Stability and weakly convergent approximations of queueing systems on a circle

Ronald Meester and Corrie Quant*

March 4, 1999

Department of Mathematics, University of Utrecht,
PO Box 80010, 3508 TA Utrecht, The Netherlands.
meester@math.uu.nl, quant@math.uu.nl

Abstract

We first consider a 'non-greedy' queueing system on a circle. We present a new and very simple proof of the stability of this system (under the appropriate condition) based on the average travel times between customers. Next we show that the same non-greedy system, with a restricted number of customers, converges weakly to this system when the restricted number goes to infinity. Finally we consider a polling network with finitely many service stations, in which the server has a 'greedy' service strategy. Under the appropriate condition, we give a new simple proof of the stability of this system.

KEYWORDS: queueing system, average travel time, stability, weak convergence, coupling.

*The research of the second author was financially supported by the Dutch organisation for scientific research NWO.

1 Introduction and results

Consider the following 'non-greedy' queueing systems:

(i) Customers arrive on the circumference of a circle with circumference one according to a Poisson process with parameter $\lambda > 0$ (we assume that the system is empty at time 0). Each customer chooses a waiting position on the circle uniformly, independently of the state of the system and of each other. A server is travelling clockwise along the circle at constant speed, and without loss of generality we assume that the server travels at speed one. When the server encounters a customer, he stops and serves that customer; after the service he continues his journey (in the same direction as before). Unless stated otherwise, we assume that the server does not travel in case there are no customers present on the circle, but this is of course arbitrary. Service times are i.i.d. with finite mean μ^{-1}. We call this system the *original system.*

(ii) A system that strongly resembles the first system, there is only one restriction: at most k customers are allowed on the circle. Customers who arrive at a moment that there are already k customers present on the circle, are sent away, and do not return. We call this system the *k-system.*

We shall first discuss two results concerning these systems. The first has to do with stability of the original system. We say that a system is *stable* if the expected length of a busy period is finite. In Section 2 we shall give a new and very simple, elementary proof of the following result, which was also obtained in [4].

Theorem 1.1 *The original system described above is stable if and only if* $\frac{\lambda}{\mu} < 1$.

Our proof is based on the average travel time between custumers, and our strategy will be as follows. Suppose that $\lambda/\mu < 1$. For the system to be unstable, the average travel time must be positive, otherwise we can essentially compare with an ordinary M/G/1 system. But if the average travel time is positive then there can be no accumulation of customers, and this essentially implies that the system must be stable.

To describe our next result, we have to say a few words about weak convergence of random counting measures. Identify the circle with the interval $[0, 1)$ in such a way that the server is always at position 0. Define X_t to be the random counting measure on $[0, 1)$, corresponding to the customers who are waiting (or being served) on the circle at time t in the original system. Similarly, let X_t^k be the random counting measure on $[0, 1)$, corresponding to the customers who are waiting (or being served) on the circle at time t in the k-system.

When $\lambda < \mu$, it follows from Theorem 1.1 that X_t is a regenerative process, with regeneration periods that have absolutely continuous distributions and finite expectations. Hence X_t converges in distribution to a limiting random counting measure X, when $t \to \infty$. Similarly, X_t^k converges in distribution to a limiting random counting measure X^k.

We shall prove the following result, which might appear obvious, but which seems surprisingly difficult to prove.

Theorem 1.2 *Let $\frac{\lambda}{\mu} < 1$. Then X^k converges weakly to X, when $k \to \infty$.*

A few words of explanation are appropriate here. In connection with weak convergence of random counting measures we recall Theorem 9.1.VI in [1] which says that weak convergence of random counting measures in the appropriate setting is equivalent to convergence of fidi distributions of continuity sets, i.e. sets whose boundary has probability zero to contain points under the limiting counting measure. This means that we need only show that the appropriate fidi distributions converge weakly.

The result is proved by making a coupling of the original system and the k-system, which is described in Section 3. In this coupling, we start in two empty systems and both systems behave identically until the moment that there are $k+1$ customers present in the original system. From that moment, the systems are not identical anymore, but in this coupling, it is the case that when the original system is empty, the k-system is also empty. So when the original system has been empty, both systems are identical for a while again, until in the original system, the level of $k+1$ customers is reached. We show that the stationary probability of being in a period where the level of the number of customers in the original system has been larger than k and since then has not achieved the zero level again, tends to zero as $k \to \infty$. This will suffice to prove the result.

In the last section we show that the idea of the stability proof as presented for the non-greedy queueing system is also applicable to a discrete 'greedy' system on a circle, in which the server always travels to the nearest customer. In fact, we hope that this idea is even the key to prove the stability of certain continuous greedy queueing systems, but until now we have not succeeded in proving so. As for Theorem 1.1, the following result was obtained earlier, this time in [2], but our proof is much simpler, and, as expressed above, hopefully can be generalised to continuous systems.

We now describe the greedy system. Consider a polling system with k waiting stations, which are numbered $1, \ldots, k$. The stations are located at equal distances on a circle with circumference 1, so the distance of a station to the two nearest stations is $\frac{1}{k}$. Each station has an infinite waiting capacity. Customers enter the system according to a Poisson process with parameter $\lambda > 0$. Each customer joins the queue at one of the stations, the choice of the station is independent of the current state of the system, each station has probability $\frac{1}{k}$ to be chosen. Service times are i.i.d. with expectation $\mu^{-1} > 0$. A server is travelling along the circle at constant speed, always in the direction of nearest nonempty station. Without loss of generality we assume that the server travels at speed 1. When he arrives at a station with waiting customers, he stops and serves all customers at this station until the station is empty. When the station is empty, he looks where the nearest nonempty station is and starts walking in that direction. In case that there are two nearest nonempty stations, he chooses one of them, each with equal probability. It is possible that the server changes direction during a walk due to an arrival of a new customer at a station which is nearer than the station to which the server was travelling originally. The server does not travel when no customer is present in the system.

Results on similar systems dealing with stability of polling systems with state dependent travelling strategies can be found in for instance [2], [3] and

[6]. We shall prove the following stability theorem:

Theorem 1.3 *The greedy polling system described above is stable if and only if* $\frac{\lambda}{\mu} < 1$.

The idea of the proof is the same as in the non-greedy queueing system described earlier. Suppose $\lambda < \mu$. For the system to be unstable, the average travel time must again be positive. But if the average travel time is positive, it is easy to show that it happens regularly that the server visits all stations in the order $(1, 2, \ldots, k, 1)$, i.e. the server makes a complete tour along all stations. This essentially implies that the number of customers on the circle can not become too large and this quickly leads to a stability proof.

2 Stability of the original system

We shall make the idea described in the introduction rigorous and we start with some notation. The number of customers that has arrived in the system until time t is denoted by $A(t)$, the length of the i^{th} service in the system by E_i. The amount of time used for serving until time t is denoted by $S(t)$, the amount of time used for travelling by $W(t)$, and $Z(t)$ denotes the amount of time until time t that the system was empty. Note that

$$S(t) + W(t) + Z(t) = t. \tag{1}$$

Finally, the travel time of the server between the $(i-1)^{th}$ service and the i^{th} service is denoted by W_i.

Lemma 2.1 *Let* $\frac{\lambda}{\mu} < 1$ *and suppose that the system is not stable. Then there exists* $\epsilon_0 > 0$ *such that with probability one,*

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} W_i > \epsilon_0.$$

Proof of Lemma 2.1 We shall prove the contrapositive. Fix some $0 < \delta < 1$, and let W_δ be the event that $\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} W_i < \delta$. Suppose that

$P(W_\delta) > 0$. The strong law of large numbers tells us that a.s. for t large enough we have

$$A(t) \leq t(\lambda + \delta). \tag{2}$$

Also with probability one, for n large enough we have

$$\frac{1}{n}\sum_{i=1}^{n} E_i \leq \frac{1}{\mu} + \delta. \tag{3}$$

Combining (2) and (3), we see that for some $K > 0$ (independent of δ), we have for t large enough,

$$S(t) \leq t\,(\lambda + \delta)\left(\frac{1}{\mu} + \delta\right) \leq t\left(\frac{\lambda}{\mu} + K\delta\right). \tag{4}$$

On the event W_δ, we know that for n large enough

$$\frac{1}{n}\sum_{i=1}^{n} W_i < 2\delta, \tag{5}$$

and since the number of travels up to time t is bounded above by the total number of customers arrived by time t, we conclude from (2), (4) and (5) that for some $K' > 0$ (again independent of δ), on the event W_δ we have with probability one for that for t large enough,

$$S(t) + W(t) \leq t\left(\frac{\lambda}{\mu} + K'\delta\right).$$

Now take δ so small that $\frac{\lambda}{\mu} + K'\delta < 1$. For these values of δ we have, using (1), that on W_δ,

$$\liminf_{t\to\infty} \frac{Z(t)}{t} > 0.$$

We conclude that if $P(W_\delta) > 0$ for *any* small enough δ, the empty state is positive recurrent with positive probability and therefore also positive recurrent almost surely. This is the contrapositive of what we wanted to prove and therefore we are done. □

Proof of Theorem 1.1 Clearly the system cannot be stable if $\frac{\lambda}{\mu} \geq 1$. Next suppose that $\frac{\lambda}{\mu} < 1$ and that the system is unstable. From Lemma 2.1 we obtain $\epsilon_0 > 0$ such that

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} W_i > \epsilon_0.$$

This implies that there exists a (random) sequence $L_1 < L_2 < \cdots$ such that for all j,

$$\frac{1}{L_j}\sum_{i=1}^{L_j} W_i > \frac{1}{2}\epsilon_0. \tag{6}$$

Denote the time at which the i^{th} service starts by T_i. Let us mark the position of the server at time 0 by $\star$. It follows from (6) that the number of times that the server has been in $\star$ until T_{L_j} is at least $\lfloor \frac{1}{2}\epsilon_0 L_j \rfloor$. Let M_i be the number of customers on the circle at the moment the server is in $\star$ for the i^{th} time. Since all these M_i customers must have been served by the time the server reaches $\star$ for the $(i+1)^{th}$ time, we see that

$$\sum_{i=1}^{\lfloor \frac{1}{2}\epsilon_0 L_j \rfloor - 1} M_i < L_j.$$

Hence there exists a constant $C > 0$ such that for all j we have

$$\frac{1}{\lfloor \frac{1}{2}\epsilon_0 L_j \rfloor - 1} \sum_{i=1}^{\lfloor \frac{1}{2}\epsilon_0 L_j \rfloor - 1} M_i < C. \tag{7}$$

From (7) it follows that the fraction of M_i's in the sum that are smaller than $2C$, is at least $1/2$. We conclude that there exists a positive constant D, such that for all j the following statement (A) is true:

(A) *the number of times before T_{L_j} that the server has been in $\star$ while at the same time the corresponding M_i is at most $2C$ is at least DL_j.*

Each time this happens, there is a uniform positive lower bound on the probability that all (at most) $2C$ customers are served before a new one arrives, and this lower bound does not depend on the past of the process. That is to say that there is another positive constant D' such that for all large j the following statement (B) is true:

(B) *the number of time intervals before T_{L_j} during which the system was empty is at least $D'L_j$.*

To complete the proof we observe that a.s. there exists a positive constant K such that

$$T_i \leq Ki, \tag{8}$$

for all i. This bound follows from the observation that $T_{i+1} - T_i$ is dominated by the sum of a service time, an interarrival time and 1 (the maximum travel time), all independent of each other, and independent for different values of i. Hence the number of time intervals until KL_j during which the system was empty is at least $D'L_j$, for j large, that is, a number *linear in time*. This implies that the expected time between two empty time intervals cannot be infinite and we are done. □

3 A coupling of the original system and the k-system

Let $\frac{\lambda}{\mu} < 1$. We construct a coupling of the k-system and the original system. In this coupling we assume that the servers do continue to travel when no customers are present on the circle. At time 0 both systems are empty. Customers arrive and depart in the original system as described in the first section. In the k-system, we let customers arrive at exactly the same moments as in the original system (of course some of them are sent away, because there are already k customers present in the system at their arrival). We call customers that arrive at the same time in both systems *corresponding customers*. The arrival location of the customer in the k-system is chosen such that at the moment of arrival, the distance between server and customer in the k-system equals the distance between the server and the corresponding customer in the original system. The service time of a customer in the k-system is equal to the service time of the corresponding customer in the original system. In this coupling, we denote the random counting measure on $[0,1)$ corresponding to the customers in the original system relative to the server by Y_t and the random counting measure on $[0,1)$ corresponding to the customers in the k-system relative to the server by Y_t^k. $(Y_t^k, Y_t)_{(t\geq 0)}$ is a coupling of $X^k_{t\,(t\geq 0)}$ and $X_{t(t\geq 0)}$. In this coupling, the following lemma holds, the proof of which is surprisingly lengthy.

Lemma 3.1 *In the coupling described above, $Y_t([0,1)) = 0$ implies that $Y_t^k([0,1)) = 0$, for all k.*

Proof The proof is by induction. Observe both systems from the first moment that a customer is sent away from the k-system until the next moment that the original system is empty again. During this time interval, let l denote the number of customers that arrived in the original system at a moment that there were less than k customers in the k-system (so the corresponding customer was not sent away in the k-system); and let n be the number of customers that arrived in the original system at a moment that that there were k customers present in the k-system (so the corresponding customer is sent away in the k-system). We call the latter customers *additional customers* since they are present in the original system, but have no corresponding customer in the k-system.

We use the following notation:

- $U(t)$ is the distance that the server has travelled in the original system until time t.
- $U_k(t)$ the distance that the server has travelled in the k-system until time t.
- T is the first moment at which the original system is empty again, after a customer was sent away from the k-system for the first time.
- S_i is the service time of the i^{th} additional customer.
- S_i^* is the service time of the i^{th} customer who arrives *and* takes place in the k-system. S_i^* is of course also the service time for the corresponding customer in the original system.
- $S_k(t)$ is the total time used for serving in the k-system, until time t.

We shall show that for all $t \leq T$

$$U_k(t) - U(t) \leq S_1 + \cdots + S_n. \tag{9}$$

We claim that (9) implies that the k-system is empty at time T. To see this, note that by definition, the original system is empty at time T, so

$$T = U(T) + S_1^* + S_2^* + \cdots + S_l^* + S_1 + S_2 + \cdots + S_n. \qquad (10)$$

Since $T = U_k(T) + S_k(T)$, we conclude from (10) that

$$U_k(T) - U(T) = S_1^* + S_2^* + \cdots + S_l^* + S_1 + S_2 + \cdots + S_n - S_k(T). \qquad (11)$$

From $U_k(T) - U(T) \leq S_1 + \cdots + S_n$ and (11) we find

$$S_1^* + S_2^* + \cdots + S_l^* + S_1 + S_2 + \cdots + S_n - S_k(T) \leq S_1 + \cdots + S_n. \qquad (12)$$

This implies that

$$S_k(T) \geq S_1^* + S_2^* + \cdots + S_l^*. \qquad (13)$$

Since the server in the k-system cannot have served more customers than the customers who have arrived in the k-system, we get that

$$S_k(T) \leq S_1^* + S_2^* + \cdots + S_l^*. \qquad (14)$$

From (13) and (14) we conclude that

$$S_k(T) = S_1^* + S_2^* + \cdots + S_l^*,$$

which implies that at time T all customers that have arrived in the k-system have been served, so that the k-system is empty at time T. It therefore suffices to prove (9).

We first prove that (9) holds for all realisations of the two systems in which $n = 1$ and $l = 0$. Next we prove that if we assume that (9) is true for all realisations of the coupling in which $n = 1$ and $l = q$, (9) must be true for all realisations of the coupling in which $n = 1$ and $l = q + 1$. Finally we show that if (9) is true for all realisations of the coupling in which $n \leq p$ and l is arbitrary, (9) is true for all realisations in which $n = p + 1$ and l is arbritrary.

In case $n = 1$ and $l = 0$, one customer has been sent away in the k-system and after that no other customers arrived until the original system was empty. To prove that $U_k(t) - U(t) \leq S_1$ we distinguish between two possibilities:

- The additional customer is the last customer served before the original system is empty. Until the server arrives at the additional customer both servers are at equal locations, the k-system is empty at the moment that the server starts the service of the additional customer. When the server serves the additional customer in the original system, the difference $U_k(t) - U(t)$ grows from zero to S_1, since during a service time of S_1 the server in the k-system travels a distance of S_1, while the server in the original system stands still. After the service of the additional customer, the original system is empty such that we are at time T. So (9) holds in this case.

- The additional customer has a waiting place between customers who arrived earlier. Until in the original system the additional customer gets served, both servers are at exactly the same location on the circles, so during that period, $U_k(t) - U(t) = 0$. During the service of the additional customer $U_k(t) - U(t) \leq S_1$, since the server in the original system does not move for a period of length S_1. After service of the additional customer $U(t) = U_k(t - S_1)$, so from that moment, $U_k(t) - U(t) \leq S_1$, since the difference $U_k(t) - U_k(t - S_1)$ is at most S_1 (this can happen if the server of the k-system travels the whole period from time $t - S_1$ until time t).

This proves (9) for all realisations of the coupling in which $n = 1$ and $l = 0$.

Next, suppose that, for all realisations of the coupling in which $n = 1$ and $l = q$, $U_k(t) - U(t) \leq S_1$, $\forall t \leq T$ (induction hypothesis). We want to prove that this implies that also $U_k(t) - U(t) \leq S_1$, $\forall t \leq T$ when we have a realisation of the coupling in which $n = 1$ and $l = q + 1$. We can prove this by looking at the last corresponding customers who arrived in both systems, before T. This is the $(q+1)^{st}$ customer that entered the systems after the additional customer arrived in the original system (that is why we shall call this customer the $(q+1)^{st}$ customer). As long as these corresponding customers are not served in both systems, these customers have no influence on the positions of the servers if we compare these positions to realisations of the systems in which this $(q+1)^{st}$

customer would never arrive in the systems. In such a realisation, the number of customers that arrived after the additional customer before T would be q, so we know by the induction hypothesis that as long as the services of these corresponding customers have not started, $U_k(t) - U(t) \leq S_1$.

Now we distinguish four cases for the position where the $(q+1)^{st}$ customer who arrived before T is situated on the circle with respect to the other customers in both systems:

1. The $(q+1)^{st}$ customer is the last customer served before time T, in both systems.
2. In both systems, the $(q+1)^{st}$ customer is not the last customer served before time T.
3. The $(q+1)^{st}$ customer is the last customer served before time T in the k-system, but is not the last customer served before time T in the original system.
4. The $(q+1)^{st}$ customer is the last customer served before time T in the original system, but is not the last customer served before time T in the k-system.

In the first case, since the induction hypothesis implies that if the $(q+1)^{st}$ customers would not have been present, the k-system would not empty later than the original system, the server in the k-system starts its journey to the $(q+1)^{st}$ customer no later than the server in the original system. Until the server in the k-system starts travelling to his $(q+1)^{st}$ customer, $U_k(t) - U(t) \leq S_1$. It is not possible that the difference $U_k(t) - U(t)$ exceeds the level S_1 before the server in the k-system arrives at his $(q+1)^{st}$ customer, since at the moment that the difference $U_k(t) - U(t)$ would equal S_1, the server in the k-system would have travelled S_1 more than the server of the original system and have served all customers in the k-system but the last one. During that period the server in the original system has also served all corresponding customers, and since he has served S_1 longer than the server in the k-system, he has also served the additional customer in his system. So if the difference $U_k(t) - U(t)$ would be equal to S_1 and the server in the k-system would be travelling, the server in the original system would be travelling too, so that the difference $U_k(t) - U(t)$ cannot grow. Observe that the server in the k-system arrives earlier at the last customer than the server in the original system arrives at the corresponding customer, since the distance which the server in the k-system has travelled at the moment that he reaches his last customer is at most S_1 larger than the distance which the server in the original system has travelled at the moment that he reaches his last customer, and in the original system the server required S_1 more time for serving his additional customer. After the server in the k-system has served the last customer the difference $U_k(t) - U(t)$ cannot get larger than S_1. Since if at a certain moment T^*, $U_k(T^*) - U(T^*)$ would be equal to S_1, we would have

$$T^* = S_1^* + \cdots + S_{q+1}^* + U_k(T^*)$$

because the server in the k-system has served all customers. Since $U_k(T^*) - U(T^*) = S_1$ we have:

$$T^* = S_1^* + \cdots + S_{q+1}^* + U(T^*) + S_1$$

so that at that moment the server in the original system has also served all customers.

In the second case, $U_k(t) - U(t) \leq S_1$ until one of the servers reached the $(q+1)^{st}$ customer, according to the induction hypothesis. When the $(q+1)^{st}$ customer is served earlier in the original system than in the k-system, the difference $U_k(t) - U(t)$ becomes larger than it would have been without the $(q+1)^{st}$ customer present. Until the $(q+1)^{st}$ customer gets served in the k-system, it is impossible that $U_k(t) - U(t) > S_1$. Since if $U_k(t) - U(t)$ would equal S_1, the server in the k-system must have arrived at the $(q+1)^{st}$ customer, because at the moment that the $(q+1)^{st}$ customers entered the systems, the server in the k-system had not travelled more than a distance S_1 extra compared to the server of the original system, according to the induction hypothesis. During the service of the $(q+1)^{st}$ customer in the k-system the difference can not get larger and after this service both servers have served the $(q+1)^{st}$ customer and the difference $U_k(t) - U(t)$ can not grow too large, since both servers have not moved for the same extra time S^*_{q+1} and the difference $U_k(t) - U(t)$ is what $U_k(t - S^*_{q+1}) - U(t - S^*_{q+1})$ would be in systems where the $(q+1)^{st}$ customers did never arrive, which is not larger than S_1 by the induction hypothesis. If the $(q+1)^{st}$ customer is served earlier in the k-system than in the original system, $U_k(t) - U(t)$ gets smaller than it would be in case the $(q+1)^{st}$ customer would not be present, when the $(q+1)^{st}$ customer is served in the original system the difference $U_k(t) - U(t)$ grows again, but it cannot grow larger than S_1. Since after the $(q+1)^{st}$ customer is served in the original system, both servers have standed still for the same (extra) time and continue as if the $(q+1)^{st}$ customers had never been present.

In the third case, as long as the $(q+1)^{st}$ customers are not served, $U_k(t) - U(t)$ can not get larger than S_1, according to the induction hypothesis. If the $(q+1)^{st}$ customer is served first in the original system, $U_k(t) - U(t)$ gets larger. As long as the $(q+1)^{st}$ customer is not served in the k-system, $U_k(t) - U(t)$ can not achieve the value S_1, since, according to the induction hypothesis, $U_k(t) - U(t) \leq S_1$ at the moment that the $(q+1)^{st}$ customers arrived. After the $(q+1)^{st}$ customer is served in the k-system the difference cannot get larger than S_1 either, since again, both servers have standed still for the same time. If the $(q+1)^{st}$ customer is served first in the k-system, the server in that system must have served all other customers. During the journey to the $(q+1)^{st}$ customer in the k-system, the difference $U_k(t) - U(t)$ can not grow larger than S_1. Suppose that would be the case, then the server in the original system could have served all customers in his system, if he had left out the service of the $(q+1)^{st}$ customer. But the $(q+1)^{st}$ customer is between the other customers, so he would have come in for his turn already. This contradicts the assumption that the $(q+1)^{st}$ customer is served earlier in the k-system than in the original system. Also when the server has finished the service of the $(q+1)^{st}$ customer $U_k(t) - U(t)$ cannot get larger than S_1. Suppose that at a certain moment T^* the difference would be equal to S_1. Then

$$T^* = S_1^* + \cdots + S_{q+1}^* + U(T^*) + S_1$$

so that at T^* all customers are served in the original system.

In the fourth case, it is impossible that the $(q+1)^{st}$ customer is served earlier in the original system than in the k-system. When the server in the original system is done with the other q customers and the additional customer, the server in the k-system could also have been, had he left out the service of the $(q+1)^{st}$ customer. Since at the moment that the $(q+1)^{st}$ customers arrived, $U_k(t) - U(t) \leq S_1$, the server in the k-system must arrive at the $(q+1)^{st}$ customer earlier than the server in the original system. As long as the server in the original system has not reached the $(q+1)^{st}$ customer, $U_k(t) - U(t) \leq S_1$ according to the induction hypothesis. $U_k(t) - U(t)$ can not grow larger than S_1 during the service of the last customer in the original system. If that would be the case, the server in the k-system would have served all customers, because the difference between the distances that the servers have travelled until they reach the last customer in their system is not larger than S_1, according to the induction hypothesis. So there would exist a time $T^* < T$ with

$$T^* = S_1^* + \cdots + S_{q+1}^* + U(T^*) + S_1.$$

So at T^* all customers in the original system would have been served, which contradicts the assumption that $T^* < T$.

Finally, we must show that if (9) holds for all realisations of the coupling in which $n \leq p$ and l is arbitrary, it holds for all realisations in which $n = p+1$ and l arbitrary. Observe that as long as the $(p+1)^{st}$ additional customer has not arrived yet, the difference $U_k(t) - U(t)$ does not grow larger than $S_1 + \cdots + S_p$ according to the induction hypothesis. Then look at the number of customers r that arrives after the $(p+1)^{st}$ additional customer in the original system. Inductively we can prove (in the same way as above) that for all $r \geq 0$: $U_k(t) - U(t) \leq S_1 + \cdots + S_{p+1}$. This proves Lemma 3.1. $\square$

4 Proof of Theorem 1.2

As mentioned before, to prove that the random counting measures X^k converge weakly to the random counting measure X it suffices to show that the fidi distributions converge weakly. That is to say that we have to prove that for all n and for all sets $D_1, D_2, \ldots, D_n$ with D_i an element of the Borel σ-algebra $\mathcal{B}$ and D_i a continuity set for X, the joint distributions of $(X^k(D_1), X^k(D_2), \ldots, X^k(D_n))$ converge weakly to the joint distribution of $(X(D_1), X(D_2), \ldots, X(D_n))$. Referring to the coupling in the previous section, it suffices to show that

$$\lim_{k\to\infty}\lim_{t\to\infty} P(Y_t^k(D_1) = k_1, \ldots, Y_t^k(D_n) = k_n) = \lim_{t\to\infty} P(Y_t(D_1) = k_1, \ldots, Y_t(D_n) = k_n)). \tag{15}$$

To prove (15) we introduce some further notation. Define

$$I_k(t) = \begin{cases} 1 & \text{if } \exists t^* < t : Y_{t^*}([0,1)) = k+1 \text{ and } \forall t' \in (t^*, t] : \; Y_{t'}([0,1)) \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that

$$\begin{aligned} &\lim_{k\to\infty}\lim_{t\to\infty} P(Y_t^k(D_1) = k_1, \ldots, Y_t^k(D_n) = k_n) = \\ &\quad \lim_{k\to\infty}\lim_{t\to\infty} P(Y_t^k(D_1) = k_1, \ldots, Y_t^k(D_n) = k_n, I_k(t) = 0) + \\ &\quad \lim_{k\to\infty}\lim_{t\to\infty} P(Y_t^k(D_1) = k_1, \ldots, Y_t^k(D_n) = k_n, I_k(t) = 1). \end{aligned} \tag{16}$$

Lemma 3.1 tells us that $I_k(t) = 0$ implies that $Y_t^k(D_i) = Y_t(D_i)$. Hence for all sets D_i, we can rewrite (16) as

$$\begin{aligned} &\lim_{k\to\infty}\lim_{t\to\infty} P(Y_t^k(D_1) = k_1, \ldots, Y_t^k(D_n) = k_n) = \\ &\quad \lim_{k\to\infty}\lim_{t\to\infty} P(Y_t(D_1) = k_1, \ldots, Y_t(D_n) = k_n, I_k(t) = 0) + \\ &\quad \lim_{k\to\infty}\lim_{t\to\infty} P(Y_t^k(D_1) = k_1, \ldots, Y_t^k(D_n) = k_n, I_k(t) = 1) \end{aligned}$$

from which we conclude that it suffices to prove that

$$\lim_{k\to\infty}\lim_{t\to\infty} P(I_k(t) = 0) = 1. \tag{17}$$

Define

$$A_t := \{t^* : t^* \leq t, \forall t' \in (t^*, t] : Y_{t'}([0,1)) \neq 0\}$$

and let

$$M(t) = \max_{t^* \in A_t} Y(t^*).$$

$M(t)$ is the maximum of the number of customers that has been in the original system since the last time before time t that the original system was empty. Since

$$I_k(t) = 1 \Leftrightarrow M(t) \geq k + 1,$$

we get

$$\lim_{k\to\infty} \lim_{t\to\infty} P(I_k(t) = 0) = \lim_{k\to\infty} \lim_{t\to\infty} P(M(t) \geq k+1) \tag{18}$$

and since $M(t)$ has a stationary distribution as $t \to \infty$, we conclude that

$$\lim_{k\to\infty} \lim_{t\to\infty} P(M(t) \geq k+1) = 0.$$

Together with (18), this proves (17), so we are done. □

5 Proof of Theorem 1.3

We start with the following elementary lemma.

Lemma 5.1 *Let $x_1, x_2, \ldots$ be a sequence with $0 \leq x_i \leq M$ for all i and let $\delta > 0$. Then*

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} x_i \geq \delta$$

implies that there exists a sequence $L_1 < L_2 < \cdots$ such that $\frac{1}{L_j} \sum_{i=1}^{L_j} \mathbf{1}_{\{x_i > 0\}} \geq \frac{\delta}{2M}$, $j = 1, 2, \ldots$.

Proof $\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} x_i \geq \delta$ implies that there exists a sequence $L_1 < L_2 < \cdots$ such that $\frac{1}{L_j} \sum_{i=1}^{L_j} x_i \geq \frac{\delta}{2}$. Since $0 \leq x_i \leq M$ for all x_i we can conclude that for all j, the fraction of the x_i's, $i \leq L_j$, which are positive must be at least $\frac{\delta}{2M}$. □

Proof of Theorem 1.3 If $\lambda \geq \mu$ it is obvious that the system is not stable (we can compare to a M/G/1 system again). So we have to prove that $\lambda < \mu$ implies stability of the system. The idea of the proof is very much the same as for the proof of Theorem 1.1, only the details are a bit trickier.

Suppose $\lambda < \mu$ and suppose that the system is not stable. Let W_i be the travel time of the server between the $(i-1)^{th}$ service and the i^{th} service in the system. Lemma 2.1 applies also to this system, giving an $\epsilon_0 > 0$ such that

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} W_i \geq \epsilon_0, \tag{19}$$

with probability 1.

Before we continue, it pays to indicate the difference between the proof for the non-greedy system as described in Section 2 and the current proof. In the non-greedy system we could conclude from (19) that the server travelled around the whole circle regularly, which implied that the number of customers on the circle could not get too large. In the current case the server can change direction and it is not immediately clear from (19) anymore that the server visits all stations regularly. We shall show that this is the case nevertheless, and once we have proved that, we can finish the proof in a same way as the proof of Theorem 1.1.

To prove that the server visits all stations regularly, we start proving that it happens regularly that the server starts a walk of positive length from (say) the first station. From the fact that with probability 1, $0 \leq W_i \leq \frac{1}{2}$, (19) and Lemma 5.1 we conclude that with probability 1, there exists a random sequence $L_1 < L_2 < \cdots$ such that for all j,

$$\frac{1}{L_j}\sum_{i=1}^{L_j} \mathbf{1}_{\{W_i>0\}} \geq \epsilon_0. \tag{20}$$

Define for $l = 1, \ldots, k$,

$$A_i^l = \begin{cases} 1 & \text{if } W_i > 0 \text{ and the } i^{th} \text{ walk starts at station } l, \\ 0 & \text{otherwise.} \end{cases}$$

Since there are only finitely many stations, we claim that there exist an l, a $\delta > 0$ and a subsequence $L_1' < L_2' < \cdots$ such that

$$\frac{1}{L_j'}\sum_{i=1}^{L_j'} A_i^l \geq \delta. \tag{21}$$

To see this, note that $\mathbf{1}_{\{W_i>0\}} = \sum_{l=1}^{k} A_i^l$. From (20) we find that

$$\frac{1}{L_j}\sum_{i=1}^{L_j}\sum_{l=1}^{k} A_i^l \geq \epsilon_0.$$

Interchanging the summation order yields for all j the existence of $l_0(j)$ such that

$$\frac{1}{L_j}\sum_{i=0}^{L_j} A_i^{l_0(j)} \geq \frac{\epsilon_0}{k}.$$

One of the l's must be equal to infinitely many $l_0(j)$'s and hence there exists an l_0, such that

$$\frac{1}{L_j}\sum_{i=1}^{L_j} A_i^{l_0} \geq \frac{\epsilon_0}{k} \quad \text{i.o.}$$

This proves (21) and without loss of generality we assume that (21) is the case for $l_0 = 1$.

Next we define B_i as follows: $B_i = 1$ if $A_i^1 = 1$ and in addition, the server does not return to station 1 before it has visited all other stations. In all other cases, B_i is defined to be 0. For instance, B_i is equal to 1 if $A_i^1 = 1$ and the server chooses stations $2, 3, \ldots, k$ (in that order) which is possible if we 'make sure' that customers are present at the appropriate stations at the appropriate time. This can be arranged by 'letting customers arrive' at certain places after the end of the $(i-1)^{th}$ service together with certain choices of the server about the next direction to go to. This makes it clear that the conditional probability that $B_i = 1$, given $A_i^1 = 1$ and the complete history of the process until the end of the $(i-1)^{th}$ service is uniformly bounded away from zero. This implies that for all large n we have, for some $\eta > 0$,

$$\eta \leq \frac{\sum_{i=1}^{n} B_i}{\sum_{i=1}^{n} A_i^1} \leq 1. \tag{22}$$

In particular, for j large enough, we have from (21) and (22) that

$$\frac{1}{L_j'}\sum_{i=1}^{L_j'} B_i \geq \delta\eta. \tag{23}$$

Next we define M_i to be the total number of customers present at stations $2, \ldots, k$ after the $(i-1)^{th}$ service if $A_i^1 = 1$; $M_i := 0$, otherwise. The remark above concerning the uniform lower bound on the condidional probability for B_i to be 1 implies that we also have for all C and n large,

$$\eta \leq \frac{\sum_{i=1}^{n} \mathbf{1}_{\{B_i=1\}} \mathbf{1}_{\{M_i>C\}}}{\sum_{i=1}^{n} \mathbf{1}_{\{A_i^1=1\}} \mathbf{1}_{\{M_i>C\}}} \leq 1. \tag{24}$$

Next we claim the following:

Statement $\star$ *There exists a C such that with probability one, for L'_j large enough we have*

$$\frac{\sum_{i=1}^{L'_j} \mathbf{1}_{\{A_i^1=1\}} \mathbf{1}_{\{M_i\leq C\}}}{\sum_{i=1}^{L'_j} \mathbf{1}_{\{A_i^1=1\}}} > \frac{1}{2}. \tag{25}$$

To see this, we assume that the converse of Statement $\star$ is true, and deduce a contradiction. This converse is the following statement: For all C, there is positive probability that there is a subsequence L'_{j_k} such that for all k,

$$\frac{\sum_{i=1}^{L'_{j_k}} \mathbf{1}_{\{A_i^1=1\}} \mathbf{1}_{\{M_i>C\}}}{\sum_{i=1}^{L'_{j_k}} \mathbf{1}_{\{A_i^1=1\}}} \geq \frac{1}{2}. \tag{26}$$

If (26) were true, (22) and (24) would give that for some positive β, the following statement is true:

Statement $\star\star$ *For all C, there is a positive probability that there is a subsequence L'_{j_k} such that for all k,*

$$\frac{\sum_{i=1}^{L'_{j_k}} \mathbf{1}_{\{B_i=1\}} \mathbf{1}_{\{M_i>C\}}}{\sum_{i=1}^{L'_{j_k}} \mathbf{1}_{\{B_i=1\}}} \geq \beta. \tag{27}$$

We claim that Statement $\star\star$ contradicts (23). To see this, just note that (23) implies that for large k, the number of indices $i \leq L'_{j_k}$ for which $B_i = 1$, is at least $\delta\eta L'_{j_k}$. Since all the M_i customers in Statement $\star\star$ are not at station 1, they can not contribute to $\sum_{j=1}^{L'_{j_k}} B_j$. Therefore Statement $\star\star$ tells us that the number of indices $i \leq L'_{j_k}$ for which $B_i = 1$ does *not* occur, is at least $(\delta\eta L'_{j_k} - 1)C\beta$. (We subtract one since after the last occurrence of $B_i = 1$ before L'_{j_k}, it is not clear that all M_i customers really count.) These two estimates are incompatible for large C.

Now we finish the argument as in the proof for the non-greedy system in Section 2. Statement $\star$ together with (21) yield that there exists a C such that for j large enough,

$$\frac{1}{L'_j}\sum_{i=1}^{L'_j} \mathbf{1}_{\{A_i^1=1\}}\mathbf{1}_{\{M_i\leq C\}} \geq \frac{\delta}{2}. \tag{28}$$

Let T_i be the time at which the i^{th} service starts. We can conclude from (28) that there exists a positive constant D such that the following statement (A') is true:

(A') *the number of times before $T_{L'_j}$ that the server has been at station 1 starting a walk of positive length, while at the same time the total number of customers at stations $2,\ldots,k$ was at most C is at least DL'_j.*

Each time this happens, there is a uniform positive lower bound on the probability that all (at most) C customers are served before a new one arrives, and this lower bound does not depend on the past of the process. So there is another positive constant D' such that for all j large enough the following statement (B') is true:

(B') *the number of time intervals before $T_{L'_j}$ during which the system was empty is at least $D'L'_j$.*

Observe that as in the non-greedy case there exists a constant K, such that $T_i \leq Ki$, since in this case, the difference $T_i - T_{i+1}$ is dominated by the sum of a service time, an interarrival time and $\frac{1}{2}$, the maximal travel time. These are all independent of i and each other. We can now finish our argument in the same way as in the non-greedy case. $\square$

Remark We would like to apply this idea to a continuous greedy system on a circle (i.e. a greedy system where the customers choose a waiting position on the circle uniformly instead of at a service station). The problem is that we are unable to show that the server travels along the whole circle regularly.

References

[1] D.J. Daley and D. Vere-Jones, An Introduction to the Theory of Point Processes, 1988.

[2] S. Foss, and G. Last, Stability of polling systems with exhaustive service policies and state-dependent routing, Annals of Applied Probability 6 (1996) 116-137.

[3] S. Foss and G. Last, On the stability of greedy polling systems with general service policies, Probability in the Engineering and Informational Sciences 12 (1998) 49-68.

[4] D.P. Kroese and V. Schmidt, A continuous polling system with general service times, The Annals of Applied Probability 2 (1992) 906-927.

[5] D.P. Kroese and V. Schmidt, Queueing Systems on a Circle, ZOR - Methods and Models of Operations Research 37 (1993) 303-331.

[6] R. Schassberger, Stability of polling networks with state-dependent server routing, Probability in the Engineering and Informational Sciences 9 (1995) 539-550.