g_0(b, c)$ for $W_1 \ni b \neq a$,

(2) $g_1(a, b) = \dfrac{g_0(a, b)}{1 - g_0(a, a)}$ for $W_1 \ni b \neq a$,

(3) $g_1(a, a) = 0$.

Theorem 3.1. *If $((W_1, g_1), e)$ is formed from a stochastic graph $((W_0, g_0), e)$ by a loop (a, a) reduction, then $((W_1, g_1), e)$ is a strong reduced graph obtained from $((W_0, g_0), e)$.*

Proof. It is sufficient to find a strong strategy which determines the reduced graph $((W_1, g_1), e)$. If $a \neq a_0$, we show that a sequence $(\gamma_0, \gamma_1, \gamma_2, \ldots)$ defined below is a (W_1, g_1)-strong strategy. Let

$$\gamma_0 = 0,$$

$$\gamma_1 = 1,$$

$$\gamma_2(\boldsymbol{x}) = \begin{cases} \min\{i > \gamma_1 : x_i(\boldsymbol{x}) \neq a\}, & \boldsymbol{x} \in \{x_{\gamma_1} = a\}, \\ \gamma_1(\boldsymbol{x}) + 1, & \boldsymbol{x} \in \{x_{\gamma_1} \neq a\}, \end{cases}$$

$$\vdots$$

$$\gamma_n(\boldsymbol{x}) = \begin{cases} \min\{i > \gamma_{n-1} : x_i(\boldsymbol{x}) \neq a\}, & \boldsymbol{x} \in \left\{x_{\gamma_{n-1}} = a\right\}, \\ \gamma_{n-1}(\boldsymbol{x}) + 1, & \boldsymbol{x} \in \left\{x_{\gamma_{n-1}} \neq a\right\}, \quad \text{etc.} \end{cases}$$

Clearly, $(\gamma_0, \gamma_1, \gamma_2, \ldots)$ is a non-decreasing family and $\gamma_n \mapsto \infty$ a.s. Moreover, γ_0, γ_1 are stopping times. To show by induction that $(\gamma_0, \gamma_1, \gamma_2, \ldots)$

are stopping times, fix $n \in N$ and suppose that $\gamma_0, \gamma_1, \ldots, \gamma_{n-1}$ are stopping times. Then for $m > n$ we have

$$\begin{aligned}\{\gamma_n = m\} &= \left\{\gamma_n = m, x_{\gamma_{n-1}} = a\right\} \cup \left\{\gamma_n = m, x_{\gamma_{n-1}} \neq a\right\} \\ &= \bigcup_{i=1}^{m-n+1} \{x_{m-i} = a, \ldots, x_{m-1} = a, x_m \neq a, \gamma_{n-1} = m-i\} \\ &\qquad \cup \left\{\gamma_{n-1} = m-1, x_{m-1} \neq a\right\} \in \mathcal{F}_m,\end{aligned}$$

hence γ_n is a stopping time.

Fix $b \in W_1$. Let $\boldsymbol{a} \in W_0^{m+1}$ satisfy $a_0 = e$, $P\left(C^0_{\boldsymbol{a}}\right) > 0$, $a_m = a$ and $\gamma_n = m$ on $C^0_{\boldsymbol{a}}$, then we have

$$\begin{aligned}P_{\boldsymbol{a}}\left(x_{\gamma_{n+1}} = b\right) &= \sum_{i=1}^{\infty} P_{\boldsymbol{a}}\left(x_{m+i} = b | \gamma_{n+1} = m+i\right) P_{\boldsymbol{a}}\left(\gamma_{n+1} = m+i\right) \\ &= \sum_{i=1}^{\infty} \frac{g_0(a,b)}{1 - g_0(a,a)} g_0(a,a)^{i-1} \left(1 - g_0(a,a)\right) = g_1(a,b)\end{aligned}$$

and for any $k \geq 1$

$$\begin{aligned}P_{\boldsymbol{a}}&\left(\gamma_{n+1} - \gamma_n = k, x_{\gamma_{n+1}} = b, x_{\gamma_n} = a\right) = P_{\boldsymbol{a}}\left(\gamma_{n+1} = m+k, x_{m+k} = b\right) \\ &= P_{\boldsymbol{a}}\left(x_{m+1} = a, \ldots, x_{m+k-1} = a, x_{m+k} \neq a, x_{m+k} = b\right) \\ &= g_0(a,a)^{k-1} g_0(a,b).\end{aligned}$$

Hence we can define

$$\mathbb{N} \in k \mapsto g_1^{(k)}(a,b) := g_0(a,a)^{k-1} g_0(a,b). \tag{3.1}$$

Suppose now, that $\boldsymbol{a} \in W_0^{m+1}$ satisfies $a_0 = e$, $P\left(C^0_{\boldsymbol{a}}\right) > 0$, $a_m = b \neq a$ and $\gamma_n(\omega) = m$ on $C^0_{\boldsymbol{a}}$. Then we have, for any $c \in W_0$

$$P_{\boldsymbol{a}}\left(x_{\gamma_{n+1}} = c\right) = P_{\boldsymbol{a}}\left(x_{m+1} = c\right) = g_0(b,c) = g_1(b,c)$$

and

$$\begin{aligned}g_1^{(k)}(b,c) &= P_{\boldsymbol{a}}\left(\gamma_{n+1} - \gamma_n = k, x_{\gamma_{n+1}} = c, x_{\gamma_n} = a\right) \\ &= P_{\boldsymbol{a}}\left(\gamma_{n+1} = m+k, x_{m+k} = b\right) \\ &= \begin{cases} g_0(b,c), & k = 1 \\ 0, & k = 2, 3, \ldots, \end{cases}\end{aligned} \tag{3.2}$$

by the definition of γ_{n+1}. Therefore $(\gamma_0, \gamma_1, \gamma_2, \ldots)$ is a (W_1, g_1)-strong strategy. If $a = a_0$, we define the family $(\gamma_0, \gamma_1, \gamma_2, \ldots)$ as follows

$$\begin{aligned}\gamma_0 &= 0, \\ \gamma_1(\boldsymbol{x}) &= \min\{i > 0 : x_i(\boldsymbol{x}) \neq a\}, \\ \gamma_2(\boldsymbol{x}) &= \gamma_1(\boldsymbol{x}) + 1,\end{aligned}$$

$$\gamma_3(\boldsymbol{x}) = \begin{cases} \min\{i > \gamma_2 : x_i(\boldsymbol{x}) \neq a\}, & \boldsymbol{x} \in \{x_{\gamma_2} = a\}, \\ \gamma_2(\boldsymbol{x}) + 1, & \boldsymbol{x} \in \{x_{\gamma_2} \neq a\}, \end{cases}$$

$$\vdots$$

$$\gamma_n(\omega) = \begin{cases} \min\{i > \gamma_{n-1} : x_i(\boldsymbol{x}) \neq a\}, & \boldsymbol{x} \in \left\{x_{\gamma_{n-1}} = a\right\}, \\ \gamma_{n-1}(\boldsymbol{x}) + 1, & \boldsymbol{x} \in \left\{x_{\gamma_{n-1}} \neq a\right\}, \ldots. \end{cases}$$

The proof that the family $(\gamma_0, \gamma_1, \gamma_2, \ldots)$ is a (W_1, g_1)- strong strategy is similar to the case of $a \neq a_0$. □

The next algorithm is called the edge reduction.

Definition 3.2. Assume that $g_0(b,b) = 0$. A pair $((W_1, g_1), e)$ is obtained from a stochastic graph $((W_0, g_0), e)$ by *the edge (a,b) reduction* if:

(1) $W_1 = W_0$ and $g_1(c,d) = g_0(c,d)$ for $W_1 \ni c \neq a$,
(2) $g_1(a,c) = g_0(a,b)\, g_0(b,c) + g_0(a,c)$ for $W_1 \ni c \neq b$,
(3) $g_1(a,b) = 0$.

Theorem 3.2. *If $((W_1, g_1), e)$ is formed from a stochastic graph $((W_0, g_0), e)$ by the edge (a,c) reduction, then $((W_1, g_1), e)$ is a strong reduced graph obtained from $((W_0, g_0), e)$.*

Proof. We find a strategy $(\gamma_0, \gamma_1, \gamma_2, \ldots)$ which determines the strategic subgraph $((W_1, g_1), e)$. Let us define

$\gamma_0 = 0,$
$\gamma_1 = 1,$
$\gamma_2(\boldsymbol{x}) = \gamma_1(\boldsymbol{x}) + 1 + \mathbf{1}_{A_1}(\boldsymbol{x}),$
where $A_1 = \{\boldsymbol{x} : x_{\gamma_1}(\boldsymbol{x}) = a, x_{\gamma_1+1}(\boldsymbol{x}) = b\}$,
$\ldots$
$\gamma_n(\boldsymbol{x}) = \gamma_{n-1}(\boldsymbol{x}) + 1 + \mathbf{1}_{A_{n-1}}(\boldsymbol{x}),$
where $A_{n-1} = \left\{\boldsymbol{x} : x_{\gamma_{n-1}}(\boldsymbol{x}) = a, x_{\gamma_{n-1}+1}(\boldsymbol{x}) = b\right\}$, ….

It is clear that $(\gamma_0, \gamma_1, \gamma_2, \ldots)$ is non-decreasing and $\gamma_n \mapsto \infty$ a.s., the variables γ_0, γ_1 are stopping times. One can show by induction that $\gamma_2, \gamma_3, \ldots$ are also stopping times. It easy to check that $(\gamma_0, \gamma_1, \gamma_2, \ldots)$ is a (W_1, g_1)-strong strategy which determines $((W_1, g_1), e)$ and

$$\begin{aligned} g_1^{(k)}(a,b) &:= P_{\boldsymbol{a}}\left(\gamma_{n+1} - \gamma_n = k, x_{\gamma_{n+1}} = b, x_{\gamma_n} = a\right) \\ &= \begin{cases} g_0(a,c), & i = 1, \\ g_0(a,b)\, g_0(b,c), & i = 2, \\ 0, & i = 3, 4 \ldots. \end{cases} \end{aligned} \tag{3.3}$$

□

The last algorithm is called the state reduction. We reduce a state $b \in W_0$ in the case when there exists only one edge (a,b) directed to b and $g_0(b,b) = 0$.

Definition 3.3. Assume that there exists only one edge (a, b) directed to b and $g_0(b, b) = 0$. A pair $((W_1, g_1), e)$ is formed from a stochastic graph $((W_0, g_0), e)$ by *the state $b \in W_0$ reduction* if:

(1) $W_1 = W_0 \setminus \{b\}$,
(2) $g_1(c, d) = g_0(c, b)\, g_0(b, d) + g_0(c, d)$, $c, d \in W_1$.

Theorem 3.3. *If $((W_1, g_1), e)$ is formed from a stochastic graph $((W_0, g_0), e)$ by the state $b \in W_0$ reduction, then $((W_1, g_1), e)$ is a strong reduced graph obtained from $((W_0, g_0), e)$.*

Proof. The (W_1, g_1)-strong strategy which determines the strong reduced graph $((W_1, g_1), e)$ is the same as in the case of the edge reduction. □

Notice that each stochastic graph $((W_0, g_0), e)$ can be reduced according to the above algorithms to the strong reduced graph $((W_r, g_r), e)$ consisting of an initial state e and all the absorbing states S, $W_r = S \cup \{e_0\}$. By Lemma 2.2 the probabilities of absorption in a reduced graph are equal to the probabilities of absorption in its initial graph, thus we have

$$p(s) = g_r(a_0, s), \quad s \in S. \tag{3.4}$$

4. Distribution of time to absorption. Let $\mathfrak{X}_0' = \{\boldsymbol{x} : \exists_{n>0}\, x_n(\boldsymbol{x}) = s\}$ with probability being conditional probability derived from P (cf. (1.3)) be a probability space. Recall that, a random variable $T_s : \mathfrak{X}_0' \mapsto \mathbb{N}$ is called time to absorption in state $s \in S$ if $T_s(\boldsymbol{x}) = \inf\{n \geq 0 : x_n(\boldsymbol{x}) = s\}$, $\boldsymbol{x} \in \mathfrak{X}_0'$.

Let $((W_1, g_1), e)$ be a strong reduced graph obtained from $((W_0, g_0), e)$ by a (W_1, g_1)-strong strategy $\tau_0, \tau_1, \tau_2, \ldots$ on $\mathfrak{X}_0$ and $((W_2, g_2), e)$ be a strong reduced graph obtained from $((W_1, g_1), e)$ by (W_2, g_2)-strong strategy η_0, η_1, $\eta_2, \ldots$ on $\mathfrak{X}_1$ – one of the described in Section 3 (see Definitions 3.1, 3.2, 3.3). For any $n \geq m \geq 0$ and for any $\boldsymbol{b} \in W_1^{m+1}$ such that $b_0 = e$, $P^1\left(C^1_{\boldsymbol{b}}\right) > 0$, $b_m = c_1 \in W_2$ and $\eta_n = m$ on $C^1_{\boldsymbol{b}}$ denote by

$$\begin{aligned} p_2^{(k)}(c_1, c_2) &= P^1_{\boldsymbol{b}}\left(\eta_{n+1} - \eta_n = k | y_{\eta_n} = c_1, y_{\eta_{n+1}} = c_2\right) \\ &= \begin{cases} \frac{g_2^{(\cdot)}(c_1,c_2)}{g_2(c_1,c_2)} & \text{if } g_2(c_1, c_2) > 0, \\ 0 & \text{if } g_2(c_1, c_2) = 0 \end{cases} \end{aligned}$$

for $k \geq 1$, *the probability distribution of time of transition in "one step" between the states c_1 and c_2* in the strong reduced graph $((W_2, g_2), e)$ obtained from (W_1, g_1), for any $c_2 \in W_1$.

By Theorem 2.1 we know that $((W_2, g_2), e)$ is also a strong reduced graph obtained from $((W, g_0), e)$ by (W_2, g_2)- strategy $\tau_{\tilde{\eta}_0}, \tau_{\tilde{\eta}_1}, \tau_{\tilde{\eta}_2}, \ldots$ on $\mathfrak{X}_0$. Hence for any $m \geq n \in \mathbb{N}$ and for $\boldsymbol{a} \in W_0^{m+1}$ such that $a_0 = e$, $P\left(C^0_{\boldsymbol{a}}\right) > 0$,

$a_m = c_1 \in W_2$ and $\tau_{\tilde{\eta}_n} = m$ on $C^0_{\boldsymbol{a}}$ we can define $\mathbb{N} \ni k \mapsto \tilde{g}^{(\cdot)}_2(c_1, c_2)$ by

$$\tilde{g}^{(\cdot)}_2(c_1, c_2) = P_{\boldsymbol{a}}\left(\tau_{\tilde{\eta}_{n+1}} - \tau_{\tilde{\eta}_n} = k, x_{\tau_{\tilde{\eta}_n}} = c_1, x_{\tau_{\tilde{\eta}_{n+1}}} = c_2\right). \tag{4.1}$$

Now, we can also denote by

$$\tilde{p}^{(k)}_2(c_1, c_2) = \begin{cases} \frac{\tilde{g}^{(\cdot)}_2(c_1,c_2)}{g_2(c_1,c_2)} & \text{if } g_2(c_1, c_2) > 0, \\ 0 & \text{if } g_2(c_1, c_2) = 0 \end{cases}$$

for $k \geq 1$, *the probability distribution of time of transition in "one step" between the states* c_1 *and* c_2 in the strong reduced graph $((W_2, g_2), e)$ obtained from the stochastic graph $((W_0, g_0), e)$. Denote by $\tilde{m}^n_2(c_1, c_2)$ nth moment of the distribution $\tilde{p}^{(\cdot)}_2(c_1, c_2)$.

Next we present formulas which allow us to compute $\tilde{p}^{(\cdot)}_2(c_1, c_2)$ for any $c_1, c_2 \in W_2$.

4.1. Loop reduction. Let $((W_2, g_2), e)$ be formed from $((W_1, g_1), e)$ by loop (a, a) reduction, $a \in W_1$. From (2.16)–(2.17) and (3.1)–(3.2) we obtain that

$$\tilde{p}^{(k)}_2(a, b) = \frac{1}{g_2(a, b)} \sum_{1 \leq i \leq k} p^{(k)}_1(\underbrace{a, a \ldots a}_{i-1 \text{ times}}, b) g_1(a, a)^{i-1} g_1(a, b), \tag{4.2}$$

for $k = 1, 2, \ldots$. Hence, we get

$$\begin{aligned} \tilde{m}^n_2(a, b) &= \frac{g_1(a, b)}{g_2(a, b)} \sum_{p=0}^{n} \binom{n}{p} m^{n-p}_1(a, a)\, m^p_1(a, b) \sum_{j=0}^{\infty} j^{n-p} g_1(a, a)^j, \\ \tilde{m}^n_2(a, a) &= 0, \quad \tilde{m}^n_2(b, c) = \tilde{m}^n_1(b, c), \end{aligned} \tag{4.3}$$

where $b \in W_2$, $b \in W_2 \setminus \{a\}$, $n \in \mathbb{N}$.

4.2. Edge reduction. Let $((W_2, g_2), e)$ be formed from $((W_1, g_1), e)$ by edge (a, b) reduction $a, b \in W_1$. From (2.16)–(2.17) and (3.3) we obtain:

$$\tilde{p}^{(k)}_2(a, c) = \frac{1}{g_2(a, c)} \left[p^{(k)}_1(a, c)\, g_1(a, c) + p^{(k)}_1(a, b, c)\, g_1(a, b) g_1(b, c)\right], \tag{4.4}$$

for $k = 1, 2, \ldots$. Hence, we get

$$\begin{aligned} \tilde{m}^n_2(a, c) &= \frac{g_1(a, c)}{g_2(a, c)} m^n_1(a, c) \\ &\quad + \frac{g_1(a, b)\, g_1(b, c)}{g_2(a, c)} \sum_{p=0}^{n} \binom{n}{p} m^{n-p}_1(a, b)\, m^p_p(b, c), \\ \tilde{m}^n_2(a, b) &= 0, \quad \tilde{m}^n_2(c, d) = \tilde{m}^n_1(c, d), \end{aligned} \tag{4.5}$$

where $c \in W_2 \setminus \{a\}$, $d \in W_2$, $n \in \mathbb{N}$.

4.3. State reduction. Let $((W_2, g_2), e)$ be formed from $((W_1, g_1), e)$ by state $b \in W_1$ reduction. Formulas for $\tilde{p}_2^{(\cdot)}(\cdot,\cdot)$ and $\tilde{m}_2^n(\cdot,\cdot)$ are the same as in the case of edge reduction.

Every stochastic graph $((W_0, g_0), e)$ can be reduced according to the algorithms described in Section 3 and formulas (4.2)–(4.5) to a stochastic graph $((W_r, g_r), e)$, $W_r = \{e\} \cup S$ determine by the strategy $(\theta_0, \theta_1, \ldots)$ with matrix of distributions of time of transition in "one step" between its states $\mathbb{P}_r = \left[p_r^{(\cdot)}(a,b)\right]_{a,b\in W_r}$ and the matrix of nth moments $M_r^n = [m_r^n(a,b)]_{a,b\in W_r}$. Therefore $T_s \overset{d}{=} p_r^{(\cdot)}(e,s)$ and $\mathbb{E}(T_s)^n = m_r^n(e,s)$, for all $s \in S$ and all $n \in \mathbb{N}$.

4.4. An example. Now we apply the described algorithms to solve a classical problem: the bold gamble.

Example 4.1. We have 2$ and you need 5$, we can reach our goal by a fair gamble. We decide on the bold strategy: at each time we stake so much of our current fortune that we come as close to our goal as possible, if we win. The bold gamble can be translated into the following stochastic graph:

$$W_0 = \{0,1,2,3,4,5\},\ S = \{0,5\},\ a_0 = 2,\ g_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \end{bmatrix},$$

where in first, second, third, fourth, fifth, sixth row of g_0 there are probabilities of transition from state 0, 5, 1, 2, 3, 4 respectively. We compute probabilities of absorption and first two moments of time to absorption. Denote by M_0^1, M_0^2 the matrices of first and second moments i.e. $M_0^1 = M_0^2 = \left[\mathbf{1}_{\{(i,j):g_0(i,j)>0\}}(i,j)\right]$. Notice that there is only one edge $(2,4)$ directed to the state 4, so we can reduce it. After that we obtain

$$g_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{bmatrix},\ \tilde{M}_1^1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 2 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix},$$

$$\tilde{M}_1^2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 4 & 0 & 0 & 4 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}.$$

Next we reduce the state 3

$$g_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{8} & \frac{1}{8} & 0 \end{bmatrix}, \ \tilde{M}_2^1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & \frac{7}{3} & 3 & 0 \end{bmatrix}, \ \tilde{M}_2^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & \frac{17}{3} & 9 & 0 \end{bmatrix}$$

and the edge $(1, 2)$

$$g_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{3}{4} & \frac{3}{16} & \frac{1}{16} & 0 \\ \frac{1}{2} & \frac{3}{8} & \frac{1}{8} & 0 \end{bmatrix}, \tilde{M}_3^1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{4}{3} & \frac{10}{3} & 4 & 0 \\ 1 & \frac{7}{3} & 3 & 0 \end{bmatrix}, \tilde{M}_3^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & \frac{34}{3} & 16 & 0 \\ 1 & \frac{17}{3} & 9 & 0 \end{bmatrix}.$$

After these three reductions we obtain a loop $(1, 1)$. We reduce the loop $(1, 1)$, thus

$$g_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{4}{5} & \frac{1}{5} & 0 & 0 \\ \frac{1}{2} & \frac{3}{8} & \frac{1}{8} & 0 \end{bmatrix}, \ \tilde{M}_4^1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{8}{5} & \frac{18}{5} & 0 & 0 \\ 1 & \frac{7}{3} & 3 & 0 \end{bmatrix},$$

$$\tilde{M}_4^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3.92 & 14.32 & 0 & 0 \\ 1 & \frac{17}{3} & 9 & 0 \end{bmatrix}.$$

Finally we reduce the state 1.

$$g_r = \begin{bmatrix} 1 & 0 & - & 0 \\ 0 & 1 & - & 0 \\ - & - & - & - \\ \frac{3}{5} & \frac{2}{5} & - & 0 \end{bmatrix}, \ M_r^1 = \begin{bmatrix} 1 & 0 & - & 0 \\ 0 & 1 & - & 0 \\ - & - & - & - \\ \frac{8}{5} & \frac{13}{5} & - & 0 \end{bmatrix},$$

$$M_r^2 = \begin{bmatrix} 1 & 0 & - & 0 \\ 0 & 1 & - & 0 \\ - & - & - & - \\ \frac{1376}{300} & 8.12 & - & 0 \end{bmatrix}.$$

References

[1] Billingsley, P., *Probability and Measure*, John Wiley & Sons, New York, 1979.

[2] Chandrasekharan, M. P., Madhusudanan Pillai, V., *An absorbing Markov chain model for production systems with rework and scrapping*, Computers and Industrial Engineering **55**, Issue 3 (2008), 695–706.

[3] Engel, A., *The probabilistic abacus*, Educational Studies in Mathematics **6** (1975), 1–22.

[4] Engel, A., *Why does the probabilistic abacus work?*, Educational Studies in Mathematics **7** (1976), 59–56.

[5] Feller, W., *An Introduction to Probability Theory and Its Applications*, John Wiley & Sons, New York, 1968.

[6] Gosselin, F., *Asymptotic behavior of absorbing Markov chains conditional on nonabsorption for application in conservation biology*, Ann. Appl. Probab. **11**, no. 1 (2001), 261–284.

[7] Iosifescu, M., *Finite Markov Processes and Their Applications*, John Wiley & Sons, Chichester, 1980.

[8] Kemeny, J. G., Snell, J. L., *Finite Markov Chain*, D. Van Nostrand, Princeton, 1960.

[9] Keming Gu, Sadiku, M. N. O., *Absorbing Markov chain solution for Possion's equation*, Southeastcon 2000. Proceedings of the IEEE (2000), 297–300.

[10] Norris, J. R., *Markov Chain*, Cambridge University Press, Cambridge, 1997.

[11] Płocki, A., *Introduction to probability calculus and mathematical statistics for teachers*, Propedeutyka rachunku prawdopodobieństwa i statystyki matematycznej dla nauczycieli, PWN, Warszawa 1992 (Polish).

[12] Swan, Y. C., Bruss, F. T., *A matrix-analytic approach to the N-player ruin problem*, J. Appl. Probab. **43**, no. 3 (2006), 755–766.

Mariusz Górajski
The Faculty of Mathematics and Computer Science
University of Łódź
ul. Stefana Banacha 22
90-238 Łódź
Poland
e-mail: `mariuszg@math.uni.lodz.pl`

Received September 17, 2008