Queues, Stores, and Tableaux

Moez Draief * Jean Mairesse † Neil O'Connell ‡

May 10, 2005

Abstract

Consider the single server queue with an infinite buffer and a FIFO discipline, either of type $M/M/1$ or Geom/Geom/1. Denote by $\mathcal{A}$ the arrival process and by s the services. Assume the stability condition to be satisfied. Denote by $\mathcal{D}$ the departure process in equilibrium and by r the time spent by the customers at the very back of the queue. We prove that $(\mathcal{D}, r)$ has the same law as $(\mathcal{A}, s)$ which is an extension of the classical Burke Theorem. In fact, r can be viewed as the departures from a dual *storage* model. This duality between the two models also appears when studying the transient behavior of a tandem by means of the RSK algorithm: the first and last row of the resulting semi-standard Young tableau are respectively the last instant of departure in the queue and the total number of departures in the store.

Keywords: Single server queue, storage model, Burke's theorem, non-colliding random walks, tandem of queues, Robinson-Schensted-Knuth algorithm

1 Introduction

The main purpose of this paper is to clarify the interplay between two models of queueing theory. The first model is the very classical single server queue with an infinite buffer and a FIFO discipline. The second, less common but very natural, model can be described as a queue operating in slotted time with batch arrivals and services. It was studied for instance in [4]. The literature on discrete-time queues is relatively less furnished than its continuous-time counterpart. The interest for discrete-time queues is however gaining ground due to their relevance in ATM-like communication systems [6, 36].

In this paper, to clearly distinguish between the two models, we choose to describe the second one with a different terminology and as a storage model.

We prove first that the two models are linked in a very strong way. We set up an abstract model with an ordered pair of input variables $(\mathcal{A}, s)$ and an ordered pair of output variables $(\mathcal{D}, r) = \Phi(\mathcal{A}, s) = (\Phi_1(\mathcal{A}, s), \Phi_2(\mathcal{A}, s))$. On the one hand, the queueing model corresponds to $\mathcal{D} = \Phi_1(\mathcal{A}, s)$ with $\mathcal{A}$ and s being respectively the arrivals and services and $\mathcal{D}$ the departures. On the other hand, the storage model corresponds to $r = \Phi_2(\mathcal{A}, s)$, with s and $\mathcal{A}$ being respectively the supplies (arrivals) and requests (services) and r the departures. The interpretation of either r for the queue or $\mathcal{D}$ for the store is much less natural.

Then we assume that the random variables driving the dynamic are either exponentially or geometrically distributed, and we consider the models in equilibrium (under the stability condition). In this situation, it is well known that a Burke's type Theorem holds: the departures and the arrivals have the same law [7, 30, 4]. This can be considered as one of the cornerstones of queueing theory, see for instance the books [5, 15, 31] for discussions and related materials. Here we prove a 'joint' version of the result: $(\mathcal{D}, r)$ has the same law as $(\mathcal{A}, s)$. This joint Burke Theorem is new, although it is similar in spirit to the results proved in [25, 18] for a variant: queues with *unused services*. Also, in the geometric case, we use an original method of proof based on the reversibility of a symmetrized version of the workload process (instead of the queue length process). As in [25, 18], the joint Burke's Theorem can be used to obtain a representation theorem for the joint law of two independent random walks with exponential or geometric jumps conditioned on never colliding.

*LIAFA, Université Paris 7, case 7014, 2 place Jussieu, 75251 Paris Cedex 05, France. draief@liafa.jussieu.fr.
†LIAFA, Université Paris 7, case 7014, 2 place Jussieu, 75251 Paris Cedex 05, France. mairesse@liafa.jussieu.fr.
‡Mathematics Institute, Univ. Warwick, Coventry CV4 7AL, United Kingdom. noc@maths.warwick.ac.uk.

A second facet of the duality between queues and stores appears when considering the Robinson-Schensted-Knuth (RSK) algorithm. The RSK algorithm is an important object in representation theory and symmetric functions theory and it has been known for some time now that is has connections with queueing theory, although these connections have remained somewhat mysterious until now. In this paper we elucidate these connections completely via the above duality. The results we obtain, apart from being interesting in their own right, can be used as a starting point for the application of symmetric functions theory (and the associated analysis) to obtain interchangeability as well as asymptotic results for models which are constructed by putting the above queues and stores in series. These results are also applicable in the context of the corresponding interacting particle systems, where there is much interest in obtaining exactly solvable microscopic models for Burger-type equations.

Let us now detail the results. Consider K queues (stores) in tandem: the departures from a queue (store) are the arrivals (supplies) at the next one. Initially, the network is empty except for an infinite number of customers (infinite supply) in the first queue (store).

The particle systems associated with the above tandems are described in Section 5.2. One model is original and worth putting forward: the *bus-stop model* which is the zero-range model associated with stores in tandem.

Here we assume that the variables driving the dynamic are $\mathbb{N}$-valued without any further assumption. Building on ideas developed in [3, 23], we can study the transient evolution as follows. Consider the family of r.v.'s $(u(i,j))_{1\leq i\leq N, 1\leq j\leq K}$ where $u(i,j)$ is the service of the i-th customer at queue j, resp. the request at time slot i in store $K+1-j$. Apply the RSK algorithm, see [17, 32, 34], to this family and let P be the resulting semi-standard Young tableau (here we do not consider the recording tableau Q). Let $\lambda_1 \geq \cdots \geq \lambda_K \geq 0$ be the lengths of the successive rows of P. Classically [17], we have $\lambda_1 = \max_{\pi\in\Pi} \sum_{(i,j)\in\pi} u(i,j)$, where Π is the set of paths in $\mathbb{N}^2$ going from $(1,1)$ to (N,K) and which are increasing and consist of adjacent points. Moreover, it is well known in queueing theory [11, 21, 37, 35] that $\max_{\pi\in\Pi} \sum_{(i,j)\in\pi} u(i,j)$ is D: the instant of departure of customer N from queue K. Pasting the two results together gives the folklore observation that $\lambda_1 = D$. Here we complete the picture by proving that $\lambda_K = R$, where R is the total number of departures from the last store in the tandem up to time slot N. Again, this identity is proved in two steps by showing that $\lambda_K = \min_{\pi\in\widetilde{\Pi}} \sum_{(i,j)\in\pi} u(i,j) = R$, where $\widetilde{\Pi}$ is a different set of paths in the lattice. To summarize, we obtain on the same Young tableau the total departures for the two tandem models.

A short version without proofs of the paper appears in the proceedings of the conference *Discrete Random Walks 03* [10].

2 Notations

We work on a probability space $(\Omega, \mathcal{F}, P)$. The indicator function of an event $E \in \mathcal{F}$ is denoted by $\mathbb{1}_E$. We use the symbol $\sim$ to denote the equality in distribution of random variables. Depending on the context, $|A|$ is the cardinal of the set A or the length of the word A. We set $\mathbb{N}^* = \mathbb{N}\setminus\{0\}$, $\mathbb{R}_+^* = \mathbb{R}_+\setminus\{0\}$, and $x^+ = x\vee 0 = \max(x,0)$. We use the convention that $\sum_{i=j}^k u_i = 0$ when $j > k$.

Below, a *point process* is a stochastic simple point process on $\mathbb{R}$ with an infinite number of positive and negative points. We identify a point process $\mathcal{A}$ with the random ordered sequence of its points: $\mathcal{A} = (A_n)_{n\in\mathbb{Z}}$ with $A_n < A_{n+1}$ for all n. Observe that the numbering of the points is defined up to a translation in the indices. For any interval I, we define the *counting* random variable: $\mathcal{A}(I) = \sum_{n\in\mathbb{Z}} \mathbb{1}_{\{A_n\in I\}}$.

A *marked point process* is a couple $(\mathcal{A}, c) = (A_n, c_n)_{n\in\mathbb{Z}}$ where $\mathcal{A} = (A_n)_{n\in\mathbb{Z}}$ is a point process and $c = (c_n)_{n\in\mathbb{Z}}$ is a sequence of r.v.'s valued in some state space. The *mark* c_n is *associated* to the point A_n of the point process. For precisions concerning point processes see [9].

Given a point process $\mathcal{A} = (A_n)_{n\in\mathbb{Z}}$, the *reversed* point process $\mathcal{R}(\mathcal{A})$ is the point process obtained by reversing the direction of time; i.e. $\mathcal{R}(\mathcal{A}) = (-A_{-n})_{n\in\mathbb{Z}}$. Given a marked point process $(\mathcal{A}, c) = (A_n, c_n)_{n\in\mathbb{Z}}$, the *reversed* marked point process is $\mathcal{R}(\mathcal{A}, c) = (-A_{-n}, c_{-n})_{n\in\mathbb{Z}}$.

Given a càdlàg, i.e. right-continuous and left-limited, random process $Y = (Y(t))_{t\in\mathbb{R}}$ valued in $\mathbb{R}$, we define the *reversed* process $\mathcal{R}\circ Y = (\mathcal{R}\circ Y(t))_{t\in\mathbb{R}}$ as the càdlàg modification of the process $(Y(-t))_{t\in\mathbb{R}}$. Denote by $\mathcal{N}_+(Y)$ and $\mathcal{N}_-(Y)$ the point processes (with a possibly finite number of points) corresponding

respectively to the positive and negative jumps of Y, that is for any interval I of $\mathbb{R}$,

$$\mathcal{N}_+(Y)(I) = \int_I \mathbb{1}_{\{Y(u)>Y(u-)\}} du,\ \mathcal{N}_-(Y)(I) = \int_I \mathbb{1}_{\{Y(u)<Y(u-)\}} du\,. \tag{1}$$

3 The Model

Let $\mathcal{A} = (A_n)_{n\in\mathbb{Z}}$ be a point process and assume that $A_0 \le 0 < A_1$. We define the $\mathbb{R}_+^*$-valued sequence of r.v's $a = (a_n)_{n\in\mathbb{Z}}$ by $a_n = A_{n+1} - A_n$. Let $s = (s_n)_{n\in\mathbb{Z}}$ be another $\mathbb{R}_+^*$-valued sequence of r.v's. The marked point process $(\mathcal{A}, s)$ is the *input* of the model.

Define the sequence of r.v.'s $\mathcal{D} = (D_n)_{n\in\mathbb{Z}}$ by

$$D_n = \sup_{k\le n} \Big[A_k + \sum_{i=k}^{n} s_i \Big]\,. \tag{2}$$

A priori the D_n's are valued in $\mathbb{R}\cup\{+\infty\}$. Assume from now on that $(\mathcal{A}, s)$ is such that the D_n's are almost surely finite. They satisfy the recursive equations:

$$D_{n+1} = \max(D_n, A_{n+1}) + s_{n+1}\,. \tag{3}$$

Set $\mathcal{D} = (D_n)_{n\in\mathbb{Z}}$ and set $d_n = D_{n+1} - D_n$. Define an additional sequence of r.v.'s $r = (r_n)_{n\in\mathbb{Z}}$, valued in $\mathbb{R}_+^*$, by

$$r_n \quad = \quad \min(D_n, A_{n+1}) - A_n\,. \tag{4}$$

The marked point process $(\mathcal{D}, r)$ is the *output* of the model. By summing (3) and (4), we get the following interesting relation between input and output variables:

$$r_n + d_n = a_n + s_{n+1}\,. \tag{5}$$

In view of the future analysis, it is convenient to define the following auxiliary variables. Let $w = (w_n)_{n\in\mathbb{Z}}$ be the sequence of r.v.'s valued in $\mathbb{R}_+$ and defined by

$$w_n = D_n - s_n - A_n = \sup_{k\le n-1} \Big[\sum_{i=k}^{n-1} (s_i - a_i) \Big]^+\,. \tag{6}$$

These r.v.'s satisfy the recursive equations:

$$w_{n+1} = [w_n + s_n - a_n]^+\,. \tag{7}$$

Using the variables w_n, we can give alternative definitions of D_n and r_n:

$$\forall l \le n,\ D_n = \Big[w_l + A_l + \sum_{i=l}^{n} s_i \Big] \vee \max_{l<k\le n} \Big[A_k + \sum_{i=k}^{n} s_i \Big]\,, \tag{8}$$

$$r_n = \min\{w_n + s_n, a_n\} = s_n + w_n - w_{n+1}\,. \tag{9}$$

At last, we define the càdlàg random process $Q = (Q(t))_{t\in\mathbb{R}}$, valued in $\mathbb{N}$, by

$$Q(t) = \sum_{n\in\mathbb{Z}} \mathbb{1}_{\{A_n\le t<D_n\}}\,. \tag{10}$$

Lemma 3.1. *We have $\mathcal{N}_+(Q) = \mathcal{A}$ and $\mathcal{N}_-(Q) = \mathcal{D}$. Furthermore, $D_{-Q(0)} \le 0 < D_{-Q(0)+1}$.*

Proof. Recall that $Q(0) = \sum_{n\in\mathbb{Z}} \mathbb{1}_{\{A_n\le 0<D_n\}}$ and that $A_0 \le 0 < A_1$. Hence, $Q(0) = |k|$ where $k = \sup\{n \in \mathbb{Z}_- \mid D_n \le 0\}$. The result follows. □

We now interpret the variables defined above in two different contexts: a queueing model and a storage model.

3.1 The single-server queue

A bi-infinite string of customers is served at a queueing facility with a single server. Each customer is characterized by an instant of arrival in the queue and a service demand. Customers are served upon arrival in the queue and in their order of arrival. Since there is a single server, a customer may have to wait in a buffer before the beginning of its service. Using Kendall's nomenclature, our model is a ././1/∞/FIFO queue.

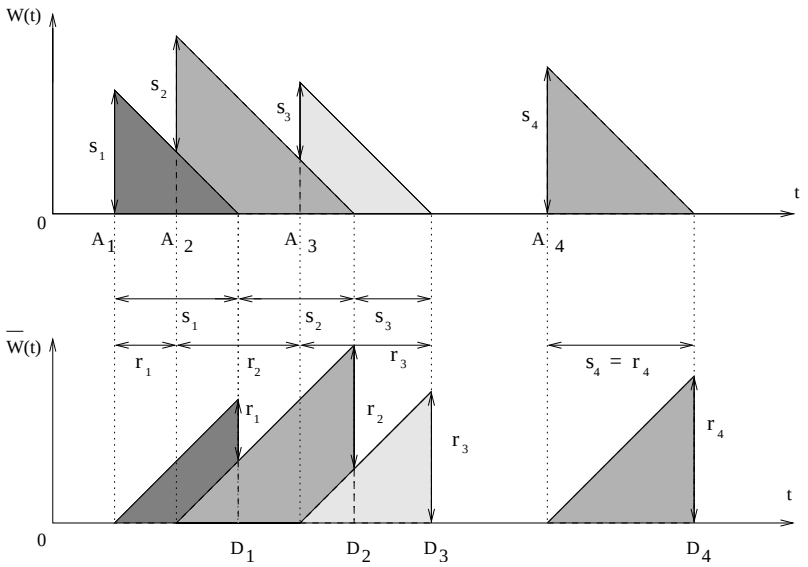

Figure 1: The dual variables $(s_n)_n$ and $(r_n)_n$

The customers are numbered by $\mathbb{Z}$ according to their order of arrival in the queue (customer 1 being the first one to arrive strictly after instant 0). Let A_n be the instant of **A**rrival of customer n and s_n its **S**ervice time. Then the variables defined in (2)-(10) have the following interpretations:

- D_n is the instant of departure of customer n from the queue, after completion of its service;
- w_n is the waiting time of customer n in the buffer between its arrival and the beginning of its service;
- $Q(t)$ is the number of customers in the queue at instant t (either in the buffer or in service); $Q = (Q(t))_t$ is called the *queue-length* process;
- r_n is the time spent by customer n at the *very back* of the queue.

The variables $(r_n)_n$ are less classical in queueing theory although they have already been considered [26]. They should be viewed as being dual to the services $(s_n)_n$ as illustrated in Figure 1. On the upper part of the figure, we have represented the *workload* process $(W(t))_t$, where $W(t)$ is the waiting time of a virtual customer arriving at instant t (see (12) for the formal definition).

3.2 The storage model

Some product P is supplied, sold and stocked in a store in the following way. Events occur at integer-valued epochs, called *slots*. At each slot, an amount of P is supplied and an amount of P is asked for by potential buyers. The rule is to meet all the demand, if possible. The demand of a given slot which is not met *is lost*. The supply of a given slot which is not sold *is not lost* and is stocked for future consideration.

Let s_n be the amount of P **S**upplied at slot $n+1$, and let a_n be the amount of P **A**sked for at the same slot. The variables in (2)-(10) can be interpreted in this context:

- w_n is the level of the stock at the end of slot n. It evolves according to (7);

- r_n is the demand met at slot $n+1$, see equation (9); it is the amount of P departing at slot $n+1$;

The variables $(D_n)_n$ and the function Q do not have a natural interpretation in this model.

The evolution of the store is summarized in Figure 2. The indices may seem weird (s_n, a_n, r_n, for time slot $n+1$). They were chosen that way to get better looking formulas in §5.

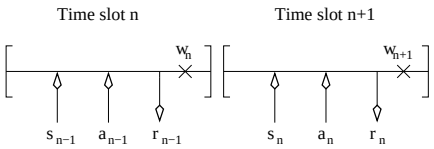

Figure 2: The storage model

It is important to remark that while the equations driving the *single server queue* and the *storage model* are exactly the same, it is not the same variables that make sense in the two models. The important variables are the ones corresponding to the departures from the system. The departures are coded in the variables $(D_n)_n$ for the single server queue and in the variables $(r_n)_n$ for the storage model. On the other hand, interpreting the variables $(r_n)_n$ in the single server queue or the variables $(D_n)_n$ in the storage model is not so immediate.

To summarize, the output variables $(D_n)_n$ and $(r_n)_n$ are equally relevant, but in different modelling contexts.

Observe that it would be possible to describe the above storage model with a "queueing" terminology (a queue with slotted time, batch arrivals (s_n) and batch services (a_n)). It is the description used in [4] for instance. We have avoided it on purpose to clearly separate the models of §3.1 and §3.2 and therefore minimize the possibility of confusion.

4 Equilibrium Behavior: Around Burke's Output Theorem

We consider the exponential or geometric version of our model in equilibrium. We prove a Burke's type result: $(\mathcal{D}, r) \sim (\mathcal{A}, s)$. The relevant result for the sequence of departures in one of the two models will then follow by forgetting one of the two variables. In the exponential case, our proof uses the queue-length process Q following [30]. Our method of proof in the geometric case is original and based on the zigzag process, a symmetrized version of the workload process. A discussion of the literature and of the increment of the present version is carried out in §4.3.

4.1 Output theorem in the exponential case

Let $\mathcal{A}$ be an homogeneous Poisson process of intensity $\lambda \in \mathbb{R}_+^*$. We set $\mathcal{A} = (A_n)_{n\in\mathbb{Z}}$ with $A_0 < 0 < A_1$. Recall that $a_n = A_{n+1} - A_n$. Then $(a_n)_{n\geq 1}$ is a sequence of i.i.d. r.v.'s with exponential distribution of parameter λ. Let $s = (s_n)_{n\in\mathbb{Z}}$ be a sequence of i.i.d. random variables, independent of $\mathcal{A}$, with exponential distribution of parameter μ. We assume that $\lambda < \mu$.

Consider the marked point process $(\mathcal{A}, s)$ as being the input of the model of Section 3. The sequence $(w_n)_n$ is a random walk valued in $\mathbb{R}_+$ with a barrier at 0. Under the *stability* condition $\lambda < \mu$, this random walk has a negative drift. It implies that the random variables D_n defined in (2) are indeed almost surely finite.

We now prove Theorem 4.1. Figure 2 provides a visual illustration of the result. The principle of the proof is the same as in Reich [30]. However, the adaptation is not completely immediate. Some care is necessary in dealing with the indices, to make sure that the variable r_n is the mark of the point D_n.

Theorem 4.1. *The marked point process $(\mathcal{D}, r)$ has the same law as the marked point process $(\mathcal{A}, s)$.*

Proof. Knowing Q, one can recover the input of the model: $(A_n, s_n)_n = \varphi(Q)$. We are going to prove that

$$\varphi \circ \mathcal{R}(Q) = (-D_{-n+Q(0)+1}, r_{-n+Q(0)+1})_n = \mathcal{R}(\mathcal{D}, r)$$

Let $\mathcal{T}_n$ be the operator mapping a point process to its n-th point, with the convention that $\mathcal{T}_0(\cdot) \leq 0 < \mathcal{T}_1(\cdot)$. For instance, $A_n = \mathcal{T}_n(\mathcal{A}) = \mathcal{T}_n \circ \mathcal{N}_+(Q)$. Taking into account the shift in the indices and using Lemma 3.1, we get that $\mathcal{T}_n(\mathcal{D}) = \mathcal{T}_n \circ \mathcal{N}_-(Q) = D_{n-Q(0)}$. So we have

$$\begin{aligned} s_n &= D_n - \max(A_n, D_{n-1}) \\ &= \mathcal{T}_{n+Q(0)} \circ \mathcal{N}_-(Q) - \max(\mathcal{T}_n \circ \mathcal{N}_+(Q), \mathcal{T}_{n+Q(0)-1} \circ \mathcal{N}_-(Q)) \,. \end{aligned}$$

Set $s_n = \mathcal{S}_n(Q)$. Now let us apply the operators $\mathcal{T}_n \circ \mathcal{N}_+$ and $\mathcal{S}_n$ to the reversed process $\mathcal{R}(Q)$:

$$\mathcal{T}_n \circ \mathcal{N}_+ \circ \mathcal{R}(Q) = -D_{-n-Q(0)+1}$$

Furthermore $\mathcal{T}_{n+\mathcal{R}(Q)(0)-1} \circ \mathcal{N}_- \circ \mathcal{R}(Q) = -A_{-n-Q(0)+2}$. Thus,

$$\begin{aligned} \mathcal{S}_n \circ \mathcal{R}(Q) &= -A_{-n-Q(0)+1} \,-\, \max(\, -D_{-n-Q(0)+1},\, -A_{-n-Q(0)+2}\,) \\ &= -A_{-n+Q(0)+1} \,+\, \min(\, D_{-n+Q(0)+1},\, A_{-n+Q(0)+2}\,) \\ &= r_{-n+Q(0)+1} \,. \end{aligned} \tag{11}$$

Hence we obtain that $\varphi \circ \mathcal{R}(Q) = \mathcal{R}(\mathcal{D}, r)$.

The process Q is a stationary birth-and-death process, hence reversible: $\mathcal{R}(Q) \sim Q$. It implies that $(\mathcal{A}, s) = \varphi(Q) \sim \varphi \circ \mathcal{R}(Q) = \mathcal{R}(\mathcal{D}, r)$. Hence, $\mathcal{R}(\mathcal{D}, r)$ is an homogeneous Poisson process marked with an i.i.d. sequence of r.v.'s. So, $\mathcal{R}(\mathcal{D}, r) \sim (\mathcal{D}, r)$, which concludes the proof. □

Corollary 4.2. *In the queueing model, the departure process $\mathcal{D}$ is a Poisson process of intensity λ. In the storage model, the sequence $(r_n)_n$ of the amounts of product P departing at successive slots, is a sequence of i.i.d. exponential r.v.'s of parameter μ.*

4.2 Output theorem in the geometric case

Let $\mathcal{A}$ be a Bernoulli point process of parameter $p \in (0,1)$, that is: all the points are integer valued, there is a point at a given integer with probability p, and the presence of points at different integers are independent. As before, set $\mathcal{A} = (A_n)_n$ with $A_0 \leq 0 < A_1$ and $a_n = A_{n+1} - A_n$. Then the sequence $(a_n)_{n\geq 1}$ is a sequence of i.i.d. geometric r.v.'s with parameter p ($\forall k \in \mathbb{N}^*$, $P\{a_1 = k\} = (1-p)^{k-1}p$). Let $(s_n)_n$ be a sequence of i.i.d. geometric r.v.'s with parameter $q \in (0,1)$, independent of $\mathcal{A}$. We assume that $p < q$ (stability condition).

Let the marked point process $(\mathcal{A}, s)$ be the input of the model of Section 3. As in §4.1, the model is stable and the output $(\mathcal{D}, r)$ is a marked point process.

We prove a Burke-type Theorem using an original approach based on a symmetrization of the workload process.

Define the (random) càdlàg maps $f_n, g_n : \Omega \times \mathbb{R} \to \mathbb{R}$, by

$$f_n(t) = \begin{cases} 0 & \text{if } t < A_n \\ D_n - t & \text{if } A_n \leq t \leq D_n \\ 0 & \text{if } t > D_n \end{cases}, \qquad g_n(t) = \begin{cases} 0 & \text{if } t < A_n \\ t - A_n & \text{if } A_n \leq t \leq D_n \\ 0 & \text{if } t > D_n \end{cases}.$$

Set

$$W(t) = \bigvee_{n\in\mathbb{Z}} f_n(t), \qquad \overline{W}(t) = \bigvee_{n\in\mathbb{Z}} g_n(t) \,. \tag{12}$$

Since $D_n = A_n + w_n + s_n$, we have $W(A_n^-) = w_n$. For the queueing model, $W(t)$ is the total amount of service remaining to be done by the server at instant t, and $W = (W(t))_t$ is called the *workload* process. The process $\overline{W} = (\overline{W}(t))_t$ is the dual of the process W, see Figure 1 or Figure 3 for an illustration.

Looking at W and $\overline{W}$ together is a way to symmetrize the workload process. Consider the ordered pair of processes $P = (W, \overline{W})$. Define

$$\mathcal{R}(P) = (\mathcal{R}(\overline{W}), \mathcal{R}(W))$$

(observe that the first element of the pair is now $\mathcal{R}(\overline{W})$, as opposed to W for the pair P). Obviously, we can recover the input of the model knowing P: $(\mathcal{A}, s) = \Psi(P)$. Also it is clear (see Figure 1 or Figure 3) that: $\mathcal{R}(\mathcal{D}, r) = \Psi \circ \mathcal{R}(P)$.

The crux of the argument in proving a Burke-type Theorem is now the next Lemma.

Lemma 4.3. *The ordered pair of the workload process and its dual is reversible, that is:* $(W, \overline{W}) \sim (\mathcal{R}(\overline{W}), \mathcal{R}(W))$.

Proof. Clearly, W and $\overline{W}$ are (time) stationary and ergodic processes. An *idle period* is a maximal non-empty time interval during which $W(t) = 0$ (or equivalently, $\overline{W}(t) = 0$). The *busy periods* are the time intervals between the idle periods. Hence $\mathbb{R}$ is partitioned by an alternating sequence of busy and idle periods.

Let $(i_n)_{n\in\mathbb{Z}}$ and $(b_n)_{n\in\mathbb{Z}}$ be the lengths of the successive idle periods and busy periods respectively. Clearly, the sequences $(i_n)_n$ and $(b_n)_n$ are independent and i.i.d. Moreover, it follows from the memoryless property of the geometric distribution that i_n is a geometric distribution of parameter p. So, to prove the reversibility of $(W, \overline{W})$, it is enough to prove the reversibility within a single busy period.

To do so, the easiest is to introduce an auxiliary process. Consider a busy period which is assumed to consist, for simplicity, of the k customers numbered from 1 up to k. Set for $n \in \{1, \dots k\}$:

$$C_n = B_n + s_n, \quad B_{n+1} = C_n + a_n \,. \tag{13}$$

This defines $(B_n)_n$ and $(C_n)_n$ up to a translation. Define

$$Z(B_1) = 0, \qquad Z(t) = \begin{cases} Z(B_n) + (t - B_n) & \text{for } t \in [B_n, C_n) \\ \big[Z(C_n) - (t - C_n)\big]^+ & \text{for } t \in [C_n, B_{n+1}) \end{cases} . \tag{14}$$

On an interval of type $[B_n, C_n)$, we have $dZ/dt = 1$ and on an interval of type $[C_n, B_{n+1})$, we have $dZ/dt = (-1)\mathbb{1}_{Z>0}$.

Now, on a busy period, the process Z is in bijection with $(W, \overline{W})$, and the time-reversed process $\mathcal{R}(Z)$ is in bijection with $(\mathcal{R}(\overline{W}), \mathcal{R}(W))$. This is illustrated in Figure 3, where we have represented a trajectory of $(W, \overline{W})$ and the corresponding trajectory of Z. The process Z is obtained by gluing together the slices of $\overline{W}$ and W alternatively.

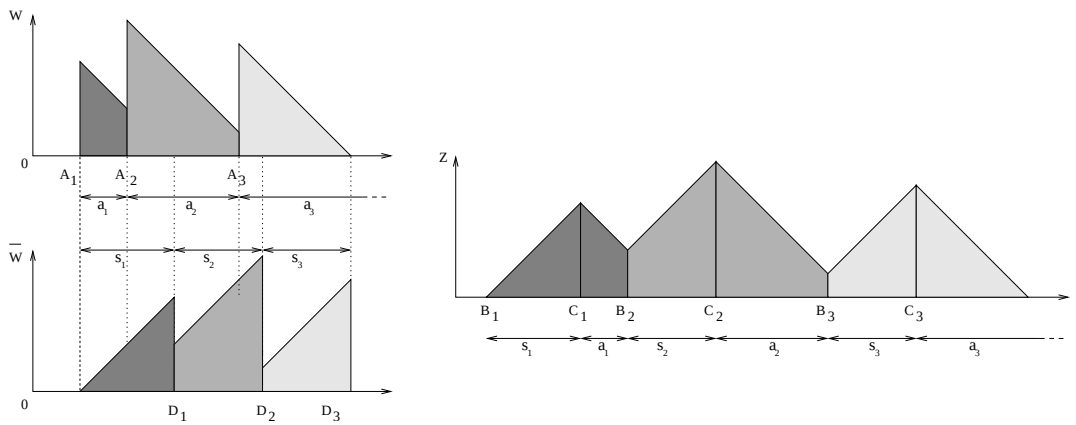

Figure 3: The workload process, the dual workload process, and the zigzag process

Hence, proving the reversibility of $(W, \overline{W})$ on a busy period, is equivalent to proving the reversibility of Z. We may call Z the *zigzag* process. This process was considered (with a different motivation) in [12] and in particular, it was proved there that the zigzag process is reversible (although not Markovian). We recall the argument for completeness.

Identify Z with the finite sequence of the lengths of its intervals of increase and decrease. Let $(l_1, \ldots, l_{2k}) \in (\mathbb{N}^*)^{2k}$, with $\sum_i l_{2i} = \sum_i l_{2i+1} = L$, be a possible value for Z. (The total length of the periods of increase has to be equal to the total length of the periods of decrease). Recall that the variables $(s_n)_n$ and $(a_n)_n$ are geometrically distributed of respective parameters q and p. We get:

$$\begin{aligned} P\big\{Z = (l_1, \ldots, l_{2k})\big\} &= (1-p)^{l_1-1}p(1-q)^{l_2-1}q \cdots (1-p)^{l_{2k-1}-1}p(1-q)^{l_{2k}-1} \\ &= (1-p)^{L-k}p^k(1-q)^{L-k}q^{k-1} \, . \end{aligned}$$

Hence, given L and k, all the different possible trajectories for Z are equiprobable. It implies that the same is true for the time-reversed process $\mathcal{R}(Z)$. This completes the proof. □

We are now ready to prove the following result:

Theorem 4.4. *The marked point process $(\mathcal{D}, r)$ has the same law as the marked point process $(\mathcal{A}, s)$.*

Proof. The proof is similar to the one of Theorem 4.1 with $(W, \overline{W})$ playing the role of Q.
Recall that $(\mathcal{A}, s) = \Psi(W, \overline{W})$ and $\mathcal{R}(\mathcal{D}, r) = \Psi \circ \mathcal{R}(W, \overline{W})$. According to Lemma 4.3, we have

$$(\mathcal{A}, s) = \Psi(W, \overline{W}) \sim \Psi \circ \mathcal{R}(W, \overline{W}) = \mathcal{R}(\mathcal{D}, r) \, .$$

We conclude that $(\mathcal{D}, r) \sim \mathcal{R}(\mathcal{A}, s) \sim (\mathcal{A}, s)$, where the last equivalence comes from the fact that $(\mathcal{A}, s)$ is a Bernouilli process with i.i.d. marks, hence is reversible. □

4.3 Comments on the different proofs of Burke Theorem

Reflecting on the above, there are three different ways to prove Burke Theorem, be it in the exponential or geometric case. The first way is by using analytic methods, the second is by using the reversibility of the queue length process Q, and the third is by using the reversibility of the zigzag process Z.

The original proof of Burke is for the exponential model using analytic methods [7]. For the geometric model, an analytic proof was given by Azizoğlu and Bedekar [4]. For the exponential model, the idea of using the reversibility of Q to get the result is due to Reich [30]. This proof has become a cornerstone of queueing theory, it has been extended to various contexts and has given birth to the concept of *product form networks* [5, 15]. Reich's proof does not translate directly to the geometric model. Of course, $(Q(n))_{n\in\mathbb{Z}}$ is a reversible birth-and-death Markov chain. However, a difficulty arises: It is not possible to reconstruct $\mathcal{A}$ and s from Q. Indeed, on the event $\{Q(n-1) = Q(n) > 0\}$, two cases may occur: there is either no departure and no arrival at instant n, or one departure and one arrival; and it is not possible to distinguish between them knowing only Q. One feasible solution is to add an auxiliary sequence that contains the lacking information but the details become quite intricate. This program has been carried out in [14], see also [18, Theorem 4.1], for a variant: the geometric model with *unused services*. The above idea of using the zigzag process to prove Burke Theorem is original. Observe that the zigzag process is also clearly reversible in the exponential model. Hence the proof used for Theorem 4.4 could also be used to get Theorem 4.1.

In the original references [7, 30] and in all the classical textbooks presenting the result [5, 15, 31], the version proved is: $\mathcal{A} \sim \mathcal{D}$: ("Poisson Input Poisson Output"). In [4], or [14] for the variant with *unused services*, the result proved is $s \sim r$. The complete version $(\mathcal{A}, s) \sim (\mathcal{D}, r)$ appears first in [25, Theorem 3] for the exponential model with *unused services*. This is extended to the geometric model with *unused services* in [18]. Brownian analogues are proved in [13, 24].

4.4 Non-colliding random walks

Following the lines of thought in [25, 18, 12], it is possible to use Theorems 4.1 and 4.4 to get representation results for non-colliding random walks.

We start by stating a preliminary lemma.

Lemma 4.5. *The Lindley's Equation (7) admits a backward analogue:*

$$w_n = (w_{n+1} + r_n - d_{n-1})^+ \, . \tag{15}$$

Proof. Recall Equation (9): $r_n = s_n + w_n - w_{n+1}$. Using (5), we get

$$d_n = a_n + s_{n+1} - r_n = a_n - w_n + w_{n+1} - s_n + s_{n+1}\,. \tag{16}$$

Therefore,

$$\begin{aligned} w_{n+1} + r_n - d_{n-1} &= w_{n+1} + s_n + w_n - w_{n+1} - a_{n-1} - w_n + w_{n-1} - s_n + s_{n-1} \\ &= w_{n-1} + s_{n-1} - a_{n-1}\,. \end{aligned}$$

Taking the maximum with 0 on both sides of the equality and using Lindley's equation (7), we obtain $(w_{n+1} + r_n - d_{n-1})^+ = w_n$. □

Proposition 4.6. *Let the sequences of r.v.'s $(a_n)_{n\in\mathbb{N}^*}$ and $(s_n)_{n\in\mathbb{N}^*}$ be as in §4.1 or as in §4.2. The conditional law of $(\sum_{i=1}^n a_i,\ \sum_{i=1}^n s_i)$, given that $\{\sum_{i=1}^k a_i \geq \sum_{i=1}^{k+1} s_i,\ \forall k \geq 0\}$ is the same as the unconditional law of $(\max_{1\leq j\leq n}\{\sum_{i=1}^j a_i + \sum_{i=j+1}^{n+1} s_i\},\ \min_{1\leq j\leq n}\{\sum_{i=2}^j s_i + \sum_{i=j+1}^n a_i\})$.*

Proof. Set $v_n = w_n + s_n = w_{n+1} + r_n$ (the *sojourn time* of customer n in the queue). By developing (15), we obtain

$$v_n = w_{n+1} + r_n = \sup_{k\geq n}\{\sum_{i=n}^k r_i - \sum_{i=n}^{k-1} d_i\}\,. \tag{17}$$

By summing the equalities (16), we obtain on the event $\{v_1 = 0\}$:

$$\begin{aligned} \sum_{i=1}^n d_i &= \sum_{i=1}^n a_i + w_{n+1} + s_{n+1} \\ &= \sum_{i=1}^n a_i + \max_{1\leq j\leq n}\{\sum_{i=j+1}^{n+1} s_i - \sum_{i=j+1}^n a_i\} \;=\; \max_{1\leq j\leq n}\{\sum_{i=1}^j a_i + \sum_{i=j+1}^{n+1} s_i\}\,, \end{aligned}$$

and summing (9) on the event $\{v_1 = 0\}$, we get:

$$\begin{aligned} \sum_{i=1}^n r_i &= \sum_{i=2}^n s_i - w_{n+1} \\ &= \sum_{i=2}^n s_i - \max_{1\leq j\leq n}\{\sum_{i=j+1}^n s_i - \sum_{i=j+1}^n a_i\} \;=\; \min_{1\leq j\leq n}\{\sum_{i=2}^j s_i + \sum_{i=j+1}^n a_i\}\,. \end{aligned}$$

To summarize, we have

$$\sum_{i=1}^n d_i = \max_{1\leq j\leq n}\{\sum_{i=1}^j a_i + \sum_{i=j+1}^{n+1} s_i\},\quad \sum_{i=1}^n r_i = \min_{1\leq j\leq n}\{\sum_{i=2}^j s_i + \sum_{i=j+1}^n a_i\}\,. \tag{18}$$

Applying Theorem 4.1 or 4.4, the distribution of $(\sum_{i=1}^n a_i, \sum_{i=1}^n s_i)$ given that $\{\sum_{i=1}^k a_i \geq \sum_{i=1}^{k+1} s_i,\ \forall k \geq 0\}$ is the same as the law of $(\sum_{i=1}^n d_i,\ \sum_{i=1}^n r_i)$ given that $\{\sum_{i=1}^k d_i \geq \sum_{i=1}^{k+1} r_i,\ \forall k \geq 0\}$. By (17), the event $\{\sum_{i=1}^k d_i \geq \sum_{i=1}^{k+1} r_i,\ \forall k \geq 0\}$ is equal to $\{v_1 = 0\}$.
Using (18), we have that the distribution of $(\sum_{i=1}^n d_i,\ \sum_{i=1}^n r_i)$ given that $\{v_1 = 0\}$ is the same as the law of $(\max_{1\leq j\leq n}\{\sum_{i=1}^j a_i + \sum_{i=j+1}^{n+1} s_i\}, \min_{1\leq j\leq n}\{\sum_{i=2}^j s_i + \sum_{i=j+1}^n a_i\})$ given that $\{v_1 = 0\}$. This last conditional law is the same as the unconditional law of $(\max_{1\leq j\leq n}\{\sum_{i=1}^j a_i + \sum_{i=j+1}^{n+1} s_i\}, \min_{1\leq j\leq n}\{\sum_{i=2}^j s_i + \sum_{i=j+1}^n a_i\})$. Indeed, the variables $(a_i)_{i\geq 1}$ and $(s_i)_{i\geq 2}$ are independent of $\{v_1 = 0\}$.
By fitting all the parts together, we obtain the desired result. □

It is possible to extend Proposition 4.6 to the limiting case $E[a_1] = E[s_1]$, and also to higher dimensions, by adapting the methods of [25, 18, 12] to the present setting.

5 Transient Behavior and RSK Representation

In Sections 5.1 and 5.3, the model described and the results obtained have a purely combinatorial nature. In particular, the results are true pathwise, without any probabilistic assumption. These results are then used in Section 5.4 for geometric queues or stores: here the pathwise arguments are coupled with probabilistic ones.

5.1 The saturated tandem

We consider another aspect of the dynamic of queues and stores: the transient evolution for the model starting empty. More precisely, consider the model of §3 under the assumption that $w_0 = A_0 = s_0 = 0$ (which implies $D_0 = r_0 = 0$) and focus on the customers, resp. time slots, from 1 onwards.

It is convenient to describe such a model with a different perspective. We first do it for the queue. View the arrivals as being the departures from a *virtual queue* having at instant 0 an infinite number of customers (labelled by $\mathbb{N}^*$) in its buffer. The service time of customer n in the virtual queue is a_{n-1}. We describe this as a *saturated tandem of two queues*.

Let us turn our attention to the store. View the supplies as being the departures from a virtual store having an infinite stock at the end of time slot 0. In the virtual store, the request (=departure) at time slot n is s_n. This is a *saturated tandem of two stores*.

Now we want to fit these two descriptions together. Denote the virtual queue/store as *queue/store 1* and the other one as *queue/store 2*. For convenience, set $u(n,1) = a_{n-1}$ and $u(n,2) = s_n$ for all $n \geq 1$. The saturated tandem is completely specified by the family $(u(n,i), n \in \mathbb{N}^*, i = 1,2)$ of input variables. These variables are the services, resp. requests, when the model is seen as a tandem of queues, resp. stores. Next table gives the input variables in the saturated tandems of two queues/stores.

Customer / Time slot	1	2	$\cdots$	n
Queue 2 / (Virtual) Store 1	$u(1,2) = s_1$	$u(2,2) = s_2$		$u(n,2) = s_n$
(Virtual) Queue 1 / Store 2	$u(1,1) = a_0$	$u(2,1) = a_1$		$u(n,1) = a_{n-1}$

A couple of observations are in order. Observe that $u(n,i)$ is the service of customer n in queue i, and the request at time slot n in store $3-i$. In other words, the elements (queues or stores) associated with a given sequence $(u(\cdot,i)), i = 1,2$, are crossed in reverse orders in the queueing/storage tandem. This is illustrated in Figure 4 (set $K = 2$). Observe also that there is a shift in the time slots for the storage model: the departure from store 1 at time slot n is the supply of store 2 at time slot $n+1$ (contrast this with the situation for the queues). This is consistent with a model in which we view a time slot as being decomposable in three consecutive stages: first, the supply arrives; second, the request is made; third, the departure occurs. See Figure 2.

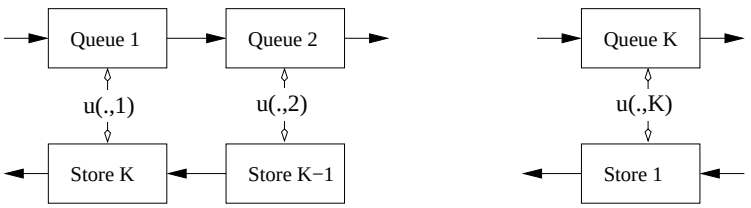

Figure 4: Queues and stores are gone through in opposite orders

The above setting is naturally extended to define the *saturated tandem of K queues/stores*. Such a model is entirely defined by a family of $\mathbb{N}$-valued variables $(u(i,j))_{i\in\mathbb{N}^*, j\in\{1,\ldots,K\}}$. (In §3, the variables were valued in $\mathbb{N}^*$. This restriction is not necessary here.) For the queueing model: (i) at instant 0, queue 1 has an infinite number of customers labelled by $\mathbb{N}^*$ in its buffer, and the other queues are empty; (ii) $u(i,j)$ is the service of customer i at queue j; (iii) the instant of departure of customer n from queue i is the instant of arrival of customer n in queue $i+1$. For the storage model: (i) at the end of time slot 0, store 1 has an infinite stock and the other stores have an empty stock; (ii) $u(i, K+1-j)$ is the request at time slot i in store j; (iii) the departure at time slot n from store i is the supply at time slot $n+1$ in store $i+1$. The models are depicted in Figure 4.

5.2 Interacting particles representation

It is helpful to describe the above tandem models as interacting particle models (see for instance [19]). Commonly, interacting particle systems are defined as continuous-time Markov processes. In contrast, the description below is pathwise and the time is discrete.

Classically, the queues in tandem may be viewed as the following *zero-range process*. The set of sites is $\{1, \ldots, K\}$ and a site is occupied by a number of particles in $\mathbb{N} \cup \{+\infty\}$. At each site, the front particle (if any) has a clock whose value is decremented by 1 at each time slot. When it reaches 0, the particle jumps to the site on its right. The right-most site has an infinite number of particles. This is illustrated in Figure 5. The variables $(u(i,j))_i$ correspond to the initial values of the clocks of the successive front particles at site j.

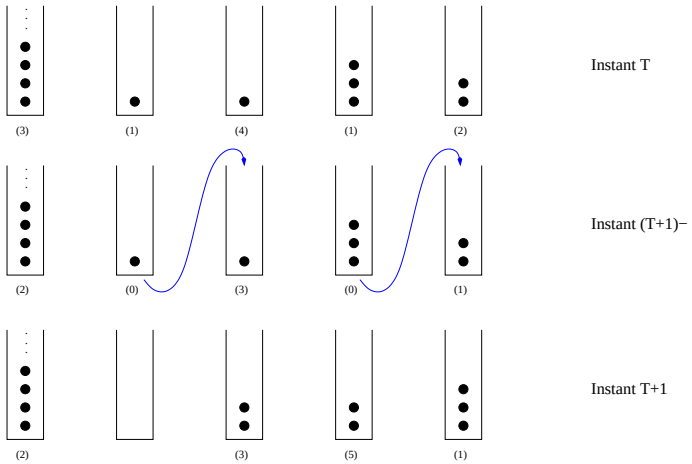

Figure 5: The zero-range process for queues in tandem

We now describe the stores in tandem as an original zero-range model: the *bus-stop model*. At the beginning of each time slot, a bus arrives at each site. The size of the bus varies from site to site, and from time slot to time slot. The particles take place in the bus until it is filled in, and are transported to the next site on the left, during the current time slot. The left-most site has an infinite number of particles. This mechanism is illustrated in Figure 6. Here the variables $(u(i,j))_i$ correspond to the successive bus sizes at site j.

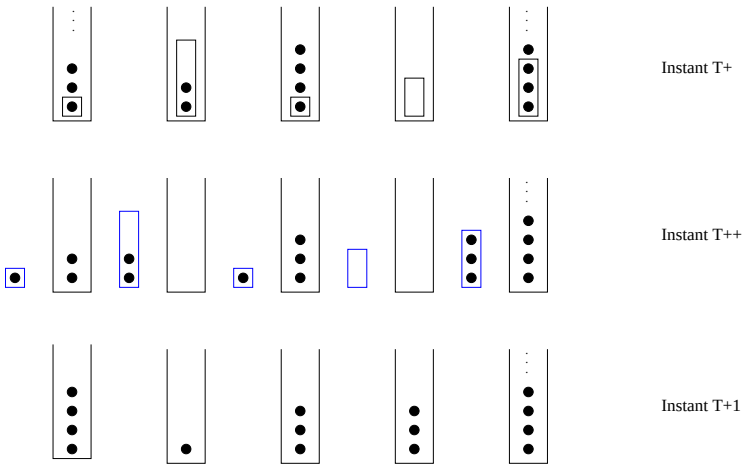

Figure 6: The zero-range process for stores in tandem

Alternatively, the above two models can also be viewed as *exclusion processes*. The set of sites is embedded in $\mathbb{N}$ and each site now contains either 0 or 1 particle.

The queues in tandem are associated with the classical TASEP or totally asymmetric exclusion process. For the stores in tandem, the exclusion process is as follows. Particles take jumps of integer length to the right which are constrained by the interdiction to overpass other particles. In one time step, particles jump in order from left to right. This model is close but different from the exclusion processes considered in [28, 33].

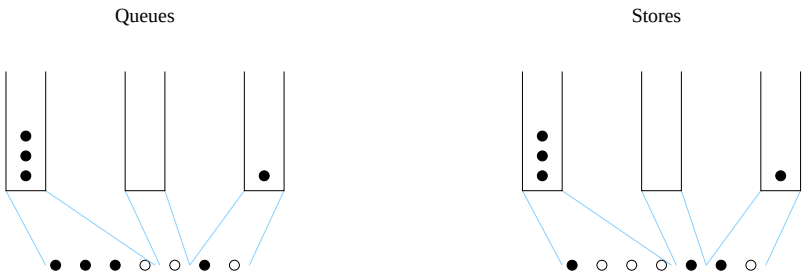

Figure 7: Exclusion processes for the queues and the stores

The mapping between zero-range process and exclusion process, both for the queues and the stores, is illustrated in Figure 7.

5.3 Robinson-Schensted-Knuth representation

A *partition* of $n \in \mathbb{N}^*$ is a sequence of integers $\lambda = (\lambda_1, \dots, \lambda_k)$ such that $\lambda_1 \geq \cdots \geq \lambda_k \geq 0$ and $\lambda_1 + \cdots + \lambda_k = n$. We use the notation $\lambda \vdash n$. By convention, we identify partitions having the same non-zero components. The *(Ferrers) diagram* of $(\lambda_1, \dots, \lambda_k) \vdash n$ is a collection of n boxes arranged in left-justified rows, the i-th row starting from the top consisting of λ_i boxes. A *(semi-standard Young) tableau* on the alphabet $\{1, \dots, \ell\}$ is a diagram in which each box is filled in by a label from $\{1, \dots, \ell\}$ in such a way that the entries are weakly increasing from left to right along the rows and strictly increasing down the columns. The *shape* of a diagram or tableau is the underlying partition. A *standard tableau* of size n is a tableau of shape $\lambda \vdash n$ whose entries are from $\{1, \dots, n\}$ and are distinct. In Figure 8 we have represented a diagram on the left and a tableau of the same shape on the right.

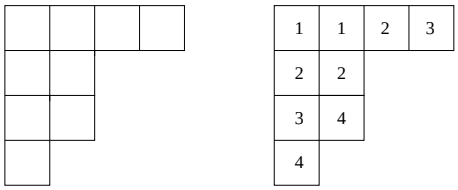

Figure 8: Ferrers diagram and semi-standard Young tableau of shape $(4, 2, 2, 1) \vdash 9$

The Robinson-Schensted-Knuth row-insertion algorithm (RSK algorithm) takes a tableau T and $i \in \mathbb{N}^*$ and constructs a new tableau $T \leftarrow i$. The tableau $T \leftarrow i$ has one more box than T and is constructed as follows. If i is at least as large as the labels of the first (upper) row of T, add a box labelled i to the end of the first row of T and stop the procedure. Otherwise, find the leftmost entry in the first row which is strictly larger than i, relabel the corresponding box by i and apply the same procedure recursively to the second row and to the bumped label. By convention, an empty row has label 0. With this convention, the above procedure stops.

Consider a word $v = v_1 \cdots v_n$ over the alphabet $\{1, \dots, k\}$. The tableau associated with v is by definition

$$P = (\cdots((T_0 \leftarrow v_1) \leftarrow v_2) \cdots \leftarrow v_n)\,,$$

where T_0 is the empty tableau. Observe that P has at most k non-empty rows. Classically, the length of the top (and longest) row of P is equal to the longest weakly-increasing subsequence in v.

Remark 5.1. While building the tableau P, it is possible to build another tableau of the same shape in which the entries, labelled from 1 to $|v|$, record the order in which the boxes are added. This

recording tableau is a standard tableau of size $|v|$. By doing this, one defines a bijection between words of $\{1,\dots,k\}^n$ and ordered pairs of tableaux of the same shape, the first being semi-standard over the alphabet $\{1,\dots,k\}$ and the second being standard and of size n. This bijection is often referred to as the Robinson-Schensted-Knuth correspondence. Here, we do not need this result and we do not consider the recording tableau.

Consider a family $U = (u(i,j))_{(i,j)\in\{1,\dots,N\}\times\{1,\dots,K\}}$ of variables valued in $\mathbb{N}$. (Here we do not make any assumption on the variables, which may or may not be random.) We associate with U the word $w(U)$ over the alphabet $\{1,\dots,K\}$ defined by

$$w(U) = w_1 \dots w_N, \text{ with } w_i = \underbrace{1\cdots 1}_{u(i,1)}\ \underbrace{2\cdots 2}_{u(i,2)}\ \cdots\ \underbrace{K\cdots K}_{u(i,K)}\ . \tag{19}$$

Set $M = |w(U)| = \sum_{i=1}^{N}\sum_{j=1}^{K} u(i,j)$. For $i = 1,\dots,K$, define

$$\forall n \le M,\quad x_i(n) = |\{j \le n \,|\, w(U)_j = i\}|\,.$$

Given two maps $x, y : \{1,\dots,N\} \longrightarrow \mathbb{N}$, define the maps $x\triangledown y, x \vartriangle y : \{1,\dots,N\} \longrightarrow \mathbb{N}$ as follows: $\forall n \le N$,

$$x\triangledown y(n) = \max_{0\le m\le n}\Big[x(m) + y(n) - y(m)\Big],\quad x \vartriangle y(n) = \min_{0\le m\le n}\Big[x(m) + y(n) - y(m)\Big]\,. \tag{20}$$

Denote by $P(U)$ the tableau obtained from $w(U)$ by applying the RSK algorithm and let $(\lambda_1,\dots,\lambda_K)$ be its shape. The following holds

$$\lambda_1 = x_1\triangledown x_2\triangledown\dots\triangledown x_K(M),\quad \lambda_K = x_K \vartriangle \dots \vartriangle x_2 \vartriangle x_1(M)\,, \tag{21}$$

where the operations $\triangledown$, $\vartriangle$ are performed from left to right (the operations $\triangledown$, $\vartriangle$ are non-associative). The expression for λ_1 follows from the fact that λ_1 is the longest weakly-increasing subsequence in $w(U)$. The expression for λ_K is proved in [23, Theorem 3.1]. In fact, the result from [23] is more general: there exists a min-max-type operator Γ_K such that $(\lambda_1,\dots,\lambda_K) = \Gamma_K(x_1,\dots,x_K)$.

A *lattice path* is a sequence $\pi = ((i_1,j_1),\dots,(i_l,j_l))$ with $(i_k,j_k) \in \mathbb{Z}^2$. The *steps* of π are the differences $(i_{k+1}-i_k, j_{k+1}-j_k)$, $k = 1,\dots l-1$.

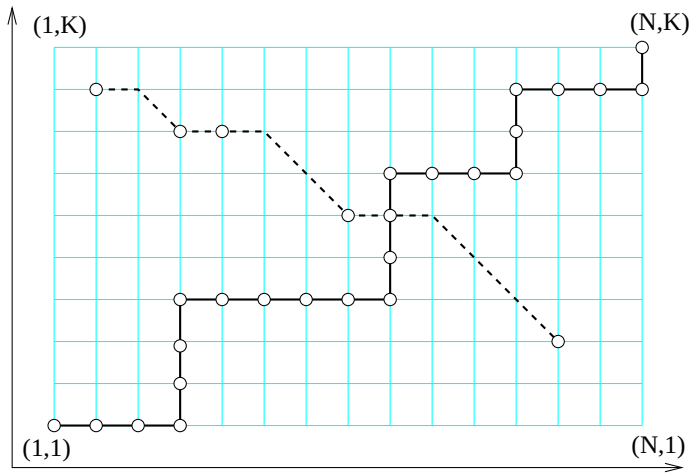

Figure 9: A path from Π (plain line) and a path from $\widetilde{\Pi}$ (dashed line)

Let Π be the set of lattice paths from (1,1) to (N,K) with steps of the type $(1,0)$ or $(0,1)$. All the paths in Π have the same number of nodes: $N+K-1$. Let $\widetilde{\Pi}$ be the set of lattice paths from one of the points in $\{(1+i, K-i), i = 0,\dots,K-1\}$ to one of the points in $\{(N-j, 1+j), j = 0,\dots,N-1\}$ with steps of the type $(1+i,-i)$, $i \ge 0$.

Observe that all the paths in $\widetilde{\Pi}$ also have the same number of nodes: $(N-K+1)^+$. In particular, $\widetilde{\Pi}$ is empty when $N < K$. An example of a path from each of the two sets is provided in Figure 9.

The following result holds. When interpreting it, recall that queues and stores are arranged in opposite orders in the tandems, see Figure 4. It is also fruitful to compare the statements of Proposition 4.6 and Theorem 5.2 for $K = 2$.

Theorem 5.2. *Consider a saturated tandem of K queues/stores with variables $u(i,j), i \in \mathbb{N}^*, j \in \{1,\dots,K\}$, as defined in §5.1. Fix $N \in \mathbb{N}^*$. Let D be the instant of departure of customer N from queue K and let R be the cumulative departures over the time slots 1 to N in store K. Set $U = (u(i,j))_{1\le i\le N\,,\,1\le j\le K}$ and define $w(U)$ as in (19). Let $(\lambda_1,\dots,\lambda_K)$ be the shape of the tableau associated with $w(U)$. We have*

$$\lambda_1 \;=\; \max_{\pi\in\Pi}\Big[\sum_{(i,j)\in\pi} u(i,j)\Big] \;=\; D \tag{22}$$

$$\lambda_K \;=\; \min_{\pi\in\widetilde{\Pi}}\Big[\sum_{(i,j)\in\pi} u(i,j)\Big] \;=\; R\,. \tag{23}$$

First, we state and prove a preliminary result.

Lemma 5.3. *We have*

$$x_K \,\triangle\, \dots \,\triangle\, x_2 \,\triangle\, x_1(M) = \min_{\pi\in\widetilde{\Pi}}\Big[\sum_{(i,j)\in\pi} u(i,j)\Big]\,. \tag{24}$$

Proof. By definition,

$$x_K \,\triangle\, \dots \,\triangle\, x_1(M) = \min_{1\le m_K<\dots<m_1=M} x_K(m_K) + x_{K-1}(m_{K-1}) - x_{K-1}(m_K) + \dots + x_1(m_1) - x_1(m_2)\,.$$

Let $(m_K^*,\dots,m_2^*)$ be a sequence of integers realizing the minimum in the above equation. Consider the set of integers

$$I = \{\sum_{i=1}^{k} |w_i| + \sum_{j=1}^{K-1} u(k+1,j),\ k\in\{0,\dots,K-1\}\,.$$

We may assume wlog that $m_K^* \in I$. Indeed assume that $m_K^* \notin I$, and let n_K be the smallest integer such that $n_K \in I, n_K > m_K^*$. Then, we clearly have $x_K(n_K) = x_K(m_K^*)$ and

$$\begin{aligned} x_K \,\triangle\, \dots \,\triangle\, x_1(M) - x_K(n_K) \;&\le\; \min_{n_K<m_{K-1}<\dots<m_2} x_{K-1}(m_{K-1}) - x_{K-1}(n_K) + \dots + x_1(M) - x_1(m_2) \\ &\le\; \min_{m_K^*<m_{K-1}<\dots<m_2} x_{K-1}(m_{K-1}) - x_{K-1}(m_K^*) + \dots + x_1(M) - x_1(m_2) \\ &=\; x_K \,\triangle\, \dots \,\triangle\, x_1(M) - x_K(m_K^*)\,. \end{aligned}$$

So we assume that $m_K^* \in I$. Let i_K be such that $m_K^* = \sum_{i=1}^{i_K-1}|w_i| + \sum_{j=1}^{K-1} u(i_K,j)$. We have $x_K(m_K^*) = \sum_{j=1}^{i_K-1} u(j,K)$. Then we prove with the same type of argument that we can assume wlog that $m_{K-1}^* \in \big\{\sum_{i=1}^{k}|w_i| + \sum_{j=1}^{K-2} u(k+1,j),\ k\in\{i_K,\dots,K-1\}\big\}$.

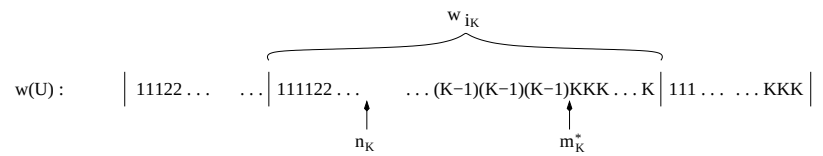

Figure 10: Illustration of the proof of Lemma 5.3

Since $m_K^* \in I$, we have $x_K(m_K^*) = \sum_{j=1}^{i_K} u(K,j)$ for some i_K. Then we prove with the same type of argument that we can assume wlog that $m_{K-1}^* \in \{\sum_{i=1}^{k}|w_i| + \sum_{j=1}^{K-2} u(k,j),\ k\in\{i_K+1,\dots,K-1\}$. It implies that $x_{K-1}(m_{K-1}) - x_{K-1}(m_K) = \sum_{j=i_K+1}^{i_{K-1}-1} u(K-1,j)$ for some i_{K-1}. By repeating the argument, we obtain the desired result. □

Proof of Theorem 5.2. The result for λ_1 is essentially due to Schensted. The left equality in (22) is (21) and the right equality appears for instance in [21, 37, 11]. As far as λ_K is concerned, Lemma 5.3 leads to

$$\lambda_K = x_K \triangle \ldots \triangle x_2 \triangle x_1(M) = \min_{\pi \in \widetilde{\Pi}} \Big[\sum_{(i,j)\in\pi} u(i,j) \Big] .$$

It remains to prove the right equality in (23). Let us denote by $r(i-1,j)$ the amount departing at time slot i from store j and $w(i,j)$ the quantity remaining at the end of time slot i in store j. According to (9), we have

$$r(n,k) = r(n-1,k-1) + w(n,k) - w(n+1,k) , \tag{25}$$

Applying (7) to our tandem model,

$$\begin{aligned} w(n+1,k) &= \big[w(n,k) + r(n-1,k-1) - u(n,K+1-k) \big]^+ \\ &= \max_{1\le m\le n} [\sum_{i=m}^{n-1} (r(i,k-1) - \sum_{i=m+1}^{n} u(i,K+1-k))] . \end{aligned}$$

Using (25), for $n \ge K$, we get

$$\begin{aligned} r(n,K) &= r(n-1,K-1) + w(n,K) - w(n+1,K) \\ &= r(n-2,K-2) + w(n-1,K-1) - w(n,K-1) + w(n,K) - w(n+1,K) \\ &= r(n-K+1,1) + \sum_{i=0}^{K-2} \big[w(n-i,K-i) - w(n-i+1,K-i) \big] \\ &= u(n-K+1,K) + \sum_{i=0}^{K-1} \big[w(n-i,K-i) - w(n-i+1,K-i) \big] . \end{aligned}$$

Hence,

$$\begin{aligned} R = \sum_{n=1}^{N} r(n,K) &= \sum_{n=K}^{N} r(n,K) \\ &= \sum_{n=K}^{N} u(n-K+1,K) + \sum_{n=K}^{N} \sum_{i=0}^{K-1} w(n-i,K-i) - w(n-i+1,K-i) \\ &= \sum_{n=1}^{N-K+1} u(n,K) + \sum_{i=0}^{K-1} \sum_{n=K}^{N} w(n-i,K-i) - w(n-i+1,K-i) \\ &= \sum_{n=1}^{N-K+1} u(n,K) + \sum_{i=0}^{K-1} w(K-i,K-i) - w(N-i+1,K-i) . \end{aligned}$$

Now recall that $w(k,k) = 0$, for all $k \ge 1$. We obtain

$$\begin{aligned} R &= \sum_{n=1}^{N-K+1} u(n,K) + \sum_{i=1}^{K} w(i,i) - \sum_{i=1}^{K} w(N-K+i+1,i) \\ &= \sum_{n=1}^{N-K+1} u(n,K) - \sum_{i=1}^{K} w(N-K+i+1,i) . \end{aligned} \tag{26}$$

Moreover, set $A = w(N+1,K) + w(N,K-1)$, we have

$$
\begin{aligned}
A &= \max_{1\le i_1\le N}\Big[\sum_{n=i_1}^{N-1} r(n,K-1) - \sum_{n=i_1+1}^{N} u(n,1)\Big] + w(N,K-1)\\
&= \max_{1\le i_1\le N}\Big[\sum_{n=i_1}^{N-1} [r(n-1,K-2) + w(n,K-1) - w(n+1,K-1)] - \sum_{n=i_1+1}^{N} u(n,1)\Big] + w(N,K-1)\\
&= \max_{1\le i_1\le N}\Big[\sum_{n=i_1-1}^{N-2} r(n,K-2) + w(i_1,K-1) - w(N,K-1) - \sum_{n=i_1+1}^{N} u(n,1)\Big] + w(N,K-1)\\
&= \max_{1\le i_1\le N}\Big[\sum_{n=i_1-1}^{N-2} r(n,K-2) + w(i_1,K-1) - \sum_{n=i_1+1}^{N} u(n,1)\Big]\\
&= \max_{1\le i_1\le N}\Big[\sum_{n=i_1-1}^{N-2} r(n,K-2) + \max_{1\le i_2\le i_1-1}[\sum_{n=i_2}^{i_1-2} r(n,K-2) - \sum_{n=i_2+1}^{i_1-1} u(n,2)\Big] - \sum_{n=i_1+1}^{N} u(n,1)]\\
&= \max_{1\le i_2<i_1\le N}\Big[\sum_{n=i_2}^{N-2} r(n,K-2) - \sum_{n=i_2+1}^{i_1-1} u(n,2) - \sum_{n=i_1+1}^{N} u(n,1)\Big].
\end{aligned}
$$

Applying this repeatedly leads to

$$
\sum_{i=1}^{K} w(N-K+i+1,i) = \max_{1\le i_{K-1}<\cdots<i_1\le N}\Big[\sum_{n=i_{K-1}}^{N-K+1} u(n,K) - \sum_{n=i_{K-1}+1}^{i_{K-2}-1} u(n,K-1) + \cdots + \sum_{n=i_1+1}^{N} u(n,1)\Big]. \tag{27}
$$

Combining (26) and (27), we obtain

$$
\begin{aligned}
R &= \min_{1\le i_{K-1}<\cdots<i_1\le N}\Big[\sum_{n=1}^{i_{K-1}-1} u(n,K) + \sum_{n=i_{K-1}+1}^{i_{K-2}-1} u(n,K-1) + \cdots + \sum_{n=i_1+1}^{N} u(n,1)\Big]\\
&= \min_{\pi\in\widetilde{\Pi}}\Big[\sum_{(i,j)\in\pi} u(i,j)\Big].
\end{aligned}
$$

□

Let us comment on Theorem 5.2. First, it is interesting to observe that the 6 quantities appearing in the equalities (22) and (23) all come from the same model. Now, consider the two identities (22) and (23) separately. The first one is classical and has led to several interesting developments for queues in series. A by-product of the newly proved identity (23) is therefore to pave the way to obtain analog results for stores in tandem. We set out this program in some details in the next Section.

5.4 Stochastic models, interchangeability, and hydrodynamic limits

If T is a tableau over the alphabet $\{1,\dots,k\}$, we write $x^T = \prod_{i=1}^k x_i^{\alpha_i}$, where α_i is the number of occurrences of the label i in the tableau. The *Schur function* s_λ associated with the partition $\lambda = (\lambda_1,\dots,\lambda_k)$ is defined by

$$
s_\lambda(x_1,\dots,x_k) = \sum_{T:\ \mathrm{sh}(T)=\lambda} x^T\,,
$$

where $\mathrm{sh}(T)$ is the shape of the tableau T. We refer the reader to the books [32, 34] for more about Schur functions and their connection to the RSK algorithm.

Suppose that the random variables $U = (u(i,j))_{(i,j)\in\{1,\dots,N\}\times\{1,\dots,K\}}$ are independent and that $u(i,j)$ is geometrically distributed with parameter q_j, for some fixed $q = (q_1,\dots,q_K) \in (0,1)^K$. Then [22] the law of the random partition $\lambda = (\lambda_1,\dots,\lambda_K)$ associated with $w(U)$ is given by

$$
P\{\lambda = l\} = a(q)^N s_l(q) s_l(1,\dots,1), \quad l \in P_K\,, \tag{28}
$$

where P_K is the set of integer partitions with at most K non-zero parts, $a(q) = \prod_j (1 - q_j)$ and s_l is the Schur function associated with the integer partition l. In particular, the law of λ is *symmetric* in the parameters $q_1, \ldots, q_K$.

This extends to the level of *processes*: if we write $\lambda = \lambda^{(N)}$ to emphasize its dependence on N, it is easy to check [22] that the random *sequence* $\lambda^{(N)}$ is a Markov chain in P_K with transition probabilities given by

$$P\{\lambda^{(N+1)} = l |\ \lambda^{(N)} = m\} = a(q) \frac{s_l(q)}{s_m(q)}\,, \tag{29}$$

for all m and l such that $l_1 \geq m_1 \geq l_2 \geq m_2 \geq \cdots$. In particular, we can see that the law of the sequence $\lambda^{(N)}$ is symmetric in the parameters $q_1, \ldots, q_K$.

These remarks give the law of the departure processes from both of the stochastic tandem queueing and storage models described above. Consider a saturated tandem of K queues/stores as defined in §5.1. Assume that the random variables $u(i,j), i \in \mathbb{N}^*, j \in \{1, \ldots, K\}$ are all independent, and that $u(i,j)$ is geometric of parameter $q_j \in (0,1)$. Let $D^{(N)}$ be the instant of departure of customer N from queue K and let $R^{(N)}$ be the cumulative departures over the time slots 1 to N in store K. By Theorem 5.2, $D^{(N)}$ and $R^{(N)}$ are, respectively, the lengths of the longest and shortest rows $\lambda_1^{(N)}$ and $\lambda_K^{(N)}$ of the random tableaux obtained by the RSK algorithm.

Interchangeability. A result by Weber [38] (see also Anantharam [1]) states that a series of ./M/1 queues arranged in tandem are interchangeable, i.e. the output process has the same distribution under any reordering of the queues. This is exactly the fact that the law of D is symmetric in the parameters and can be seen now as a direct consequence of the symmetry of the Schur polynomials. We can now extend this result as follows:

Theorem 5.4. *The law of $(D^{(N)}, R^{(N)})_N$ is unchanged if the parameters of the geometric laws are modified from $(q_1, \ldots, q_K)$ to $(q_{\sigma(1)}, \ldots, q_{\sigma(K)})$, for any permutation σ of $\{1, \ldots, K\}$.*

In particular, stores arranged in tandem are interchangeable.

Hydrodynamic limits. Proving 'shape theorems' (also known as 'hydrodynamic limits') and fluctuation results is an important issue in interacting particle systems, see for instance [16]. The law of $D^{(N)}$ is well-known and the connection with RSK has been exploited before to obtain detailed asymptotic and fluctuation results for the corresponding interacting particle system. The tandem storage (and bus-stop) model however has not been discussed in the literature. A by product of our analysis is that this model is amenable to some very precise analysis and development.

Indeed, consider Theorem 5.4 and the important special case when the parameters are equal, $q_j = q$ say. Then the random variables D and R are jointly distributed as the largest and smallest 'eigenvalues' of the Meixner ensemble. In the exponential case, replace Meixner by Laguerre. For example, if the $u(i,j)$ are exponentially distributed with parameter 1, then the cumulative departures $R^{(N)}$ over the time slots 1 to N in the corresponding tandem of K stores is distributed as the smallest eigenvalue of the $K \times K$ random matrix AA^*, where A is a $K \times N$ matrix with i.i.d. standard complex normal entries. This distribution is known, see for example [29]. When $N = K$ it is exponential with mean K. This can be used as a starting point for the analysis of the fluctuations and hydrodynamic limit corresponding to the bus-stop model of Section 5.2, a programme which is beyond the scope of this paper.

Fixed point results. Finally we remark that the identity $\left[\min_{\pi \in \widetilde{\Pi}} \sum_{(i,j)\in\pi} u(i,j) = R\right]$ can also be used as a cornerstone for proving fixed point and related weak convergence results for stores in tandem. In the context of queues in tandem, these questions have been pursued for some time and the identity $\left[\max_{\pi \in \Pi} \sum_{(i,j)\in\pi} u(i,j) = D\right]$ was a key ingredient in solving them [8, 2, 20, 27].

References

[1] V. Anantharam. Probabilistic proof of the interchangeability of /M/1 queues in series. *Queueing Systems Theory Appl.*, 2(4):387–392, 1987.

[2] F. Baccelli, A. Borovkov, and J. Mairesse. Asymptotic results on infinite tandem queueing networks. *Probab. Theory Relat. Fields*, 118(3):365–405, 2000.

[3] Y. Baryshnikov. GUEs and queues. *Probab. Theory Relat. Fields*, 119(2):256–274, 2001.

[4] A. Bedekar and M. Azizoglu. The information-theoretic capacity of discrete-time queues. *IEEE Trans. Info. Theory*, 44(2):446–461, 1998.

[5] P. Brémaud. *Point Processes and Queues-Martingale Dynamics*. Springer Verlag, Berlin, 1981.

[6] H. Bruneel and B.G. Kim. *Discrete-Time Models for Communication Systems Including ATM*. Kluwer Academic Publishers, Boston, 1993.

[7] P. Burke. The output of a queueing system. *Operations Research*, 4:699–704, 1956.

[8] C.S. Chang. On the input-output map of a G/G/1 queue. *J. Appl. Prob.*, 31(4):1128–1133, 1994.

[9] D. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes*. Springer Series in Statistics. Springer-Verlag, 1988.

[10] M. Draief, J. Mairesse, and N. O'Connell. Joint Burke's theorem and RSK representation for a queue and a store. In C. Banderier and C. Krattenthaler, editors, *Discrete Random Walks, DRW'03*, volume AC, pages 69–82. Discrete Mathematics and Theoretical Computer Science, 2003.

[11] P. Glynn and W. Whitt. Departures from many queues in series. *Ann. Appl. Probab.*, 1(4):546–572, 1991.

[12] B. Hambly, J.. Martin, and N. O'Connell. Pitman's 2M-X theorem for skip-free random walks with markovian increments. *Elect. Comm. in Probab.*, 6:73–77, 2001.

[13] J. Harrison and R. Williams. On the quasireversibility of a multiclass Brownian service station. *Ann. Probab.*, 18(3):1249–1268, 1990.

[14] J. Hsu and P. Burke. Behavior of tandem buffers with geometric input and Markovian output. *IEEE Trans. Comm.*, COM-24(3):358–361, 1976.

[15] F. Kelly. *Reversibility and Stochastic Networks*. Wiley, New-York, 1979.

[16] C. Kipnis and C. Landim. *Scaling limits of interacting particle systems*. Springer-Verlag, Berlin, 1999.

[17] D. Knuth. Permutations, matrices, and generalized Young tableaux. *Pac. J. Math.*, 34:709–727, 1970.

[18] W. König, N. O'Connell, and S. Roch. Non-colliding random walks, tandem queues, and discrete orthogonal polynomial ensembles. *Elect. J. Probab.*, 7:1–24, 2002.

[19] T. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*. Springer-Verlag, Berlin, 1999.

[20] J. Mairesse and B. Prabhakar. The existence of fixed points for the $\cdot/GI/1$ queue. *Ann. Probab.*, 31(4):2216–2236, 2003.

[21] E. Muth. The reversibility property of production lines. *Management science*, 25(2):152–158, 1979.

[22] N. O'Connell. Conditioned random walks and the RSK correspondence. *J. Phys. A*, 36:3049–3066, 2003.

[23] N. O'Connell. A path-transformation for random walks and the Robinson-Schensted correspondence. *Trans. Amer. Math. Soc.*, 355:3669–3697, 2003.

[24] N. O'Connell and M. Yor. Brownian analogues of Burke's theorem. *Stoch. Proc. Appl.*, 96(2):285–304, 2001.

[25] N. O'Connell and M. Yor. A representation for non-colliding random walks. *Elect. Comm. in Probab.*, 7:1–12, 2002.

[26] B. Prabhakar and R. Gallager. Entropy and the timing capacity of discrete queues. *IEEE Trans. Inform. Theory*, 49(2):357–370, 2003.

[27] B. Prabhakar. The attractiveness of the fixed points of a $\cdot/GI/1$ queue. *Ann. Probab.*, 31(4):2237–2269, 2003.

[28] N. Rajewsky, L. Santen, A. Schadschneider, and M. Schreckenberg. The asymmetric exclusion process: comparison of update procedures. *J. Statist. Phys.*, 92(1-2):151–194, 1998.

[29] T. Ratnarajah, R. Vaillancourt, and M. Alvo. Eigenvalues and condition numbers of complex random matrices. *SIAM J. Matrix Anal. Appl.*, 26(2):441–456, 2004/05.

[30] E. Reich. Waiting times when queues are in tandem. *Ann. Math. Stat.*, 28:527–530, 1957.

[31] P. Robert. *Stochastic networks and queues*, volume 52 of *Applications of Mathematics*. Springer-Verlag, Berlin, 2003.

[32] B. Sagan. *The symmetric group. Representations, combinatorial algorithms, and symmetric functions*. Number 203 in Graduate Texts in Mathematics. Springer-Verlag, 2001. 2nd edition.

[33] T. Seppäläinen. A variational coupling for a totally asymmetric exclusion process with long jumps but no passing. In *Hydrodynamic limits and related topics*, volume 27 of *Fields Inst. Commun.*, pages 117–130. Amer. Math. Soc., Providence, RI, 2000.

[34] R. Stanley. *Enumerative Combinatorics, Vol. 2*. Number 62 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999.

[35] W. Szczotka and F. Kelly. Asymptotic stationarity of queues in series and the heavy traffic approximation. *Annals of Probability*, 18:1232–1248, 1990.

[36] H. Takagi. *Queueing analysis: a foundation of performance evaluation. Vol. 3: Discrete-time systems*. North-Holland Publishing Co., Amsterdam, 1993.

[37] S. Tembe and R. Wolff. The optimal order of service in tandem queues. *Oper. Res.*, 24:824–832, 1974.

[38] R. Weber. The interchangeability of./M/1 queues in series. *J. Appl. Probab.*, 16(3):690–695, 1979.