

# Relations between irreducible and absorbing Markov chains

G. Rubino

AMS Sectional Meeting

Riverside, November 2019

# Introduction

- For illustration purposes, or for reference, I will take as global application areas of the talk, the quantitative analysis of (complex) systems using Markov models. This can be decomposed into
  - performance evaluation of systems,
  - dependability evaluation of systems.
- The typical model for performance evaluation is the queue. The Markov model is typically irreducible, it lives in continuous time, and the analysis is typically performed in equilibrium. Both finite and infinite processes appear frequently.
- In dependability, the model is typically absorbing (sooner or later the system is dead), it is also in continuous time, and in general the process is finite.

## A reference performance model

- $X$ : the  $M/M/1$  queue

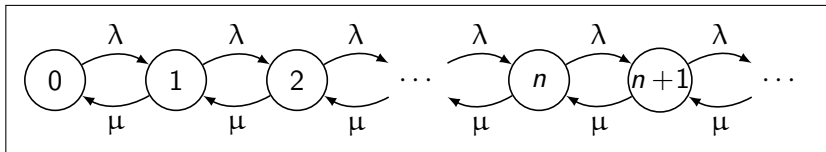


Figure: The  $M/M/1$  basic model.

- The queue is stable iff  $\lambda < \mu$ . In this case, the stationary distribution is given by  $\pi_n = (1 - \rho)\rho^n$ ,  $n \in \mathbb{N}$ , with  $\rho = \lambda/\mu < 1$ .
- Typical metric of interest: the mean # of units in the queue,

$$\mathbb{E}(X(\infty)) = \sum_{n \in \mathbb{N}} n \pi_n = \frac{\rho}{1 - \rho}.$$

## A classic dependability model

- 2 independent components with failure rate  $\lambda$ ; if one alive and one down, the latter repairs the former with rate  $\mu$ ; if both down, system is dead.

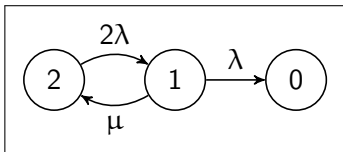


Figure: A basic dependability model.

- Observe that  $\pi$  is trivial here, but useless: taking the states in the order  $(2, 1, 0)$ , we have, whatever the initial state,  $\pi = (0 \ 0 \ 1)$ .
- Typical object of interest: the r.v. “system lifetime  $T$ ”. For instance, starting at state 2,

$$\mathbb{E}(T) = \frac{3\lambda + \mu}{2\lambda^2}.$$

## Solving

- Computing the stationary distribution  $\pi$  of an irreducible chain, or computing  $\mathbb{E}(T)$  with  $T$  the absorption time of an absorbing chain, are both *linear* problems:
  - in the first case,  $\pi$  is the unique solution to the equilibrium equations having unit norm;
  - in the second case, denote by  $P_{u,v}$  the transition probability of moving from  $u$  to  $v$  ( $P_{u,v} =$  transition rate  $Q_{u,v}$  of  $u \rightarrow v$  divided by departure rate  $d_u$  from  $u$ ); then, writing  $\alpha_x = \mathbb{P}(X(0) = x)$  and denoting  $\tau_x = \mathbb{E}(T \mid X(0) = x)$ , we have that  $(\tau_x)$  is the solution to the system

$$\tau_x = \frac{1}{d_x} + \sum_{y: y \neq x, a} P_{x,y} \tau_y, \quad \text{for all state } x \neq a.$$

- Next result connects the two problems.

## First result

- **Theorem 1:**

- Start from an absorbing chain  $X$  on the finite state space  $S$ , with a single absorbing state  $a$  and initial distribution  $\alpha$ .
- To avoid trivialities, assume that  $\alpha_a = 0$  and that from every state  $x \neq a$  there is a path to state  $a$ . This means that every state  $x \neq a$  is transient.
- Build an irreducible process  $Y$  from  $X$  by adding to  $X$  the transition  $(a, x)$  with rate  $r \alpha_x$ , for all  $x \neq a$ , where  $r$  is an arbitrary real  $> 0$ .
- Let  $\pi$  be the stationary distribution of  $Y$  and let  $T$  be the absorption time of  $X$ . Then,

$$\mathbb{E}(T) = \frac{1 - \pi_a}{r \pi_a},$$

or equivalently,

$$\pi_a = \frac{1}{1 + r \mathbb{E}(T)}.$$

## Example

- Using previous example in dependability, the  $Y$  irreducible model is the following:

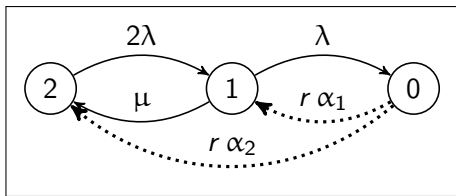


Figure:  $Y$ :  $X$  with feedback, where  $X$  is previous 3-state dependability example;  $\alpha_2 + \alpha_1 = 1$ .

- On  $Y$ , we have

$$\pi = (\pi_2 \quad \pi_1 \quad \pi_0) = \frac{1}{\frac{r(\alpha_2\lambda + \mu)}{2\lambda^2} + \frac{r}{\lambda} + 1} \begin{pmatrix} \frac{r(\alpha_2\lambda + \mu)}{2\lambda^2} & \frac{r}{\lambda} & 1 \end{pmatrix}.$$

- If we use  $\alpha_2 = 1$ ,  $\alpha_1 = 0$ , we obtain

$$\pi = \frac{1}{\frac{r(\lambda + \mu)}{2\lambda^2} + \frac{r}{\lambda} + 1} \begin{pmatrix} \frac{r(\lambda + \mu)}{2\lambda^2} & \frac{r}{\lambda} & 1 \end{pmatrix}$$

and we easily check here the formulas of previous Theorem 1.



## Proof

- Based on the concept of pseudo-aggregation:
  - Let  $U$  be an irreducible Markov model and let  $\mathcal{C} = (C_k)_{k \in \mathcal{K}}$  be a partition of its state space.
  - The aggregation of  $U$  with respect to  $\mathcal{C}$  is the process  $V$  with values on  $\mathcal{K}$ , defined by  $V(t) = k \iff U(t) \in C_k$ .
  - In general,  $V$  is not Markov, not homogenous, etc.
  - But we can always *define* a Markov process with values on  $\mathcal{K}$ , the *pseudo-aggregation* of  $U$  w.r.t.  $\mathcal{C}$ , using the stationary distribution  $\pi$  of  $U$ .
  - In continuous time, for instance, if  $Q$  is the rate matrix of  $U$ , the rate matrix  $\hat{Q}$  of  $V$  is *defined* by

$$\hat{Q}_{m,\ell} = \frac{\sum_{x \in C_m} \pi_x Q_{x,C_\ell}}{\pi_{C_m}}.$$

- Then, if  $\hat{\pi}$  denotes the stationary distribution of  $V$ , we have that for all class  $C_k$ ,

$$\hat{\pi}_k = \pi_{C_k}.$$

## Cut Lemma

- Consider an irreducible and ergodic Markov process  $X$  on the state space  $S$ , with stationary distribution  $\pi$ .
- Let  $(B, C)$  be a partition of  $S$ .
- Then, we have equilibrium or mean flow conservation between the two subsets of states  $B$  and  $C$ :

$$\sum_{x \in B} \pi_x Q_{x,C} = \sum_{y \in C} \pi_y Q_{y,B}.$$

- This is also called Kelly's Lemma (Lemma 1.4 in *Reversibility and Stochastic Networks*, Frank Kelly, Cambridge University Press, 2011).
- The Cut Lemma is a very useful tool in deriving equations necessarily satisfied by the stationary distribution of an irreducible Markov process, when such a distribution exists.

## Cuts in absorbing chains

- Consider an absorbing finite Markov chain  $X$  on  $S$ , with a single absorbing state  $a$ ; assume all states in  $S \setminus \{a\}$  are transient.
- If we index the states putting  $a$  at the end, the limiting distribution of  $X$  is  $\pi = (0 \ 0 \ \cdots \ 0 \ 1)$ , as stated before, trivial and useless.
- Consider a partition  $(B, C)$  of  $S \setminus \{a\}$ .
- Using Theorem 1, we can connect  $X$  to an irreducible process  $Y$  on  $S$ , then apply the Cut Lemma, and come back to  $X$ .
- This idea leads to new properties now related to cuts in absorbing models.

## Relevant variables in this context

- The variables that appear following this idea are the  $\sigma_{i,x}$ s and  $\sigma_x$ s where

$$\sigma_x = \mathbb{E} \left( \int_0^\infty 1(X(t) = x) dt \right),$$

and

$$\sigma_{i,x} = \mathbb{E} \left( \int_0^\infty 1(X(t) = x) dt \mid X(0) = i \right).$$

- In words,  $\sigma_x$  is the mean total time spent by  $X$  in state  $x$  until absorption, and  $\sigma_{i,x}$  is the same average but conditional to starting at state  $i$ .
- Matrix  $(\sigma_{u,v})_{u,v}$  is also called the *potential matrix* of the chain.
- How to compute these metrics?

- We have an absorbing chain  $X$  on the finite state space  $S$ , with a single absorbing state  $a$  and initial distribution  $\alpha$  (with  $\alpha_a = 0$ ). From every state  $x \neq a$  there is a path to state  $a$ ; then, every state  $x \neq a$  is transient.
- The  $\sigma_{i,x}$ s are the unique solution to the following linear system: fix  $i \neq a$ ; for any  $x \neq i, a$ ,

$$\sigma_{i,x} = \sum_{j:j \neq a} P_{i,j} \sigma_{j,x};$$

in the case of  $x = i$ ,

$$\sigma_{x,x} = \frac{1}{d_x} + \sum_{j:j \neq a} P_{x,j} \sigma_{j,x}.$$

- For the unconditional case,

$$\sigma_x = \sum_i \alpha_i \sigma_{i,x}.$$

## “Conditional” Cut Lemma

Using Theorem 1, we obtain

- **Theorem 2:** In the context of previous setting, for any  $i \in B$ ,

$$\sum_{j \in B} \sigma_{i,j} (Q_{j,C} + Q_{j,a}) = 1 + \sum_{k \in C} \sigma_{i,k} Q_{k,B}.$$

Of course, if  $i \in C$ , we symmetrically write

$$1 + \sum_{j \in B} \sigma_{i,j} Q_{j,C} = \sum_{k \in C} \sigma_{i,k} (Q_{k,B} + Q_{k,a}).$$

- **Equivalent formulation:** Let  $N_{B,C}$  be the r.v. “# of transitions from  $B$  to  $C$  before absorption” and analogously for  $N_{C,B}$ ,  $N_{B,a}$ , etc. Observe that  $N_{B,a} + N_{C,a} = 1$ . We have

- $\sum_{j \in B} \sigma_{i,j} Q_{j,C} = \mathbb{E}_i(N_{B,C}),$
- $\sum_{k \in C} \sigma_{i,k} Q_{k,B} = \mathbb{E}_i(N_{C,B}),$
- $\sum_{j \in B} \sigma_{i,j} Q_{j,a} = \mathbb{E}_i(N_{B,a}) = \mathbb{P}(X \text{ gets absorbed from } B).$

- After some algebra, this leads to

$$\mathbb{E}_i(N_{B,C}) = \mathbb{E}_i(N_{C,B}) + \mathbb{P}_i(X \text{ gets absorbed from } C) \quad (i \in B).$$

- Again, we can write this in the case we called  $C$  the partition class containing the initial state  $i$ .

## “Unconditional” Cut Lemma

- **Theorem 3:** In the context of previous setting,,

$$\sum_{j \in B} \sigma_j (Q_{j,C} + Q_{i,a}) = \alpha_B + \sum_{k \in C} \sigma_k Q_{k,B}.$$

- **Equivalent formulation:**

$$\mathbb{E}(N_{B,C}) + \mathbb{P}(X \text{ gets absorbed from } B) = \mathbb{E}(N_{C,B}) + \alpha(B).$$

By symmetry, we also have

$$\mathbb{E}(N_{B,C}) + \alpha(C) = \mathbb{E}(N_{C,B}) + \mathbb{P}(X \text{ gets absorbed from } C).$$



## Illustration

- As we know, the Cut Lemma (irreducible models) is typically useful when we have a state space with a tree-like topology, for instance, a birth-death process.
- Consider now the absorbing process  $X^{(n)}$  with the dynamics indicated in the picture.

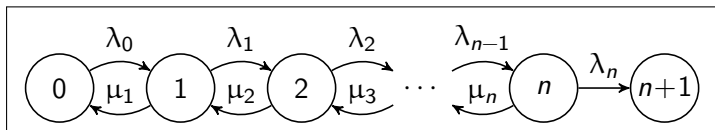


Figure: A quasi-birth-death process  $X^{(n)}$ .

Assume we need to compute the  $\sigma_j$ s,  $j = 0, 1, \dots, n$ , when the process starts, say, at state 0.

- Using the basic relations given on slide 13, we obtain 3-order difference equations.
- Using our Lemmas for absorbing processes, we immediately write

$$\sigma_{0,j}\lambda_j = 1 + \sigma_{0,j+1}\mu_{j+1}, \quad j = 0, 1, \dots, n,$$

an order-2 difference equation.

This comes from the partition  $(\{0, 1, \dots, j\}, \{j + 1, \dots, n\})$  of  $S \setminus \{n + 1\}$ .

- From this, after some algebra, we obtain

$$\sigma_{0,j} = \frac{1 + \Psi_{j+1} + \Psi_{j+2} + \dots + \Psi_n}{\lambda_j},$$

where

$$\Psi_h = \frac{\mu_h \mu_{h+1} \cdots \mu_n}{\lambda_h \lambda_{h+1} \cdots \lambda_n}.$$

# Applications

- We have two kinds of applications of these results that we can't develop here for lack of room/time.
- First, these properties can lead to numerical procedures that can analyze both
  - irreducible models, moving the problem to a new one defined on an absorbing chain,
  - or the other way: the problem is set in terms of absorbing processes, and it is solved by moving it to an irreducible setting.
- We started to explore the first path by means of an extension of Theorem 1. No time to develop this further. See in `arxiv` soon.

- Second, they are useful for deriving bounds of metrics of interest. So far, we did it for metrics defined on irreducible models, such as the asymptotic availability in dependability analysis, or more generally, the mean reward when extending the model to a Markov process with rewards on the transitions or on the states.
- Again, it isn't possible to expand this here. See in [arxiv](#) soon, or attend the JMM meeting at Denver, Colorado, in next January, for details on this work.