



HAL
open science

The Exponential-Dual Matrix Method: Applications to Markov Chain Analysis

Gerardo Rubino, Alan Krinik

► **To cite this version:**

Gerardo Rubino, Alan Krinik. The Exponential-Dual Matrix Method: Applications to Markov Chain Analysis. R. Swift, A. Krinik, J. Switkes and J. Park. Stochastic Processes and Functional Analysis: New Perspectives, 774, , pp.1-23, 2021, Contemporary Mathematics. hal-03260400

HAL Id: hal-03260400

<https://inria.hal.science/hal-03260400v1>

Submitted on 14 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Exponential-Dual Matrix Method: Applications to Markov Chain Analysis

Gerardo Rubino

Alan Krinik

INRIA, CAMPUS DE BEAULIEU, 35042 RENNES, FRANCE
Email address: `gerardo.rubino@inria.fr`

CAL POLY POMONA, 3801 WEST TEMPLE AVENUE, POMONA, CA
91768

Email address: `ackrinik@cpp.edu`

*Dedicated to two beloved colleagues of M. M. Rao at Univ. of California, Riverside:
Professor Neil E. Gretskey (1941-2015; also a graduate student of M. M. Rao at
Carnegie-Mellon Univ., graduated 1967) and Professor Victor L. Shapiro (1924-2013)*

1991 *Mathematics Subject Classification*. 60J27, 60J35, 60K25, 15A04,
15A16

Key words and phrases. Markov Chains, Markov Processes, transient analysis, dual process, exponential-dual matrix, duality, generator, exponential matrix, transition rate matrix, catastrophes, uniformization, closed-forms.

ABSTRACT. Classic performance evaluation using queueing theory is usually done assuming a stable model in equilibrium. However, there are situations where we are interested in the transient phase. In this case, the main metrics are built around the model's state distribution at an arbitrary point in time. In dependability, a significant part of the analysis is done in the transient phase. In previous work, we developed an approach to derive distributions of some continuous time Markovian models, built around uniformization (also called Jensen's method), transforming the problem into a discrete time one, and the concept of stochastic duality. This combination of tools provides significant simplifications in many cases. However, stochastic duality does not always exist. Recently, we discovered that an idea of algebraic duality, formally similar to stochastic duality, can be defined and applied to any linear differential system (or equivalently, to any matrix). In this case, there is no limitation, the transformation is always possible. We call it the exponential-dual matrix method. In the article, we describe the limitations of stochastic duality and how the exponential-dual matrix method operates for any system, stochastic or not. These concepts are illustrated throughout our article with specific examples, including the case of infinite matrices.

1. Introduction

In this article, we first review how the concept of stochastic duality in continuous time Markov chains with discrete state spaces, as defined in [16] and used in [1], coupled with the uniformization method [6], allows us to obtain analytical expressions of transient distributions of fundamental Markovian queueing models. The quantity of literature on the transient analysis of these stochastic models is huge, both in the number of research papers and in books [12], [4], [2], [15]. There are many uses of the word “duality” in science, and also inside probability theory (see for instance [5] in the Markovian case). When we say “dual” here, or more precisely “stochastic dual”, we mean for Markov chains, as defined in [16] and developed by Anderson in [1].

Analyzing transient behavior in stochastic models is extremely important in performance and in dependability analysis, in many engineering branches including computer science and communication networks. The combination of duality with uniformization (also called randomization or Jensen's method) has proved very useful in obtaining closed-form expressions, see our previous articles [3, 9, 10, 8]. In particular, in [10] the

$M/M/1$ and $M/M/1/H$ queueing systems are analyzed using these methods, together with a variant of the $M/M/1/H$ where “catastrophes” (transitions from every state to the empty state) are also included.

However, the use of the duality transformation has some limitations. For example, the stochastic dual exists only if a strong monotonicity property of the original chain holds, see [1] (a property that makes sense with respect to a total order relation on the state space). The problem is that in some cases, there is no such monotonicity and thus, no dual process. Moreover, the ordering matters, which means that the dual may exist for a specific ordering of the states, and not for a different one. Another possible restriction is that the dual may exist iff the transition rates of the original Markov chain satisfy some specific inequalities.

In this article, we first describe the stochastic dual concept and review how the method is used to find analytic expressions of the transient distribution of some fundamental Markov chain models. In particular, we highlight the main structure of the duality/uniformization approach. Examples when the dual does not exist and the dependency upon ordering of states or other required conditions are explored. Another possibility not usually covered in the literature, occurs when the state space is infinite making the dual transformation a nonconservative matrix. We illustrate this with an example and show how to deal with it and still obtain transient distributions (Subsection 4.1).

In our work with stochastic duality, we realized that many nice properties of duality rely only upon algebraic relations and not upon the stochastic properties of the models such as monotonicity. In fact, a similar transformation can be defined for any matrix, that is, not necessarily just for infinitesimal generators. This generalization of the stochastic dual is what we call “exponential-dual”. It turns out that many of the main properties of the dual concept for Markov chains also hold unchanged for the exponential-dual. However, the exponential-dual always exists, independently of the specific ordering of the states and without any condition of the model’s parameters. Of course, when we deal with a Markov model and the dual exists, the exponential-dual coincides with it.

Our article has the following outline. This introductory Section 1 describes its context and content. Uniformization is discussed in Section 2 and Section 3 is devoted to stochastic duality. The *uniformization plus duality* approach to transient solutions is explained and illustrated by examples in Section 4. The generalization of the stochastic dual to the exponential-dual matrix, along with its main properties and results, is presented in Section 5. In this article, we mainly consider finite matrices. Conclusions and future work comprise Section 6.

2. Uniformization

The uniformization procedure is a way of transforming a problem specified on a continuous time Markov chain $X = \{X(t)\}_{t \geq 0}$ into a similar problem defined on an associated discrete time Markov chain $Y = \{Y_n\}_{n \geq 0}$. The transformation is such that solving the problem defined on Y allows to get immediately an answer to the original problem defined in terms of X . Both chains X and Y share the same discrete state space S and the same initial distribution α (when writing matrix relations using α , it will be considered a row vector). Let A be the infinitesimal generator of X , that is, $A_{i,j} \geq 0$ is the transition rate from i to j for $i \neq j$; $A_{i,i} = -d_i \leq 0$ is the departure rate from state i multiplied by -1 . The transformation has a scalar real parameter Λ that must satisfy $\Lambda \geq \sup_{i \in S} d_i$. This is why X is said to be *uniformizable* iff $\sup_{i \in S} d_i < \infty$. We are only interested in uniformizable processes here. Let U be the transition probability matrix of chain Y . Its definition is $U = I + A/\Lambda$, where I is the identity matrix indexed on S .

If $N = \{N(t)\}_{t \geq 0}$ is the counting process of a Poisson process having rate Λ and independent of Y , then the process $Z = \{Z(t)\}_{t \geq 0}$ defined on S by $Z(t) = Y_{N(t)}$ is stochastically equivalent to X . In particular, this means that the distribution of $X(t)$, seen as a row vector $p = (p(t))$, satisfies

$$(2.1) \quad p(t) = \sum_{n \geq 0} e^{-\Lambda t} \frac{(\Lambda t)^n}{n!} q_n,$$

where (q_n) is the distribution of Y_n also considered a row vector indexed on S . To evaluate, for instance, the law of $X(t)$ with an absolute error $< \varepsilon$, we use the previous relation and look first for integer N defined by

$$N = \min \left\{ K \geq 0 : \sum_{k=0}^K e^{-\Lambda t} \frac{(\Lambda t)^k}{k!} \geq 1 - \varepsilon \right\}.$$

This can be done in negligible computing time. Using the vector norm $\|\cdot\| = \|\cdot\|_1$, we have

$$\begin{aligned} \left\| p(t) - \sum_{n=0}^N e^{-\Lambda t} \frac{(\Lambda t)^n}{n!} q_n \right\| &= \left\| \sum_{n > N} e^{-\Lambda t} \frac{(\Lambda t)^n}{n!} q_n \right\| \\ &\leq \sum_{n > N} e^{-\Lambda t} \frac{(\Lambda t)^n}{n!} \|q_n\| \\ &\leq \sum_{n > N} e^{-\Lambda t} \frac{(\Lambda t)^n}{n!} \\ &= 1 - \sum_{n=0}^N e^{-\Lambda t} \frac{(\Lambda t)^n}{n!} \\ &\leq 1 - (1 - \varepsilon) = \varepsilon. \end{aligned}$$

Now, let us illustrate what happens if we attack the problem of evaluating the transient distribution of the $M/M/1$ queue using uniformization. Consider this chain, illustrated in FIGURE 1, with parameters $\lambda > 0$ and $\mu > 0$. This chain is uniformizable, and we can use any uniformization rate $\Lambda \geq \lambda + \mu$.

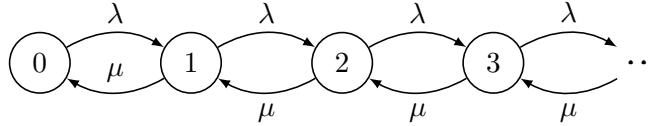


FIGURE 1. The $M/M/1$ model with arrival rate λ and service rate μ .

Uniformizing this chain with respect to the uniformizing rate $\Lambda = \lambda + \mu$ leads to the discrete time chain depicted in FIGURE 2. If we can calculate the

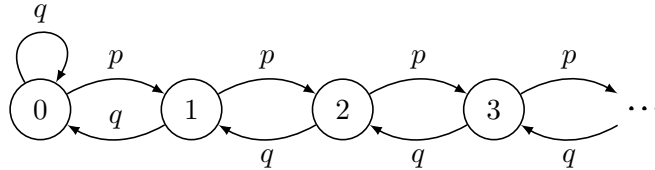


FIGURE 2. The uniformized $M/M/1$ model of FIGURE 1, with respect to the uniformizing rate $\Lambda = \lambda + \mu$, with the notation $p = \lambda/\Lambda$ and $q = 1 - p = \mu/\Lambda$.

transient distribution of Y , which we denoted (q_n) , we obtain an expression of that of the original $M/M/1$ queue using (2.1). This can be done using counting path techniques as described in [9, 10] for basic queues such as the $M/M/1$ or the $M/M/1/H$ having a finite storage capacity, and variations.

In next section, we explore the concept of duality and its limitations.

3. Stochastic duality

In this section we follow [1] using a notation slightly different from Anderson's. We start from a continuous time Markov chain X where the state space is $S = \mathbb{N}$ or $S = \{0, 1, 2, \dots, n-1\}$, for some integer $n \geq 1$. Define matrix $P = P(t)$ by $P_{i,j}(t) = \mathbb{P}(X(t) = j \mid X(0) = i)$, often also denoted $\mathbb{P}_i(X(t) = j)$. Recall that A denotes X 's generator. When $\exp(A)$ exists (for instance, when S is finite), we have $P(t) = \exp(At)$. Seen as a function of t , matrix P is called the *transition function* of X . In our case, where we avoid all possible "pathologies" of X , function P has all the information that we need to work with X .

To fix the ideas and simplify the presentation, consider the case of $S = \mathbb{N}$. The transition function P of the continuous time Markov chain X is said

to be *stochastically increasing* iff for all $t \geq 0$ and for all states $i, j, k \in \mathbb{N}$, the ordering $\mathbb{P}_i(X(t) \geq k) \leq \mathbb{P}_{i+1}(X(t) \geq k)$ holds. In [16], Siegmund proved that if the transition function P of Markov chain X is stochastically increasing, then there exists another Markov process X^* defined also on \mathbb{N} , such that for all $t \geq 0$ and $i, j \in \mathbb{N}$, $\mathbb{P}_i(X^*(t) \leq j) = \mathbb{P}_j(X(t) \geq i)$. Process X^* is the dual of X . We will say here stochastic dual, to underline the difference with our generalization in Section 5. Matrix function P^* is also stochastically increasing. Between P and P^* , the following relations hold:

$$(3.1) \quad P_{i,j}^*(t) = \sum_{k=0}^{i-1} [P_{j-1,k}(t) - P_{j,k}(t)],$$

$$(3.2) \quad P_{i,j}(t) = \sum_{k=0}^i [P_{j,k}^*(t) - P_{j+1,k}^*(t)].$$

By convention, if some index takes a value not in S , then the corresponding term is 0, and if the index summation space is empty, then the sum is 0. In the finite case where the state space of X has n states (we will fix them to the set $\{0, 1, \dots, n-1\}$), then the state space of X^* is $\{0, 1, \dots, n\}$. Relation (3.1) holds for $j \leq n-1$, and this makes that $P_{n,j}^*(t) = 0$. When $j = n$, we have $P_{i,n}^*(t) = 1 - \sum_{k=0}^{n-1} P_{i,k}^*(t)$, and $P_{n,n}^*(t) = 1$.

Relations (3.1) and (3.2) can be also written

$$(3.3) \quad \mathbb{P}_i(X(t) = j) = \mathbb{P}_j(X^*(t) \leq i) - \mathbb{P}_{j+1}(X^*(t) \leq i),$$

$$(3.4) \quad = \mathbb{P}_{j+1}(X^*(t) > i) - \mathbb{P}_j(X^*(t) > i),$$

$$(3.5) \quad \mathbb{P}_i(X^*(t) = j) = \mathbb{P}_{j-1}(X(t) \leq i-1) - \mathbb{P}_j(X(t) \leq i-1),$$

$$(3.6) \quad = \mathbb{P}_j(X(t) > i) - \mathbb{P}_{j-1}(X(t) > i),$$

where we see monotonicity at work. This is the central result of [16], called the Duality Theorem in [8].

Between X and X^* we also have other similar relations: if A^* is the generator of X^* ,

$$(3.7) \quad A_{i,j}^* = \sum_{k=0}^{i-1} [A_{j-1,k} - A_{j,k}] = \sum_{k \geq i} [A_{j,k} - A_{j-1,k}],$$

$$(3.8) \quad A_{i,j} = \sum_{k=0}^i [A_{j,k}^* - A_{j+1,k}^*] = \sum_{k \geq i-1} [A_{j+1,k}^* - A_{j,k}^*].$$

If we consider discrete time Markov chains instead, the construction remains valid, changing (3.1) to (3.8) to their immediate discrete versions (in the last two relations (3.7) and (3.8), replacing generators by transition probability matrices).

Suppose that the continuous time Markov chain X has generator A , and that we define A^* using (3.7). If A^* is also a generator and if we build a continuous time Markov chain X^* on \mathbb{N} having generator A^* , then X^* is the

dual of X . The analogous reciprocal property holds if we start from A^* and use (3.8) to build A . These relations also hold if we consider discrete time Markov chains. See [7] for some elements of stochastic duality in discrete time.

The stochastic dual matrix A^* has some interesting properties that will be addressed in later sections. For example, it turns out that A and A^* share the same spectrum (see [11] for details). In Section 5, some of these properties arise again with regard to exponential-dual matrices.

3.1. On the dual's existence. The monotonicity condition leading to the dual existence is a strong one. Let us illustrate this with a few very simple examples, using a generic one taken from [3]. See first FIGURE 3, where $\alpha, \beta, \gamma > 0$. The cyclic structure clearly suggests that the monotonicity property can't hold, which can easily be verified, and we then have an example of a Markovian model without a dual, for any value of the transition rates α, β, γ .

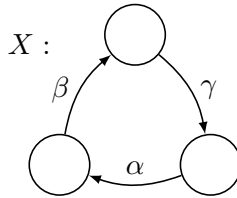


FIGURE 3. In this circular Markov process the transition diagram is not stochastically increasing, whatever the value of the rates *and whatever the numbering of the states*.

In FIGURE 4, we have a small birth-death process that satisfies the monotonicity condition for any value of its transition rates if the states are numbered as shown in part (a). This is valid for a general birth-death process [1]. If instead we number them as in part (b) of the same figure, the new process never has a dual. The verification of these claims is straightforward using the definition of duality and previous relations.

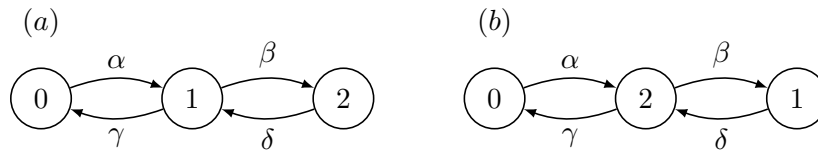


FIGURE 4. Assume that $\alpha, \beta, \gamma, \delta$ are all > 0 . In (a), we have a birth-death process that is always stochastically increasing. Changing the numbering of the states in (b) leads to a process that is never stochastically increasing, no matter which values α, β, γ and δ take.

Consider the Markov process in FIGURE 5. In previous examples, the dual process either existed or not, regardless of the values of the transition rates. Here, the dual exists iff $\beta > \nu$.

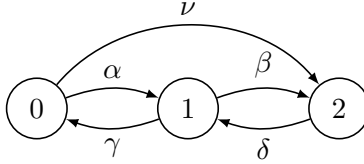


FIGURE 5. Assume that $\alpha, \beta, \gamma, \delta, \nu$ are all > 0 . The dual of this process exists iff $\beta > \nu$ (see [3]).

Before leaving this section, let us point out that the concept of stochastic duality appears in many studies after its definition by Siegmund in [16]. For a sample of recent work, see [14], [13], [18] and the references therein.

4. Transient analysis using uniformization and duality

Before going to the main topic of this section, observe that the basic uniformization relation (2.1) can be written in matrix form. Using the notation of previous section, we have

$$(4.1) \quad P(t) = \sum_{n \geq 0} e^{-\Lambda t} \frac{(\Lambda t)^n}{n!} U^n.$$

The basic idea of using the dual combined with uniformization to obtain the transient distribution of a given uniformizable continuous time Markov chain X defined on \mathbb{N} and having generator A , goes as follows. We first construct A^* and check if it is a generator. If it is, then we form the dual chain X^* of X and uniformize it. Call Y^* the result. We then determine the transient distribution of Y^* , which allows us through (2.1) to obtain that of X^* . Finally, (3.2) is used to obtain the distribution of X .

REMARK 4.1. In the process of determining Y^* , the order of the transformations can be reversed. That is, we can first uniformize X , to obtain Y , and then, in discrete time, construct the dual Y^* of Y . The result will be the same Y^* as before. FIGURE 6 illustrates that these operations commute. This is immediate from the definitions.

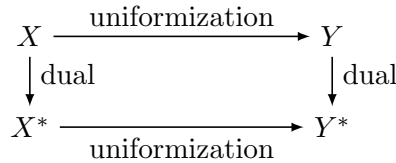


FIGURE 6. Commutativity of the operators “dual” and “uniformization”.

Here is a simple example to illustrate the preceding material. Consider the chain depicted in FIGURE 7. The dual of this chain is X^* , depicted in FIGURE 8.



FIGURE 7. The two-state irreducible continuous time Markov chain X , where $\lambda > 0$ and $\mu > 0$.



FIGURE 8. The dual of the chain X given in FIGURE 7.

Let us check the basic relations (3.4) and (3.6). It is straightforward to compute the transition functions of X and X^* . Using the notation $\Lambda = \lambda + \mu$, $p = \lambda/\Lambda$ and $q = \mu/\Lambda = 1 - p$, we obtain:

$$(4.2) \quad P(t) = e^{At} = \begin{pmatrix} pe^{-\Lambda t} + q & p(1 - e^{-\Lambda t}) \\ q(1 - e^{-\Lambda t}) & qe^{-\Lambda t} + p \end{pmatrix}$$

and

$$(4.3) \quad P^*(t) = e^{A^*t} = \begin{pmatrix} 1 & 0 & 0 \\ p(1 - e^{-\Lambda t}) & e^{-\Lambda t} & q(1 - e^{-\Lambda t}) \\ 0 & 0 & 1 \end{pmatrix}.$$

Then, for (3.4), we have

$$\begin{aligned} P_{0,0}(t) &= P_{0,0}^*(t) - P_{1,0}^*(t) = 1 - p(1 - e^{-\Lambda t}) = q + pe^{-\Lambda t}, \\ P_{1,0}(t) &= [P_{0,0}^*(t) - P_{1,0}^*(t)] + [P_{0,1}^*(t) - P_{1,1}^*(t)] \\ &= [q + pe^{-\Lambda t}] + [0 - e^{-\Lambda t}] = q(1 - e^{-\Lambda t}), \\ P_{0,1}(t) &= P_{0,0}^*(t) - P_{1,0}^*(t) = p(1 - e^{-\Lambda t}) - 0 = p(1 - e^{-\Lambda t}), \\ P_{1,1}(t) &= [P_{1,0}^*(t) - P_{2,0}^*(t)] + [P_{1,1}^*(t) - P_{2,1}^*(t)] \\ &= [p(1 - e^{-\Lambda t}) - 0] + [e^{-\Lambda t} - 0] = p + qe^{-\Lambda t}. \end{aligned}$$

We can also consider the uniformization of X and that of X^* with respect to the uniformization rate $\Lambda = \lambda + \mu$, respectively denoted by Y and Y^* . They are depicted in FIGURES 9 and 10.

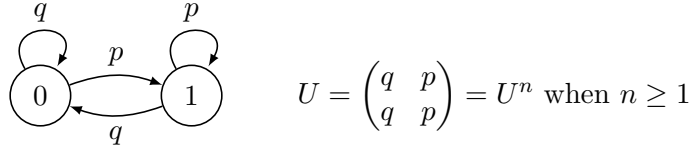


FIGURE 9. The uniformization of the chain X of FIGURE 7, denoted by Y , with respect to the uniformization rate $\Lambda = \lambda + \mu$, using the notation $\Lambda = \lambda + \mu$, $p = \lambda/\Lambda$ and $q = \mu/\Lambda = 1 - p$.

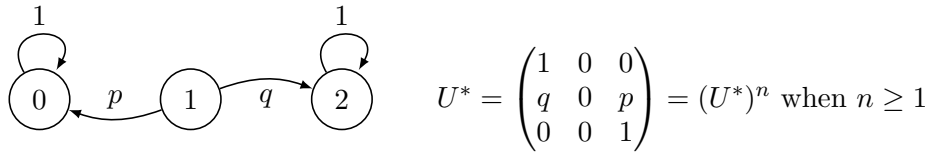


FIGURE 10. The uniformization of the chain X^* of FIGURE 8, denoted by Y^* , with respect to the uniformization rate $\Lambda = \lambda + \mu$, using the notation $\Lambda = \lambda + \mu$, $p = \lambda/\Lambda$ and $q = \mu/\Lambda = 1 - p$.

As a final check, consider (2.1), choosing the case of X . First, observe that (4.2) can be written

$$P(t) = U + e^{-\Lambda t} \begin{pmatrix} p & -p \\ -q & q \end{pmatrix} = U + e^{-\Lambda t}(I - U).$$

Now, following (4.1),

$$\begin{aligned} P(t) &= \sum_{n=0}^{\infty} e^{-\Lambda t} \frac{(\lambda + \mu)^n t^n}{n!} U^n \\ &= e^{-\Lambda t} I + \sum_{n=1}^{\infty} e^{-\Lambda t} \frac{(\lambda + \mu)^n t^n}{n!} U \\ &= e^{-\Lambda t} I + (1 - e^{-\Lambda t})U. \end{aligned}$$

The second equality uses the fact that $U^n = U$ when $n \geq 1$. The same computation can be checked on the pair X^*, Y^* .

REMARK 4.2. The use of the dual plus uniformization is useful for evaluating the distribution of $X(t)$ when finding the discrete time distribution of Y_n^* is easier than doing the same with $X(t)$, or with Y_n where Y is the uniformization of X (with respect to some appropriate rate Λ). This occurs, for instance, for queueing systems $M/M/1$ and $M/M/1/H$, where formal representations (closed forms) were obtained in this way [10].

4.1. Problems with infinite state spaces. Consider a simple immigration process with catastrophes, denoted as usual by X , our target, as illustrated in FIGURE 11.

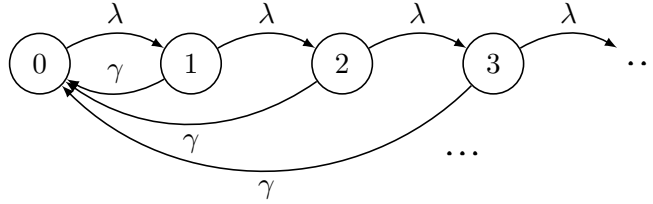


FIGURE 11. Immigration (or birth) process with “catastrophes”.

The transient analysis of this process has been the object of many papers, using different approaches (see for instance, [17]). It can also be done by means of the procedure described here. First, the generator of X is

$$(4.4) \quad A = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 & \dots \\ \gamma & -(\lambda + \gamma) & \lambda & 0 & 0 & \dots \\ \gamma & 0 & -(\lambda + \gamma) & \lambda & 0 & \dots \\ \gamma & 0 & 0 & -(\lambda + \gamma) & \lambda & \dots \\ & & \dots & & & \dots \end{pmatrix}.$$

Applying (3.7), we obtain

$$(4.5) \quad A^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \dots \\ \lambda & -(\lambda + \gamma) & 0 & 0 & 0 & \dots \\ 0 & \lambda & -(\lambda + \gamma) & 0 & 0 & \dots \\ 0 & 0 & \lambda & -(\lambda + \gamma) & 0 & \dots \\ & & \dots & & & \dots \end{pmatrix}.$$

This is not a generator. If we use the trick of adding an artificial auxiliary state Δ (as in [1, Proposition 1.1]) and the transitions that convert the new transition rate matrix into a generator, we obtain a Markov process whose graph is depicted in FIGURE 12.

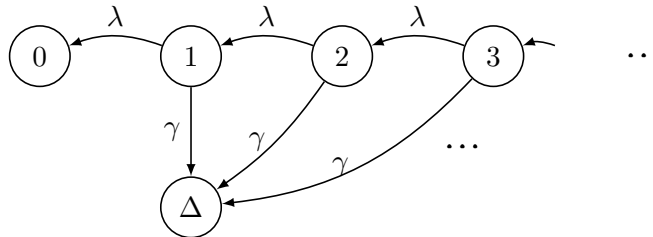


FIGURE 12. Dual of the immigration (or birth) process with “catastrophes” given in FIGURE 11.

If we decide that, by construction, Δ is greater than any integer i , then, the generator of X^* is symbolically given in (4.6),

$$(4.6) \quad A^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \lambda & -(\lambda + \gamma) & 0 & 0 & 0 & \cdots & \gamma \\ 0 & \lambda & -(\lambda + \gamma) & 0 & 0 & \cdots & \gamma \\ 0 & 0 & \lambda & -(\lambda + \gamma) & 0 & \cdots & \gamma \\ & & \cdots & & & & \end{pmatrix}.$$

where the index runs on $\{0, 1, 2, \dots\}$ followed by Δ . Now, it is easy to verify that recovering $A_{i,j}$ for $i, j \in \mathbb{N}$ from A^* works as before.

Denoting by Y^* the uniformization of the dual chain with respect to the rate $\Lambda = \lambda + \gamma$, with the additional notation $p = \lambda/\Lambda$ and $q = \gamma/\Lambda = 1 - p$, the resulting chain is shown in FIGURE 13.

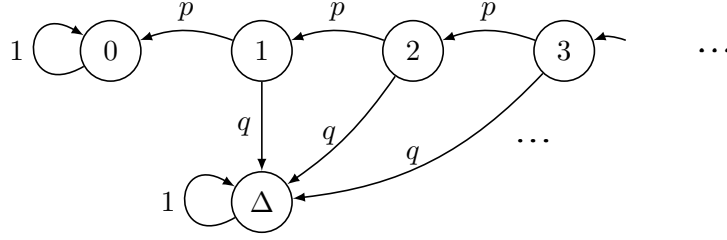


FIGURE 13. uniformization of the chain given in FIGURE 12, with respect to the uniformization rate $\Lambda = \lambda + \gamma$, with $p = \lambda/\Lambda$ and $q = \gamma/\Lambda = 1 - p$.

It is clear now how to obtain the transition function of this discrete time chain. If U^* is its transition probability matrix, and if we denote

$$((U^*)^n)_{i,j} = \mathbb{P}(Y^* = j \mid Y_0^* = i),$$

for any $n \geq 0$ and possible states i, j of Y^* , with $i \neq 0$, $i \neq \Delta$ and $j \in \mathbb{N} \cup \{\Delta\}$, we have:

- If $n = 0$, $((U^*)^0)_{i,j} = 1(i = j)$; for any $n \geq 0$, we have $((U^*)^n)_{0,0} = ((U^*)^n)_{\Delta,\Delta} = 1$.
- From this point, consider $n \geq 1$ and $i \neq 0$, $i \neq \Delta$ (so, $i \geq 1$).
 - Starting from i , we have that $((U^*)^n)_{i,j} \neq 0$ iff $j \neq \Delta$, $i > j$ and $n = i - j$, and in that case, its value is $p^{i-j} = p^n$.
 - For $n = 1, 2, \dots, i - 1$, $((U^*)^n)_{i,\Delta} = 1 - p^{n-1}$; then, for all $n \geq i$, $((U^*)^n)_{i,\Delta} = 1 - p^i$.

From these expressions, the distribution of X^* follows using (2.1), and then, that of X from the dual inversion formula (3.2). For instance, let us just check the $P_{0,0}$ case to compare with other papers, e.g. with [17]).

First, $P_{0,0}(t) = P_{0,0}(t) - P_{1,0}(t)$ using (3.2). We have $P_{0,0}^*(t) = 1$ and $P_{1,0}^*(t) = (1 - e^{-\Lambda t})p$. So,

$$P_{0,0}(t) = 1 - (1 - e^{-\Lambda t})p = q + pe^{-\Lambda t}.$$

5. Generalization of the stochastic-dual: the exponential-dual matrix

The transient distribution of the continuous time Markov process X is the solution to the Chapman-Kolmogorov equation $p'(t) = p(t)A$ where $p(t)$ is the distribution of $X(t)$ seen as a row vector, or in matrix form, using the transition function of X , $P'(t) = P(t)A$. In this section, we consider the finite case only. We know that the transition function $P(t)$ of X , the solution to the previous matrix differential equation, is $P(t) = e^{At}$ (in the infinite state space case, this exponential may not exist).

Assume now that we are given *an arbitrary matrix* A (possibly of complex numbers). Our goal is to solve the linear differential system whose matrix is A , that is, to compute e^{At} , that we denote here by $E(t)$, to avoid the notation $P(t)$ reserved for the Markov setting. We then define the exponential-dual of A as a new matrix A^* following the same formal rules as before.

We will also use the notation $E^*(t) = e^{A^*t}$ instead of $P^*(t)$ for the same reason as before: here, we have left the stochastic area.

As already said, we will limit ourselves to the case of finite chains, to focus on the algebraic work and not on the particular case of the analysis of queueing systems, where the state space may be infinite.

As before, our vectors are row vectors. When we will need a column vector we will use the transpose operator denoted $()^T$. For the sake of clarity, we denote by $\mathbf{0}$ a (row) vector only composed of 0's, and by $\mathbf{1}$ a (row) vector only composed of 1's.

5.1. Exponential-dual definition. Let A be an arbitrary square matrix of reals (or of complex numbers). We are interested in the computation of e^{At} for $t \geq 0$. Let $n < \infty$ be the dimension of A , whose elements are indexed on $\{0, 1, \dots, n-1\}$.

DEFINITION 5.1. We define the exponential-dual of A as the matrix of dimension $n+1$, indexed on $\{0, 1, \dots, n\}$, defined as follows: for any $i, j \in \{0, 1, \dots, n-1, n\}$,

$$A_{i,j}^* = \sum_{k=i}^{n-1} (A_{j,k} - A_{j-1,k}),$$

where $A_{u,v} = 0$ if (u, v) is out of the A -range (that is, if $(u, v) \notin \{0, 1, \dots, n-1\}^2$), and where we adopt the usual convention that $\sum_{k=u}^v \dots = 0$ if $u > v$.

We now describe some immediate properties of the exponential-dual matrix.

LEMMA 5.2 (Basic properties of the exponential-dual matrix). *Matrix A^* satisfies the following immediate properties:*

- the sum of the elements of any row of A^* is 0;
- the last row of A^* is only composed by 0's.

PROOF.

Case 1. Let $0 \leq i \leq n - 1$. Summing the elements of the i th row of A^* ,

$$\begin{aligned} \sum_{j=0}^n A_{i,j}^* &= \sum_{j=0}^n \sum_{k=i}^{n-1} (A_{j,k} - A_{j-1,k}) \\ &= \sum_{j=0}^n \sum_{k=i}^{n-1} A_{j,k} - \sum_{j=0}^n \sum_{k=i}^{n-1} A_{j-1,k} \\ &= \sum_{j=0}^{n-1} \sum_{k=i}^{n-1} A_{j,k} - \sum_{\ell=0}^{n-1} \sum_{k=i}^{n-1} A_{\ell,k} \\ &= 0, \end{aligned}$$

where in the penultimate '=' we use the fact that rows "-1" and "n" of A are out of the index space of the matrix (which is $\{0, \dots, n-1\}$), so, the corresponding elements of the matrix are all null.

Case 2. For the last row of A^* , the definition makes that the sum defining element $A_{n,j}$ is empty for any $j \in \{0, 1, \dots, n\}$, and then $A_{n,j}^* = 0$. \square

Let us illustrate previous definition.

$$\text{If } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \text{ then } A^* = \begin{pmatrix} a+b & c+d-(a+b) & -(c+d) \\ b & d-b & -d \\ 0 & 0 & 0 \end{pmatrix}.$$

$$\text{A numerical example: if } A = \begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix}, \text{ then } A^* = \begin{pmatrix} -1 & 0 & 1 \\ -2 & -2 & 4 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$\text{Another one: } A = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} \text{ (a generator), leads to } A^* = \begin{pmatrix} 0 & 0 & 0 \\ 1 & -3 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

(which is also a generator).

5.2. The exponential of the exponential-dual. Remember that we denote $E^*(t) = e^{A^*t}$. The given properties of A^* imply some general properties for $E^*(t)$.

LEMMA 5.3 (Initial properties of $E^*(t)$). *Matrix $E^*(t)$ satisfies the following properties:*

- the sum of the elements of any row of $E^*(t)$ is 1;

- the last row of $E^*(t)$ is composed of 0's except for its last element, $E_{n,n}^*(t)$, which is equal to 1.

PROOF.

Case 1. Let $0 \leq i \leq n$. Recall that $\mathbf{1}^T$ denotes a column vector only composed of 1s, whose dimension is defined by the context. We must prove that $E^*(t)\mathbf{1}^T = \mathbf{1}^T$. We know that $A^*\mathbf{1}^T = \mathbf{0}^T$, where $\mathbf{0}^T$ is a column vector only composed of 0s, whose dimension is defined by the context.

By definition,

$$E^*(t) = e^{A^*t} = I + \sum_{\ell \geq 1} \frac{t^\ell}{\ell!} (A^*)^\ell.$$

After right-multiplying by $\mathbf{1}^T$, we have

$$E^*(t)\mathbf{1}^T = \mathbf{1}^T + \sum_{\ell \geq 1} \frac{t^\ell}{\ell!} (A^*)^\ell \mathbf{1}^T = \mathbf{1}^T + \mathbf{0}^T = \mathbf{1}^T.$$

Case 2. For the last row of E^* , consider the decomposition of A^* in blocks as follows:

$$A^* = \left(\begin{array}{c|c} \widetilde{A}^* & -\widetilde{A}^*\mathbf{1}^T \\ \hline \mathbf{0} & 0 \end{array} \right).$$

This decomposition of A^* in blocks corresponds to the partition of $\{0, 1, \dots, n\}$ in two sets, $\{0, 1, \dots, n-1\}$ and $\{n\}$. If we index the block-decomposition on $\{0, 1\}$, block $(0, 0)$ is \widetilde{A}^* , with denotes the restriction of A^* to its first $n-1$ elements (square sub-matrix with dimension n); block $(0, 1)$ is the column vector $-\widetilde{A}^*\mathbf{1}^T$ (this follows from the fact that $A^*\mathbf{1}^T = \mathbf{0}^T$); block $(1, 0)$ is $\mathbf{0}$, a row vector of 0s (size $n-1$) and block $(1, 1)$ is the number 0.

A basic property of matrix exponentials then says that

$$e^{A^*} = \left(\begin{array}{c|c} e^{\widetilde{A}^*} & \mathbf{1}^T - e^{\widetilde{A}^*}\mathbf{1}^T \\ \hline \mathbf{0} & e^0 = 1 \end{array} \right),$$

so, in the same way,

$$e^{A^*t} = \left(\begin{array}{c|c} e^{\widetilde{A}^*t} & \mathbf{1}^T - e^{\widetilde{A}^*t}\mathbf{1}^T \\ \hline \mathbf{0} & e^0 = 1 \end{array} \right).$$

□

Consider the previous numerical example $A = \begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix}$, leading to

$$A^* = \begin{pmatrix} -1 & 0 & 1 \\ -2 & -2 & 4 \\ 0 & 0 & 0 \end{pmatrix}. \text{ We have:}$$

$$e^{A^*t} = \begin{pmatrix} e^{-t} & 0 & 1 - e^{-t} \\ -2e^{-t} + 2e^{-2t} & e^{-2t} & 1 + 2e^{-t} - 3e^{-2t} \\ 0 & 0 & 1 \end{pmatrix}.$$

5.3. Inversion lemma for the exponential-dual. Knowing the exponential-dual A^* of A , we can recover A using the following result:

LEMMA 5.4 (Inversion lemma for the exponential-dual matrix). *For $0 \leq i, j \leq n-1$, we have*

$$A_{i,j} = \sum_{k=0}^i (A_{j,k}^* - A_{j+1,k}^*).$$

PROOF.

Re-write the definition of A^* using the following notation: for $0 \leq j, k \leq n$,

$$A_{j,k}^* = \sum_{\ell=j}^{n-1} (A_{k,\ell} - A_{k-1,\ell}).$$

Summing the first $i+1$ elements of row j of A^* gives

$$\begin{aligned} \sum_{k=0}^i A_{j,k}^* &= \sum_{k=0}^i \sum_{\ell=j}^{n-1} (A_{k,\ell} - A_{k-1,\ell}) \\ &= \sum_{\ell=j}^{n-1} \sum_{k=0}^i (A_{k,\ell} - A_{k-1,\ell}) \\ &= \sum_{\ell=j}^{n-1} [A_{0,\ell} + A_{1,\ell} + \cdots + A_{i,\ell} - (A_{-1,\ell} + A_{0,\ell} + \cdots + A_{i-1,\ell})] \\ &= \sum_{\ell=j}^{n-1} A_{i,\ell}. \end{aligned}$$

Writing the obtained equality $\sum_{k=0}^i A_{j,k}^* = \sum_{\ell=j}^{n-1} A_{i,\ell}$ again but replacing j by $j+1$ produces $\sum_{k=0}^i A_{j+1,k}^* = \sum_{\ell=j+1}^{n-1} A_{i,\ell}$. Subtracting now both equalities gives

$$\sum_{k=0}^i A_{j,k}^* - \sum_{k=0}^i A_{j+1,k}^* = \sum_{\ell=j}^{n-1} A_{i,\ell} - \sum_{\ell=j+1}^{n-1} A_{i,\ell},$$

that is,

$$\sum_{k=0}^i (A_{j,k}^* - A_{j+1,k}^*) = A_{i,j}. \quad \square$$

5.4. Main result for the exponential-dual. Our main result is the following one:

THEOREM 5.5. *Define matrix function F using the following relation: for any i and j belonging to the index set $\{0, 1, \dots, n-1\}$,*

$$F_{i,j}(t) = \sum_{k=0}^i (E_{j,k}^*(t) - E_{j+1,k}^*(t)).$$

Then, $F(t) = e^{At}$.

Before proving this main theorem, we need the following lemma.

LEMMA 5.6 (Inversion lemma for matrix function E^*). *Knowing F , we can recover matrix E^* using the following relations: for $0 \leq i \leq n$ and $1 \leq j \leq n-1$, we have (we are omitting ‘ t ’ here, for more clarity in the text)*

$$E_{i,j}^* = \sum_{k=i}^{n-1} (F_{j,k} - F_{j-1,k});$$

for the last column of E^* , we have, for any $i \in \{0, 1, \dots, n\}$,

$$E_{i,n}^* = 1 - \sum_{k=i}^{n-1} F_{n-1,k}.$$

PROOF. Let us re-write the definition of F with the following notation:

$$F_{j,k} = \sum_{\ell=0}^j (E_{k,\ell}^* - E_{k+1,\ell}^*),$$

Summing the last components of row j (which is $\leq n-1$) of F , starting at column i , gives

$$\begin{aligned} \sum_{k=i}^{n-1} F_{j,k} &= \sum_{k=i}^{n-1} \sum_{\ell=0}^j (E_{k,\ell}^* - E_{k+1,\ell}^*) \\ &= \sum_{\ell=0}^j \sum_{k=i}^{n-1} (E_{k,\ell}^* - E_{k+1,\ell}^*) \\ &= \sum_{\ell=0}^j [E_{i,\ell}^* + E_{i+1,\ell}^* + \dots + E_{n-1,\ell}^* - (E_{i+1,\ell}^* + E_{i+2,\ell}^* + \dots + E_{n,\ell}^*)] \\ &= \sum_{\ell=0}^j E_{i,\ell}^*. \end{aligned}$$

We used the fact that, since in the sums on ℓ , we always have $\ell < n$, then $E_{n,\ell}^* = 0$.

Writing now the obtained equality $\sum_{k=i}^{n-1} F_{j,k} = \sum_{\ell=0}^j E_{i,\ell}^*$ and the equality obtained by changing j by $j-1$, that is, $\sum_{k=i}^{n-1} F_{j-1,k} = \sum_{\ell=0}^{j-1} E_{i,\ell}^*$ and subtracting them, we get

$$\sum_{k=i}^{n-1} (F_{j,k} - F_{j-1,k}) = E_{i,j}^*.$$

For the case of $j = n-1$, we start from the same expression (replacing j by $n-1$):

$$F_{n-1,k} = \sum_{\ell=0}^{n-1} (E_{k,\ell}^* - E_{k+1,\ell}^*),$$

We sum on k from i to $n-1$:

$$\begin{aligned} \sum_{k=i}^{n-1} F_{n-1,k} &= \sum_{k=i}^{n-1} \sum_{\ell=0}^{n-1} (E_{k,\ell}^* - E_{k+1,\ell}^*) \\ &= \sum_{\ell=0}^{n-1} \sum_{k=i}^{n-1} (E_{k,\ell}^* - E_{k+1,\ell}^*) \\ &= \sum_{\ell=0}^{n-1} [E_{i,\ell}^* + E_{i+1,\ell}^* + \cdots + E_{n-1,\ell}^* - (E_{i+1,\ell}^* + E_{i+2,\ell}^* + \cdots + E_{n,\ell}^*)] \\ &= \sum_{\ell=0}^{n-1} E_{i,\ell}^* \\ &= 1 - E_{i,n}^*, \end{aligned}$$

leading to

$$E_{i,n}^* = 1 - \sum_{k=i}^{n-1} F_{n-1,k}. \quad \square$$

PROOF OF THE MAIN THEOREM. Starting from some given matrix A , we construct A^* , compute $E^*(t) = e^{A^*t}$, and construct F . We must prove that $F(t) = e^{At}$, or equivalently, that $F' = FA$, or that $F' = AF$.

Let us use the notation $G(t) = E^*(t)$ to avoid all these $*$, and omit ' t ' as before for the sake of clarity.

We will prove that

$$(5.1) \quad F' = AF.$$

Fix i, j with $0 \leq i, j \leq n-1$, and focus on the right hand side of (5.1).

$$(AF)_{i,j} = \sum_{u=0}^{n-1} A_{i,u} F_{u,j}$$

$$(5.2) \quad = \sum_{u=0}^{n-1} A_{i,u} \sum_{v=0}^u (G_{j,v} - G_{j+1,v}).$$

Now, we look at the l.h.s. From the definition of F ,

$$(5.3) \quad F'_{i,j} = \sum_{k=0}^i (G'_{j,k} - G'_{j+1,k}).$$

We know that $G' = GA^* = A^*G$. Let us use here the second equality.

$$\begin{aligned} G'_{j,k} &= \sum_{v=0}^n G_{j,v} A_{v,k}^* \\ &= \sum_{v=0}^{n-1} G_{j,v} A_{v,k}^* \\ &\quad \swarrow \text{because the last row of } A^* \text{ is only composed of 0s} \\ &= \sum_{v=0}^{n-1} G_{j,v} \sum_{u=v}^{n-1} (A_{k,u} - A_{k-1,u}). \end{aligned}$$

Coming back to $F'_{i,j}$, we have

$$\begin{aligned} F'_{i,j} &= \sum_{k=0}^i (G'_{j,k} - G'_{j+1,k}) \\ &= \sum_{k=0}^i \sum_{v=0}^{n-1} (G_{j,v} - G_{j+1,v}) \sum_{u=v}^{n-1} (A_{k,u} - A_{k-1,u}) \\ &= \sum_{v=0}^{n-1} G_{j,v} \sum_{u=v}^{n-1} \sum_{k=0}^i (A_{k,u} - A_{k-1,u}) - \sum_{v=0}^{n-1} G_{j+1,v} \sum_{u=v}^{n-1} \sum_{k=0}^i (A_{k,u} - A_{k-1,u}) \\ &= \sum_{v=0}^{n-1} (G_{j,v} - G_{j+1,v}) \sum_{u=v}^{n-1} \sum_{k=0}^i (A_{k,u} - A_{k-1,u}) \\ &\quad \swarrow \text{after moving the first sum on } k \text{ to the end} \\ &= \sum_{v=0}^{n-1} (G_{j,v} - G_{j+1,v}) \sum_{u=v}^{n-1} A_{i,u} \\ &\quad \swarrow \text{after observing that } \sum_{k=0}^i (A_{k,u} - A_{k-1,u}) \text{ is a telescopic sum} \\ &= \sum_{u=0}^{n-1} A_{i,u} \sum_{v=0}^u (G_{j,v} - G_{j+1,v}) \\ &\quad \swarrow \text{after interchanging the summation order,} \end{aligned}$$

which is exactly (5.2). \square

Observe now that we have proved the equivalent of relations (3.1) and (3.2) but in our generalization, where A is an arbitrary square matrix: using the notation E and E^* , we have

$$E_{i,j}^*(t) = \sum_{k=0}^{i-1} [E_{j-1,k}(t) - E_{j,k}(t)],$$

$$E_{i,j}(t) = \sum_{k=0}^i [E_{j,k}^*(t) - E_{j+1,k}^*(t)].$$

This is equivalent to the Duality Theorem mentioned in [8], here a direct consequence of Theorem 5.5.

Example: let us look at the generic 2-state case. If

$$E^* = \begin{pmatrix} E_{0,0}^* & E_{0,1}^* & 1 - E_{0,0}^* - E_{0,1}^* \\ E_{1,0}^* & E_{1,1}^* & 1 - E_{1,0}^* - E_{1,1}^* \\ 0 & 0 & 1 \end{pmatrix},$$

then we have:

$$P = \begin{pmatrix} E_{0,0}^* - E_{1,0}^* & E_{1,0}^* \\ E_{0,0}^* + E_{0,1}^* - (E_{1,0}^* + E_{1,1}^*) & E_{1,0}^* + E_{1,1}^* \end{pmatrix}.$$

Reconsider example $A = \begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix}$, whose exponential-dual was $A^* =$

$$\begin{pmatrix} -1 & 0 & 1 \\ -2 & -2 & 4 \\ 0 & 0 & 0 \end{pmatrix}. \text{ We had}$$

$$E^*(t) = \begin{pmatrix} e^{-t} & 0 & 1 - e^{-t} \\ -2e^{-t} + 2e^{-2t} & e^{-2t} & 1 + 2e^{-t} - 3e^{-2t} \\ 0 & 0 & 1 \end{pmatrix}.$$

Using the inversion formulas and Theorem 5.5, we obtain

$$F(t) = \begin{pmatrix} 3e^{-t} - 2e^{-2t} & -2e^{-t} + 2e^{-2t} \\ 3e^{-t} - 3e^{-2t} & -2e^{-t} + 3e^{-2t} \end{pmatrix} = e^{At}.$$

6. Conclusions

In Chapter 3, we first describe stochastic duality (or simply duality, and meaning in the sense of [16] or [1]), and use it as a tool to find transient distributions of basic Markovian queuing models when combined with uniformization. This approach makes sense when the analysis of the dual is simpler than that of the initial model, and this has been the case in several articles [9], [10] dealing with fundamental queuing systems. However, there is a limitation to this approach, which is the fact that some Markovian models have no dual, or that they come with restrictions.

After discussing these drawbacks of dual usage, we define an algebraically similar concept but without any reference to stochastic processes. Because of its similarity to the dual and its role in computing matrix exponentials, we call it the exponential-dual, and we show that it coincides with the dual when we are in a Markov setting and the dual exists. The advantage of the exponential-dual is that it exists for any given matrix.

Future work will explore this concept in more depth by separating the discrete and continuous “time” cases, and the connections between them. Another direction that deserves attention is the exploration of the relations with other basic matrix transformations, and more generally, with spectral analysis.

Bibliography

1. W.J. Anderson, *Continuous-time Markov chains: an applications-oriented approach*, Springer, New York, 1991.
2. E. C. Bhattacharya, R.N. Waymire, *Stochastic processes with applications*, Siam, 2009.
3. M. L. Green, A. Krinik, C. Mortensen, G. Rubino, and R. J. Swift, *Transient probability functions: A sample path approach*, Discrete Mathematics and Theoretical Computer Science (2003), 127–136.
4. J. L. Jain, S. G. Mohanty, and W. Böhm, *A course on queueing models*, Chapman & Hall/CRC, 2007.
5. S. Jansen and N. Kurt, *On the notion(s) of duality for Markov processes*, Probability Surveys **11** (2014), no. 1, 59–120.
6. A. Jensen, *Markoff chains as an aid in the study of Markoff processes*, Scandinavian Actuarial Journal **1953** (1953), 87–91.
7. A. Krinik et al., *Explicit Transient Probabilities of Various Markov Models*, in Stochastic Processes and Functional Analysis, edited by Randall Swift, Alan Krinik, Jennifer Switkes and Jason Park, to appear, AMS, 2021.
8. A. Krinik and S. G. Mohanty, *On batch queueing systems: A combinatorial approach*, Journal of Statistical Planning and Inference **140** (2010), 2271–2284.
9. A. Krinik, C. Mortensen, and G. Rubino, *Connections between birth-death processes*, Stochastic Processes and Functional Analysis (Alan C. Krinik and Randall J. Swift, eds.), Marcel Dekker, 2004, pp. 219–240.
10. A. Krinik, G. Rubino, D. Marcus, R. J. Swift, H. Kasfy, and H. Lam, *Dual processes to solve single server systems*, Journal of Statistical Planning and Inference **135** (2005), no. 1, 702–713.
11. S. Lyche, *On Deep Learning and Neural Networks*, Master’s thesis, California State Polytechnic University, Pomona, California, US, 2018.
12. J. Medhi, *Stochastic models in queueing theory, 2nd edition*, Academic Press, 2003.
13. A. G. Pakes, *Convergence Rates and Limit Theorems for the Dual Markov Branching Process*, Journal of Probability and Statistics **17** (2017), published online.
14. L. Pawel, *Generalized gambler’s ruin problem: explicit formulas via Siegmund duality*, Methodology and Computing in Applied Probability **19** (2017), no. 2, 603–613.
15. G. Rubino and B. Sericola, *Markov Chains and Dependability Theory*, Cambridge University Press, 2014.
16. D. Siegmund, *The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes*, Ann. Prob. **6** (1976), 914–924.
17. R. J. Swift, *A simple immigration-catastrophe process*, Math. Sci. **25** (2000), no. 1, 32–36.
18. P. Zhao, *Siegmund duality for continuous time Markov chains on Z_+^d* , Acta Mathematica Sinica, English Series **17** (2018), published online.