



**HAL**  
open science

# Two algebraic matrix transformations with applications to Markov chain analysis

Gerardo Rubino

► **To cite this version:**

Gerardo Rubino. Two algebraic matrix transformations with applications to Markov chain analysis. JMM 2024 - Joint Mathematics Meetings, AMS - American Mathematical Society, Jan 2024, San Francisco (CA), United States. hal-04756101

**HAL Id: hal-04756101**

**<https://inria.hal.science/hal-04756101v1>**

Submitted on 28 Oct 2024

**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.

Public Domain

# Two algebraic matrix transformations with applications to Markov chain analysis

**G. Rubino**, Inria, Rennes, France  
(joint work with A. Krinik, CalPoly, Pomona, US)

JMM'24

January 5, 2024





# Starting with Siegmund duality

- The Siegmund dual of a Markov process  $X$  is a probabilistic transformation that **exists if  $X$  satisfies specific monotonicity conditions**.
- If we write

$$X \longrightarrow \text{SiegmundDual}(X),$$

this can be also seen as a mapping between matrices, that is, transition probability matrices (t.p.m.s) or infinitesimal generators (i.g.s), that we can similarly write, for a t.p.m.  $P$ ,

$$P \longrightarrow \text{SiegmundDual}(P) = \tilde{P},$$

or for an i.g.  $A$

$$A \longrightarrow \text{SiegmundDual}(A) = \hat{A}.$$

The difference in the notation is because **the matrix transformation is slightly different in case of p.t.m.s and of i.g.s.**

# Interest

- In some cases (for instance, for many fundamental queueing models starting with the  $M/M/1$  and  $M/M/1/H$  ones), the use of Siegmund duality allows for a significantly easier analysis of the transient behavior of a given Markov process, and specifically to obtain closed-form expressions.
- The main problem to face here is the conditions under which such a dual process exists.
- When looking for closed-forms, the original model must have a small number of parameters, otherwise the combinatorics of the problem can make it too hard.

## 2 ideas

- First idea in this happens: what happens if we define the same mappings for **arbitrary** square matrices now? In other words, if we ignore any restriction or condition.
- Second idea: looking at things this way forces to abandon probabilistic arguments and to **move to an algebraic point of view**. This appears to be fruitful.
- We call *power-dual* of an *arbitrary* square matrix the mapping corresponding formally to Siegmund-duality in DTMCs, and *exponential-dual* the one formally corresponding to CTMCs.
- Basically, the **power-dual** appears when we are interested in **matrix powers** and the **exponential-dual** appears if we look at **matrix exponentials**.





# The power-dual of a matrix

- Let  $M$  be an arbitrary square matrix of reals (or of complex numbers, etc.) with dimension  $n$ , indexed on  $\{0, 1, \dots, n-1\}$ .
- The power-dual of  $M$  is a new matrix  $\tilde{M}$ , also denoted here  $\mathcal{PD}(M)$ , with dimension  $n+1$ , indexed on  $\{0, 1, \dots, n-1, n\}$ , defined as follows: for  $i = 0, 1, \dots, n-1$ ,

$$\tilde{M}_{i,j} = \begin{cases} \sum_{k=i}^{n-1} M_{j,k} - \sum_{k=i}^{n-1} M_{j-1,k} & \text{if } j \leq n-1, \\ 1 - \sum_{k=0}^{n-1} \tilde{M}_{i,k} & \text{when } j = n, \end{cases}$$

and for the last row ( $i = n$ ), we have  $(0 \ 0 \ \dots \ 0 \ 1)$ .

- Example, dimension 2:

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightsquigarrow \tilde{M} = \begin{pmatrix} a+b & c+d-(a+b) & 1-(c+d) \\ b & d-b & 1-d \\ 0 & 0 & 1 \end{pmatrix}$$

- Example, dimension 3:

$$M = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & -2 \\ 1 & -3 & 3 \end{pmatrix} \rightsquigarrow \tilde{M} = \begin{pmatrix} 2 & -2 & 1 & 0 \\ 1 & -3 & 2 & 1 \\ 2 & -4 & 5 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- A stochastic matrix:

$$P = \begin{pmatrix} 2/5 & 0 & 3/5 \\ 1/6 & 1/2 & 1/3 \\ 0 & 1/4 & 3/4 \end{pmatrix} \rightsquigarrow \tilde{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3/5 & 7/30 & 1/6 & 0 \\ 3/5 & \boxed{-4/15} & 5/12 & 1/4 \\ 0 & 0 & 0 & 1 \end{pmatrix};$$

as we see, in this case the power-dual is not a stochastic matrix.

- Another stochastic matrix:

$$P = \begin{pmatrix} 2/5 & 3/5 \\ 1/6 & 5/6 \end{pmatrix} \rightsquigarrow \tilde{P} = \begin{pmatrix} 1 & 0 & 0 \\ 3/5 & 1/15 & 1/3 \\ 0 & 0 & 1 \end{pmatrix};$$

here, the power-dual is also a stochastic matrix.

## Some properties

- If  $\tilde{M}$  is the power-dual of matrix  $M$ , then the sum of the elements of any row of  $\hat{M}$  is 1, and the last row of  $\hat{M}$  is composed only of 0s except for the last element which is 1.
- Inversion Lemma: for  $0 \leq i, j \leq n-1$ , we have

$$M_{i,j} = \sum_{k=0}^i \tilde{M}_{j,k} - \sum_{k=0}^i \tilde{M}_{j+1,k}.$$

- Central result: for any two square matrices  $A$  and  $B$  with the same dimension, we have

$$\tilde{AB} = \tilde{B}\tilde{A}.$$

Corollary (and Main result): for any square matrix  $M$  and any nonnegative integer  $k$ ,

$$\boxed{\tilde{M}^k = \tilde{M}^k}, \text{ written also } \boxed{\mathcal{PD}(M^k) = \mathcal{PD}(M)^k},$$

the power-dual of the  $k$ th power of  $M$  is the  $k$ th power of  $M$ 's power-dual.



# The exponential-dual of a matrix

- Let  $M$  be a square matrix (of reals, of complex numbers, ...) with dimension  $n$ , indexed on  $\{0, 1, \dots, n-1\}$ . The exponential-dual of  $M$  is a new matrix  $\widehat{M}$ , also denoted  $\mathcal{ED}(M)$  in these slides, having dimension  $n+1$ , indexed on  $\{0, 1, \dots, n-1, n\}$ , defined as follows: for  $i < n$ ,

$$\widehat{M}_{i,j} = \sum_{k=i}^{n-1} M_{j,k} - \sum_{k=i}^{n-1} M_{j-1,k};$$

the last row of  $\widehat{M}$  is  $(0 \ 0 \ \dots \ 0)$ .

- Example, dimension 2:

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightsquigarrow \widehat{M} = \begin{pmatrix} a+b & c+d-(a+b) & -(c+d) \\ b & d-b & -d \\ 0 & 0 & 0 \end{pmatrix}$$

- Example, dimension 3:

$$M = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 2 & -3 \\ 0 & 2 & -2 \end{pmatrix} \rightsquigarrow \tilde{M} = \begin{pmatrix} 6 & -8 & 2 & 0 \\ 5 & -6 & 1 & 0 \\ 3 & -6 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- An infinitesimal generator:

$$M = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} \rightsquigarrow \tilde{M} = \begin{pmatrix} 0 & 0 & 0 \\ \lambda & -(\lambda + \mu) & \mu \\ 0 & 0 & 0 \end{pmatrix};$$

as we see, the exponential-dual is also an infinitesimal generator.

- Another infinitesimal generator:

$$M = \begin{pmatrix} -3 & 1 & 2 \\ 6 & -7 & 1 \\ 1 & 2 & -3 \end{pmatrix} \rightsquigarrow \tilde{M} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & -9 & 5 & 1 \\ 2 & \boxed{-1} & -4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix};$$

here, the exponential-dual is not an infinitesimal generator.

## Some properties

- If  $\widehat{M}$  is the exponential-dual of matrix  $M$ , then the sum of the elements of any row of  $\widehat{M}$  is 0, and the last row of  $\widehat{M}$  is composed only of 0s.
- Inversion Lemma:  $0 \leq i, j \leq n - 1$ , we have

$$M_{i,j} = \sum_{k=0}^i \widehat{M}_{j,k} - \sum_{k=0}^i \widehat{M}_{j+1,k}.$$

- Important result: for any two square matrices  $A$  and  $B$  with the same dimension, we have

$$\widehat{AB} = \widehat{B} \widehat{A}.$$

Corollary: for any square matrix  $M$  and any integer  $k \geq 1$ ,

$$\widehat{M}^k = \widehat{M}^k.$$

In words, the exponential-dual of the  $k$ th power of  $M$  is the  $k$ th power of  $M$ 's exponential-dual.

- Let  $A$  and  $B$  be two square matrices with the same dimension and  $\alpha$  a scalar. We have

$$\widehat{A+B} = \widehat{A} + \widehat{B}, \quad \widehat{\alpha A} = \alpha \widehat{A}.$$

- Let  $(M_k)_{k \geq 0}$  a matrix sequence, with  $M_k \rightarrow M$  as  $k \rightarrow \infty$ . Then,

$$\widehat{M}_k \rightarrow \widehat{M} \text{ as } k \rightarrow \infty.$$

If  $\sum_{k \geq 0} M_k = M$ , then

$$\widehat{\sum_{k \geq 0} M_k} = \sum_{k \geq 0} \widehat{M}_k = \widehat{M}.$$



# Main result

- Let  $M$  be an arbitrary square matrix. Using previous properties, we obtain

$$\boxed{e^{\widehat{M}t} = \widetilde{e^{Mt}}} \text{ also written } \boxed{\mathcal{PD}(e^{Mt}) = e^{\mathcal{ED}(M)t}}.$$

In words, the **power**-dual of the exponential of  $Mt$  is the exponential of the exponential-dual of  $M$  times  $t$ .

- This result explains how the evaluation of the power-dual appears when we work on matrix exponentials.
- In this short presentation we skipped some technical results that are needed to prove previous theorem.

# A path to compute $e^{Mt}$

Instead of

$$M \xrightarrow{\text{any algorithm}} e^{Mt},$$

we can try

$$M \xrightarrow[\text{(linear)}]{\uparrow} \hat{M} \xrightarrow[\text{any algorithm}]{\uparrow} e^{\hat{M}t} = \widetilde{e^{Mt}} \xrightarrow[\text{Inversion Lemma of power-dual (linear)}]{\uparrow} e^{Mt}$$

## Other properties of both transformations

- If  $P$  is a stochastic matrix, then the first row of  $\tilde{P}$  is  $(1\ 0\ 0\ \dots\ 0)$ . Recall that  $\tilde{P}$  doesn't need to be stochastic.
- If  $A$  is an infinitesimal generator, then the first row of  $\hat{A}$  is  $(0\ 0\ \dots\ 0)$ . Recall that  $\hat{A}$  doesn't need to be an infinitesimal generator.
- Connection with the dual of a Markov process as defined by Siegmund in 1976 and developed in Anderson's book: if  $A$  is the infinitesimal generator of a Markov process (CTMC)  $X$  whose Siegmund dual exists, its generator is  $\hat{A}$ .

The details for the theory corresponding to the exponential-dual transformation have been published in "*The exponential-dual matrix method: Applications to Markov chain analysis*", G. Rubino and A. Krinik, AMS Contemporary Mathematics Series, Vol. 774, pages 217–235, 2021.

The corresponding paper for the power-dual and the relations between both transformations is in preparation.



# Useful notation

Let  $M$  be a matrix having dimension  $n \geq 2$  and let us define two auxiliary matrices  $L$  and  $R$ , also of dimension  $n$ , as follows:

- $L_{i,j} = 1(i \leq j)$ , where  $1(S)$  is the indicator function of statement  $S$  (it is equal to 1 if  $S$  is true, and to 0 if it is false);
- concerning matrix  $R$ , we have, for  $0 \leq i \leq n-1$ ,  $R_{i,i} = 1$ , then, for  $0 \leq i \leq n-2$ ,  $R_{i,i+1} = -1$ , and for the rest of the matrix the elements are all equal to 0.

For instance, in dimension 4,

$$L = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

- Matrices  $L$  and  $R$  are invertible, and when they have the same dimension,  $RL = I$  ( $R$  is the inverse of  $L$  and vice versa).
- Given a matrix  $U$ , we denote by  $\text{nw}(U)$  (and we say the “northwest” of  $U$ ) the matrix obtained from  $U$  by removing its last row and column. “Graphically”,

$$U = \left( \begin{array}{c|c} \text{nw}(U) & \vdots \\ \hline \dots & \end{array} \right),$$

For instance, if we say  $L_h$  or  $R_h$  for previous matrices when their dimension is  $h$ , we have  $\text{nw}(L_{n+1}) = L_n$  and  $\text{nw}(R_{m+1}) = R_m$ .

# Similarity property

- Matrix version of the power-dual definition:  $M^T$  and  $\text{nw}(\tilde{M})$  verify

$$\text{nw}(\tilde{M}) = L M^T R,$$

meaning that they are **similar matrices**.

- For instance, as a corollary, we have that the spectrums of  $M$  and of  $\text{nw}(\tilde{M})$  are identical.
- We can see here, in matrix form, the inverse transformation:  
 $M = L^T \text{nw}(\tilde{M}^T) R^T$  (we used  $(\text{nw}(M))^T = \text{nw}(M^T)$ ).

# An application of the power-dual: the gambler ruin problem

- Let  $\tilde{X}$  be a Markov chain with  $n + 1$  states labelled  $0, 1, \dots, n$ , where states  $0$  and  $n$  are absorbing and states  $1$  to  $n - 1$  are transient.
- Let  $\tilde{P}$  be the transition probability matrix of  $\tilde{X}$ . Matrix  $\tilde{P}$  has the following structure:

$$\tilde{P} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ \tilde{P}_{1,0} & \boxed{\tilde{P}'} & & & \tilde{P}_{1,n} \\ \vdots & & & & \vdots \\ \tilde{P}_{n-1,0} & & & & \tilde{P}_{n-1,n} \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

- We want to compute the absorption probabilities  $x_i = \mathbb{P}_i(\tilde{X}(\infty) = 0)$ , for all state  $i$ .
- Observe that since all states  $\neq 0, n$  are transient, then for all state  $i$ ,  $\mathbb{P}_i(\tilde{X}(\infty) = n) = 1 - x_i$ .



- Observation: computing the  $x_i$ s is a simple linear problem:
- We have that  $x_0 = 1$ ,  $x_n = 0$ , and for all  $i \neq 0, n$ ,

$$x_i = \tilde{P}_{i,0} + \sum_{j \neq 0, n} \tilde{P}_{i,j} x_j.$$

- The fact that all states  $\neq 0, n$  are transient makes that this system has a unique solution.
- Denote  $y = (x_1 \dots x_{n-1})$  and  $p = (\tilde{P}_{1,0} \dots \tilde{P}_{n-1,0})$ .
  - In matrix form, the system to solve is  $y^T = p^T + \tilde{P}y^T$ .
  - The solution is  $y^T = (I - \tilde{P})^{-1}p^T$ .

# Absorption probabilities using the power-dual

- Let us put a name on  $\text{nw}(\tilde{P})$ , say  $\bar{P}$ , that is,

$$\tilde{P} = \left( \begin{array}{c|c} & \\ \hline & \bar{P} \\ \hline \dots & \vdots \end{array} \right).$$

- Call  $P$  the **unique** matrix whose power-dual is  $\tilde{P}$ .  
Property:  $P$  is **somewhat stochastic** (that is,  $P1^T = 1^T$ , but  $P$  is not necessarily stochastic).

- **Result:** there exists a vector  $\pi$  with dimension  $n$  such that matrix  $P^\ell \rightarrow \mathbf{1}^T \pi$  as  $\ell \rightarrow \infty$ , where  $\pi \mathbf{1}^T = 1$ .

**Sketch of proof:** we know that  $\bar{P} = L P^T R$ , so,  $P^T = R \bar{P} L$ , which implies that for any integer  $\ell \geq 1$ ,

$$(P^T)^\ell = R \bar{P}^\ell L.$$

Now, by construction of  $\tilde{P}$ , we have that

$$\bar{P}^\infty = \left( \begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ x_1 & & & \\ x_2 & & & \\ \vdots & & & \\ x_{n-1} & & & 0 \end{array} \right),$$

where  $x_i = \mathbb{P}_i(\tilde{X}(\infty) = 0)$ , for all transient states  $i$ .

Then, we check that  $R \bar{P}^\infty L$  can be written  $(\mathbf{1}^T \pi)^\top$ , where for  $i = 0, 1, \dots, n-1$ , we have  $\pi_i = x_i - x_{i+1}$  if we extend the  $x_i$ s with  $x_0 = 1$  and  $x_n = 0$ .

How to obtain vector  $\pi$ ? (observe that we have  $\pi 1^T = 1$  but that some components of  $\pi$  can be negative).

- **Result:** the system  $\{zP = z \text{ and } z1^T = 1\}$  has a unique solution, and the solution is  $\pi$ .

**Sketch of proof:** we first prove that  $\pi P = \pi$ , for instance, by writing the sequence  $(p(k))_{k \geq 0}$  defined by  $p(k) = \alpha P^k$  where  $\alpha$  satisfies  $\alpha 1^T = 1$  but is otherwise arbitrary, then making  $k \rightarrow \infty$  in  $p(k+1) = p(k)P$ .

Then, if  $v$  is a vector such that  $v = vP$  and  $v1^T = 1$ , again, after observing that  $vP^k = v$  and making  $k \rightarrow \infty$ , we get  $vP^\infty = v$ , but  $vP^\infty = v1^T \pi = \pi$  (associativity), so,  $v = \pi$ .

- So, once  $\pi$  known (by solving the linear system  $\pi P = \pi$ ), we obtain the  $x_i$ s inverting the relations we have from  $x$  to  $\pi$ :
  - $x_1 = 1 - \pi_0$ ,
  - $x_2 = 1 - \pi_0 - \pi_1$ ,
  - ...
  - $x_i = 1 - (\pi_0 + \pi_1 + \dots + \pi_{i-1})$ , or  $x_i = \pi_i + \pi_{i+1} + \dots + \pi_{n-1}$ ,
  - ...
  - $x_{n-1} = \pi_{n-1}$ .

# Resuming

- To obtain vector  $\pi$  we solve the linear system  $\pi P = \pi$ , as we do when  $P$  is, for instance, the t.p.m. of an irreducible chain, even if here, we can have negative elements in  $P$  (and in the solution  $\pi$ ).
- Then, we apply previous formulas. These can be simply written

$$x = \pi L \quad (\text{and } \pi = x R).$$



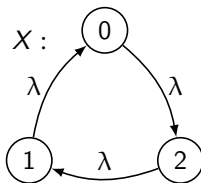
- The power-dual and the exponential-dual extend somehow Sigmund duality to arbitrary matrices, so, at the same time, they move the transformation from Markov processes to general linear algebra.
- Sometimes, the transformation simplifies significantly the analysis, that is, working with these duals is, in some cases, much easier than with the original matrices.
- The main task that remains to be explored (work in progress) is the case of infinite state spaces / infinite matrices (linear operators).



**Thanks for your time.**

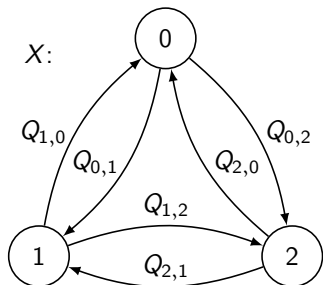
# On Siegmund dual's existence (Markov processes)

- This concerns the historical origin of this work, in the stochastic processes area.
- For instance, consider an homogeneous “ring” illustrated here for 3 states:

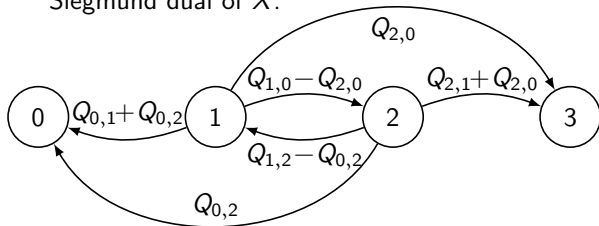


- Whatever the size of the ring and the value of  $\lambda$ , the associated Markov process has no Siegmund dual (for those familiar with duality in Markov processes, the problem is the monotonicity condition that doesn't hold here).

Consider now the general Markov process with 3 states:



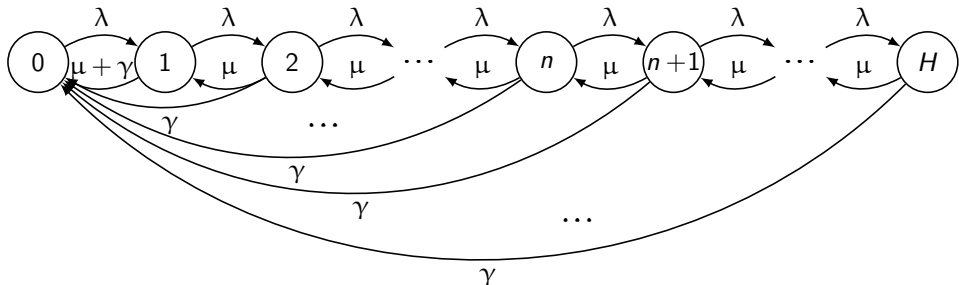
Siegmund dual of  $X$ :



Here, the Siegmund dual exists iff  $Q_{1,0} > Q_{2,0}$  and  $Q_{1,2} > Q_{0,2}$ .

# An example where duality helps

Consider here the  $M/M/1/H$  queue plus “catastrophes” (or failures) as shown in the following Markovian graph:



- Take  $H = 4$ , to shorten the slides. Denote  $\Lambda = \lambda + \mu + \gamma$ . We have

$$A = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 \\ \mu + \gamma & -\Lambda & \lambda & 0 & 0 \\ \gamma & \mu & -\Lambda & \lambda & 0 \\ \gamma & 0 & \mu & -\Lambda & \lambda \\ \gamma & 0 & 0 & \mu & -(\mu + \gamma) \end{pmatrix}$$

- If we move to discrete time using Uniformization, the computation of  $e^{At}$  translates into that of evaluating  $U^k$  where using the notation  $p = \lambda/\Lambda$ ,  $q = \mu/\Lambda$ ,  $r = \gamma/\Lambda$ , matrix  $U = I - A/\Lambda$  is

$$U = \begin{pmatrix} q+r & p & 0 & 0 & 0 \\ q+r & 0 & p & 0 & 0 \\ r & q & 0 & p & 0 \\ r & 0 & q & 0 & p \\ r & 0 & 0 & q & p \end{pmatrix}$$

The powers of  $U$  are hard to obtain in closed-form here.

- If we use the exponential-dual (plus Uniformization), we must compute the powers of

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ p & 0 & q & 0 & 0 & r \\ 0 & p & 0 & q & 0 & r \\ 0 & 0 & p & 0 & q & r \\ 0 & 0 & 0 & p & q+r & r \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Due to those '1' at the corners, the use of lattice path combinatorics is much easier here than with matrix  $U$  of previous slide. This is an example of situation where we can develop symbolically the computations until the end.

- The infinite case can also be solved in the same way.

The formal solution of the ODEs (Chapman-Kolmogorov equations) is given below, using the notation  $\rho = \lambda/\mu$ , for the case of the queue starting empty at time 0 (the general case is similar while more complex):

$$p_j(t) = \pi_j - \frac{2\mu\rho^{j/2}}{H+1} \sum_{1 \leq h \leq H} \left\{ \frac{e^{-(\lambda + \mu + \gamma - 2\sqrt{\lambda\mu} \cos(\frac{h\pi}{H+1}))t}}{\lambda + \mu + \gamma - 2\sqrt{\lambda\mu} \cos(\frac{h\pi}{H+1})} \times \right. \\ \left. \times \sqrt{\rho} \sin\left(\frac{h\pi}{H+1}\right) \left[ \sin\left(\frac{j h \pi}{H+1}\right) - \sqrt{\rho} \sin\left(\frac{(j+1)h\pi}{H+1}\right) \right] \right\}.$$

In the formula,  $\pi_j$  denotes  $p_j(\infty)$  (steady-state), easy to evaluate symbolically.