Delay analysis for the fixed-cycle traffic-light queue


Document status and date: Published: 01/01/2006

Document Version: Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at: openaccess@tue.nl
providing details and we will investigate your claim.
DELAY ANALYSIS FOR
THE FIXED-CYCLE TRAFFIC-LIGHT QUEUE

J.S.H. VAN LEEUWAARDEN

Abstract. We consider the fixed-cycle traffic-light (FCTL) queue, where vehicles arrive at an intersection controlled by a traffic light and form a queue. The traffic light signal alternates between green and red periods, and delayed vehicles are assumed to depart during the green period at equal time intervals.

Most of the research done on the FCTL queue assumes that the vehicles arrive at the intersection according to a Poisson process and focuses on deriving formulas for the mean queue length at the end of green periods and the mean delay. For a class of discrete arrival processes, including the Poisson process, we derive the probability generating function of both the queue length and delay, from which the whole queue length and delay distribution can be obtained. This allows for the evaluation of performance characteristics other than the mean, such as the variance and percentiles of the distribution.

We discuss the numerical procedures that are required to obtain the performance characteristics, and give several numerical examples.

1. Introduction

The fixed-cycle traffic light (FCTL) queue is one of the best-studied models in traffic engineering. Vehicles arrive at an intersection controlled by a traffic light and form a queue. The traffic light signal alternates between green and red periods of effective durations $g$ and $r$, and delayed vehicles are assumed to depart during the green periods at equal time intervals.

The vast majority of the research devoted to the FCTL queue is based on the simplifying assumption that vehicles arrive at the traffic light according to a Poisson process. In this subfield, Webster’s formula [26] is recognized as the most famous result. It gives the mean delay of a vehicle in closed form, and is an expression that is partly based on simulation. It may not come as a surprise that many researchers felt challenged to come up with a full analytical solution for the average delay. McNeill [18] partially succeeded by providing the exact expression up to one unknown: the mean size of the overflow queue (the mean stationary queue length at the end of a green period). We denote this unknown by $EX_g$. McNeill came to this result by deriving the mean queue length within a cycle—where a cycle consists of a consecutive green period and red period, and is of $c = g + r$ length—and by using Little’s theorem. From McNeill’s formula, it became clear that providing an exact formula for the mean delay was equivalent with providing an exact formula for $EX_g$. Darroch [10] did this in 1964, using an approach that is both analytically and computationally involved.
This must be why Ohno [23] gives a detailed algorithmic description of the computation of Darroch’s rather complicated expression for $E X_g$. Ohno further presents a thorough overview of the research that uses the Poisson assumption. In particular, Ohno compares the various formulas for the mean delay: Webster [26], Webster & Cobbe [27], McNeill [18], Miller [20] and Newell [22]. The latter three were obtained by approximating $E X_g$. Ohno shows that the differences between the formulas are extremely small, and so he concludes that approximative formulas, in particular Miller’s and Newell’s, are the ones to be preferred, since these are easy to evaluate. Despite Ohno’s conclusion, there are still many questions left to be answered in at least two directions: (1) non-Poisson arrivals and (2) characteristics of the queue length and delay distribution other than the mean.

1.1. Non-Poisson arrivals. Most of the research is devoted to Poisson arrivals. Notable exceptions are Newell [22] who derives an approximation for $E X_g$ using fairly general arguments, and Darroch [10] and McNeill [18] who both consider a compound Poisson process. The assumption of (compound) Poisson arrivals allows the fixed-cycle traffic-light queue to be modelled at embedded points, namely the times just after the departure of a delayed vehicle, see Darroch [10]. We generalize this by assuming that the number of vehicles that arrive per time slot follow some discrete distribution (the Poisson and compound Poisson are also discrete distributions), and therewith model the queue in discrete-time. This allows one to consider distributions with a larger coefficient of variation, distributions with a finite support, or distributions fitted to empirical data. For this class of discrete arrival distributions, Van den Broek et al. [6] recently presented several bounds and approximations for $E X_g$.

1.2. Beyond the mean. By far, most of the research has been focused on deriving formulas for the mean of performance characteristics such as the queue length and delay. Meissl [19] and Darroch [10] derived, independently, the probability generating function (pgf) of the stationary overflow queue using almost identical methods. From this pgf, one could attempt to derive the pgf of the stationary delay, but to the best of our knowledge, no research has been done in that direction in the subsequent years. Most likely, the reason for this is that the pgf only allows for numerical evaluation, because it requires the determination of the roots of some characteristic function. The complexity of the solution, together with the potential difficulties of root-finding, prohibited the transform solutions from finding practical application. Consensus was reached that easier approximations were to be preferred (see e.g. [3, 22, 23] and the references therein). A noteworthy contribution is that of Heidemann [14]. For a Poisson arrival process, Heidemann derives, from the pgf of the queue length at arrival instants, the Laplace transform of the delay distribution. This method is often applied in queueing theory and relies heavily on the Poisson assumption. Heidemann recognizes that the method will not work for a more general arrival process.

Root-finding is no longer a difficulty. In Chaudhry et al. [8] every effort is made to dispel the scepticism towards such root-finding in queueing theory. It is demonstrated that the root-finding is well-structured, in the sense that the roots are distinct for most models and that their location is well-predictable, so that numerical problems are not likely to occur. This is also the case for the FCTL queue, and we return to the issue of root-finding in Sec. 5.1.

Once the pgf of the delay is known and completed with the numerically determined roots, a second issue is the inversion of the pgf, so as to obtain the probability distribution. This can be done analytically, by evaluating derivatives of the pgf, but this would give numerical problems. However, the numerical inversion of a pgf, as for the root-finding, has become a
more or less standard exercise due to the availability of efficient algorithms and increased computational power. A good reference is [1].

From the distribution of the queue length and delay, one can obtain, next to the mean, other characteristics such as percentiles, variability and the output process. Percentiles of the queue-length distribution, particularly at the end of the red period, are used for determining the length of turning lanes, so that the risk of a blockage in the through lanes can be quantified (see e.g. [28]). Also, from the driver’s perspective, percentiles of the delay distribution may express the probability that a driver experiences a delay longer than a certain threshold value.

The delay experienced by vehicles at a signalized intersection is usually subject to a large variation due to the randomness of the traffic arrival and the interruption by the traffic light. The variance of the delay is a critical component for the planning and design of signal controls, because it provides information on the confidence limits on the mean delay (see e.g. [13]). Also, Darroch’s solution to the FCTL queue leads to a third result of interest:

The output process of the FCTL queue is defined by the departures of vehicles during green periods. The output process is of importance, because it might determine the input for another signalized intersection in a network (see [25]). Moreover, at many intersections, the left-turning vehicles (at least in continental Europe and the United States) from the opposing stream are expected to filter through the gaps in the output process during the residual or unused green period. Hence, the output process determines the maximum filtration rate. Among other things, Darroch’s solution yields a full description of the output process. Tarko [25] derives approximations for the variance of the output process based on existing approximations for the mean overflow queue $\mathbb{E}X_g$. Tarko does not consider the exact solution of Darroch. Cowan [9] studies the maximum filtration rate in a model that is closely related to the FCTL queue.

1.3. Model limitations. The FCTL queue was one of the very first queueing models invoked for the analysis of delay and queue length at traffic signals. The pioneering work on the FCTL queue stems from the 1950s and 1960s, and there have been many developments since.

One limitation of the FCTL queue is that it focuses on the steady-state (equilibrium) behavior, which requires the mean arrival rate to be less than the capacity rate. Obviously, for many signals this is not true, especially in peak hours. Another limitation of the FCTL queue is that it requires the assumption of independent and identically distributed (i.i.d.) arrivals. While this assumption could very well be valid in the case of an isolated signal, an i.i.d. arrival process cannot incorporate the impact of adjacent signals and control. As a final limitation, we should mention that the FCTL queue has, by definition, a fixed-time control, and so it does not cover more sophisticated actuated and adaptive signal control systems.

The above limitations obviously have been observed before, and a huge research effort has led to other types of models that do capture nonequilibrium or oversaturated conditions, time-dependency, coordinated signals, and actuated or adaptive signal control systems. For a good overview we refer to Rouphail et al. [24].

One could argue that the FCTL queue is obsolete, but, evidently, we tend to disagree. The FCTL queue is classic textbook material and distinguishes itself from most modern models in the sense that it allows for a detailed and illustrative exact analysis. The FCTL queue fulfills a prominent role in the literature on queueing models for traffic signals and is referred to by many researchers to this day.
1.4. Our contribution. Under certain assumptions (see Sec. 2), Darroch [10] derives for a compound Poisson arrival process the pgf of the length of the overflow queue. We first show in Sec. 3 how the same result holds for a more general discrete arrival process, and how one can obtain from the pgf of the overflow queue the pgf of the queue length throughout the whole cycle. Next, we derive in Sec. 4 the pgf of the delay. The transform solutions to the queue length and delay can be inverted by means of the procedures discussed in Sec. 5, which results in the probability distributions of the queue length and delay. From the probability distribution we can obtain performance characteristics such as the variance and percentiles of the distribution.

For Poisson arrivals, approximations of the variance of the delay have been derived (see [13] and the references therein), and the percentiles of the queue length and delay distribution have been estimated in [28] using regression analysis. To the best of our knowledge, for both performance characteristics, no explicit description has appeared in the literature.

Our approach leads to an explicit solution of the pgf of the queue length and delay in the FCTL queue for a general arrival distribution, and provides as such a broad framework. For practical application, though, some numerical work is necessary. The issues of root-finding and inversion of a pgf are addressed in Sec. 5. As it turns out, both issues are rather straightforward. We provide numerical examples in Sec. 6, followed by some conclusions in Sec. 7.

2. Basic assumptions

In most studies on the FCTL queue that do not rely on the Poisson assumption, e.g. [5, 10, 19, 21], the following two assumptions are made:

**Assumption 1.** (discrete-time assumption) The time axis is divided into constant time intervals of unit length, so-called slots, where each slot corresponds to the time needed for a delayed vehicle to depart from the queue. The green and red periods, and thus the cycle time $c$ (with $c = g + r$), are assumed to be fixed multiples of one slot. Hence, $g$ and $r$ are integers expressed in slots. Those vehicles that arrive at the queue and are delayed, join the queue at the end of the slot in which they arrive.

**Assumption 2.** (independence assumption) Let $Y_{k,n}$ denote the number of vehicles that arrive at the intersection during slot $k$ in cycle $n$. The random variables $Y_{k,n}$ are assumed to be independent and identically distributed (i.i.d.) for all $k, n$, according to some discrete random variable $Y$ with pgf $Y(z)$.

These two assumptions allow one to model the queue length at the end of time slots as a discrete-time Markov chain. We also work under these assumptions and assess their implications at several places in this paper. Note that a Poisson arrival process satisfies the independence assumption. Furthermore, although it might seem rather restrictive, the independence assumption is frequently made and allows for a far larger class of arrival distributions than just the Poisson case.

The following assumption is always made for the FCTL queue:

**Assumption 3.** (FCTL assumption) For those cycles in which the queue clears before the green period terminates, all vehicles that arrive during the residual green period pass through the system and experience no delay whatsoever.

The FCTL assumption is legitimate in the sense that the vehicles that arrive during the residual green period can pass the intersection without slowing down, and therefore
the discharge rate of these vehicles is much higher than the discharge rate of the delayed vehicles (one per time unit). However, the FCTL assumption does lead to the anomaly that some of these vehicles would be delayed for at least the duration of the red phase if any reasonable constraint upon the spacing of departures were imposed. This means that, due to the FCTL assumption, the delay for a small proportion of vehicles is understated.

However, because of the huge difference in discharge rates of delayed vehicles (these vehicles have to accelerate) and those vehicles covered by the FCTL queue, we think it nevertheless is a reasonable assumption.

The FCTL assumption does, however, have some severe consequences for the analysis of the queue length. Let $X_{k,n}$ denote the queue length at time $k$ in cycle $n$ (time expressed in slots). Then, in cycle $n$, $X_{0,n}$ is the queue length at the beginning of the green period, and $X_{g,n}$ the overflow defined as the queue length at the end of the green period (and thus the beginning of the red period). Let $A_n$ denote the total number of vehicles that arrive at the intersection in between the two measurements of the overflow $X_{g,n}$ and $X_{g,n+1}$. Thus $A_n$ are the arrivals from $X_{g,n}$ onwards in a consecutive red and green period, and $A_n = A_n^d + A_n^p$, where $A_n^d$ denotes the delayed vehicles and $A_n^p$ those vehicles that pass without delay on behalf of the FCTL assumption.

The overflow queue can then be defined as

$$X_{g,n+1} = \max\{X_{g,n} + A_n^d - g, 0\}.$$  

The fact that $A_n^d$ depends on both $X_{g,n}$ and the exact specification of when the arrivals occur makes (2.1) hard to analyze. The analysis could be simplified if all vehicles were delayed, so that all vehicles arrive while the queue length is at least one, and $A_n = A_n^d$. The variables $A_n^d$ are then i.i.d. and (2.1) reduces to the classical bulk service queue, first solved by Bailey [4]. He derived the pgf of the stationary queue length $X_g$, defined as $\lim_{n \to \infty} X_{g,n}$, that exists if $\mathbb{E}A < g$. The pgf requires the determination of $g$ (complex-valued) roots on and within the unit circle of some characteristic equation.

Beckmann et al. [5] and Newell [21] assumed that $Y(z) = 1 - \alpha + \alpha z$. On this assumption of zero or one arrivals per time slot, $A_n = A_n^d$, and Bailey’s solution can be applied to derive the exact value of $\mathbb{E}X_g$. When $Y$ can take values larger than one, the bulk service queue is obviously an approximation and yields an upper bound on the overflow queue. For a compound Poisson process, McNeill [18] used Bailey’s solution to derive an upper bound on $\mathbb{E}X_g$. Although one would want an upper bound to be easy to compute, McNeill’s upper bound is not, because it still requires the numerical calculation of the $g$ roots.

Darroch [10] derived, under the discrete-time assumption, the solution to the true FCTL queue that is of the same complexity as Bailey’s solution to the bulk service queue, again requiring the roots of a characteristic equation. The effort put into determining the roots is in Darroch’s case more justifiable, though, because it leads to the exact solution.

3. Darroch’s solution

We now present Darroch’s solution to the FCTL queue. Darroch allows for a slightly more general departure process of the delayed vehicles, which we omit for reasons of clarity.

Let $Y$ be a discrete random variable with pgf $Y(z) = \mathbb{E}(z^Y)$ (Darroch assumes $Y$ to be of compound Poisson type), and assume all moments of $Y$ exist. We denote the mean and variance of $Y$ by $\mu_Y$ and $\sigma_Y^2$, respectively. Clearly, to have stability, it is required that the number of arriving vehicles is less than the maximum number of vehicles that can depart, and hence $Y$ should satisfy

$$c \mu_Y < g.$$
The following recursive relation holds for $k = 0, 1, \ldots, g - 1$:

\[ X_{k+1,n} = \begin{cases} X_{k,n} + Y_{k+1,n} - 1, & X_{k,n} \geq 1, \\ 0, & X_{k,n} = 0, \end{cases} \]

and $X_{k+1,n} = X_{k,n} + Y_{k+1,n}$ for $k = g, g + 1, \ldots, c - 1$. Hence, for $k = 0, 1, \ldots, g - 1$ we have

\[ P(X_{k+1,n} = j) = \sum_{p=1}^{j+1} P(X_{k,n} = p)P(Y_{k+1,n} = j - p + 1), \quad j = 1, 2, \ldots \]

\[ P(X_{k+1,n} = 0) = P(X_{k,n} = 0) + P(X_{k,n} = 1)P(Y_{k+1,n} = 0). \]

Let $X_{k,n}(z)$ denote the pgf of $X_{k,n}$. It then follows from (3.3-3.4) that

\[ X_{k+1,n}(z) = P(X_{k,n} = 0) + \sum_{p=1}^{\infty} \sum_{j=p-1}^{\infty} P(X_{k,n} = p)P(Y_{k+1,n} = j - p + 1)z^j. \]

Some rewriting yields

\[ X_{k+1,n}(z) = z^{-1}Y(z)X_{k,n}(z) + (1 - z^{-1}Y(z))P(X_{k,n} = 0). \]

Hence, it follows from (3.6) that

\[ X_{g,n}(z) = (z^{-1}Y(z))^gX_{0,n}(z) + (1 - z^{-1}Y(z)) \sum_{k=0}^{g-1} P(X_{k,n} = 0)(z^{-1}Y(z))^{g-k-1}. \]

Note that $X_{0,n+1}(z) = X_{g,n}(z)Y(z)$ and in equilibrium $X_{0,n+1}(z) = X_{0,n}(z)$. Denote by $X_k$ the stationary distribution of $X_{k,n}$. After some rewriting, we then obtain from (3.7)

\[ X_g(z) = \frac{Y(z)^g(\zeta(z) - 1)\sum_{k=0}^{g-1} q_k \zeta(z)^k}{z^g - Y(z)^c}, \]

where $q_k = P(X_k = 0)$ and $\zeta(z) = z/Y(z)$. This expression still contains $g$ unknowns $q_0, \ldots, q_{g-1}$, which can be found using the following classical approach (see e.g. Bailey [4], Darroch [10]). With Rouché’s theorem, it can be shown that the denominator of (3.8) has $g$ zeros on or within the unit circle $|z| \leq 1$. Because a pgf is analytic and well-defined in $|z| \leq 1$, the numerator of $X_g(z)$ should vanish at each of the zeros. This gives $g$ equations. One of the zeros equals 1, and leads to a trivial equation. However, the normalization condition $X_g(1) = 1$ provides an additional equation. Using l’Hôpital’s rule, this condition is found to be

\[ \sum_{k=0}^{g-1} q_k = \frac{g - c\mu_Y}{1 - \mu_Y} =: \eta, \]

which can be written as

\[ g - \sum_{k=1}^{g-1} q_k = \left( c - \sum_{k=1}^{g-1} q_k \right)\mu_Y. \]

The right-hand side of (3.10) represents the mean number of delayed vehicles that arrive per cycle $\mathbb{E}A^d$, see (2.1). The left-hand side of (3.10) represents the mean number of slots per green period used for the departure of delayed vehicles.
Denote the $g$ roots of $z^g = Y(z)^c$ on and within the unit circle by $z_0 = 1, z_1, \ldots, z_{g-1}$. The $g$ unknowns $q_0, \ldots, q_{g-1}$ then follow from solving the set of linear equations

\begin{equation}
\begin{pmatrix}
1 & 1 & 1 & \ldots & 1 \\
1 & \zeta(z_1) & \zeta(z_1)^2 & \ldots & \zeta(z_1)^{g-1} \\
1 & \zeta(z_2) & \zeta(z_2)^2 & \ldots & \zeta(z_2)^{g-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & \zeta(z_{g-1}) & \zeta(z_{g-1})^2 & \ldots & \zeta(z_{g-1})^{g-1}
\end{pmatrix}
\begin{pmatrix}
q_0 \\
q_1 \\
q_2 \\
\vdots \\
q_{g-1}
\end{pmatrix}
= \begin{pmatrix}
\eta \\
0 \\
0 \\
\vdots \\
0
\end{pmatrix}.
\end{equation}

(3.11)

The above system can be solved by applying Cramer’s rule. The system can then be transformed into a Vandermonde system, leading to the following explicit solution for $q_0, \ldots, q_{g-1}$ (with $\tau_k = \zeta(z_k)$):

\begin{equation}
q_j = \eta(-1)^j \frac{1}{\prod_{k=1}^{g-1}(\tau_k - 1)} \sum_{1 \leq i_1 < i_2 < \cdots < i_{g-1-j} \leq g-1} \tau_{i_1} \tau_{i_2} \cdots \tau_{i_{g-1-j}}.
\end{equation}

(3.12)

Note that the probabilities $q_0, \ldots, q_{g-1}$ describe the output process of the FCTL queue. The mean stationary overflow queue follows from $\frac{d}{dz} X_g(z)|_{z=1}$, i.e.,

\begin{equation}
E X_g = \frac{c \sigma_Y^2 + r^2 \mu_Y^2 - g^2 (1 - \mu_Y)^2}{2(g - c \mu_Y)} - \frac{\sigma_Y^2}{2(1 - \mu_Y)} + \frac{1 - \mu_Y}{g - c \mu_Y} \sum_{k=0}^{g-1} k q_k.
\end{equation}

(3.13)

3.1. Mean stationary delay. In the sequel we will only work with stationary variables. From the pgf of the overflow queue, the pgf of the queue length at every other time point can be determined. In a way similar to (3.7), we find for $k = 1, 2, \ldots, g$

\begin{equation}
X_k(z) = X_g(z) Y(z)^r (Y(z) z^{-1})^k + (1 - Y(z) z^{-1}) \sum_{i=0}^{k-1} \tau_i (Y(z) z^{-1})^{k-i-1}.
\end{equation}

(3.14)

For $k = g + 1, \ldots, c - 1$ we have

\begin{equation}
X_k(z) = X_g(z) Y(z)^{k-g}.
\end{equation}

(3.15)

The mean queue length at the beginning of an arbitrary slot is given by

\begin{equation}
E \bar{X} = \frac{1}{c} \sum_{k=0}^{c-1} E X_k.
\end{equation}

(3.16)

From (3.16) the mean delay can be determined, where we define delay as:

**Definition 1.** (delay) The delay of a vehicle, denoted by $D$, is defined as the number of slots from the beginning of the first slot after the slot in which the vehicle arrives, until the end of the slot in which the delayed vehicle departs from the queue.

From Little’s theorem it then follows that the mean delay of a vehicle is given by $E \bar{X} / (c \mu_Y)$, which can be shown to be

\begin{equation}
ED = \frac{r}{2c \mu_Y (1 - \mu_Y)} \left[ \frac{\sigma_Y^2}{1 - \mu_Y} + r \mu_Y + 2E X_g \right]
\end{equation}

(3.17)
3.2. On the definition of delay. Delay in the FCTL queue can be defined in several ways, depending on how one handles the delay of a vehicle experienced within the slot of its arrival. We use Definition 1, which does not include this part of the delay; we assume, as does Darroch [10], that the vehicle joins the queue at the beginning of the next slot after its arrival. This is in line with the discrete-time assumption, where we assume that the arrival of vehicles only occurs at the end of a time slot; the vehicles then arrive as a batch. In reality, the vehicles arrive one-by-one, each vehicle arriving at some different time point during the slot, and the total delay $D_T$ satisfies $D_T = D + D_R$ where $D$ defined as in Definition 1, and $D_R$ the residual delay within the slot of arrival, $D_R \in [0, 1]$.

McNeill [18] does include $D_R$. In comparison with (3.17), McNeill’s expression has an additional term $r/(2c\mu_Y(1 - \mu_Y))$, which can be easily shown to be $ED^R$. For Poisson arrivals, Ohno [23] gives a comparison between McNeill’s expression, Darroch’s expression (3.17) (where Ohno’s formula (17) for Darroch’s mean delay formula is incorrect), and several other approximations. Ohno shows that the differences are only marginal.

To be consistent with the discrete-time assumption, we do not include $D_R$ in the delay analysis that is presented in Sec. 4, although the results derived for $D$ can be easily extended to $D_T$.

4. Full delay analysis

Although in a different context (buffer management in cable access networks), in van Leeuwaarden et al. [17] the pgf of the vehicle delay was derived for the discrete bulk service queue. In Sec. 2 we have stressed that the crucial difference between the FCTL queue and the discrete bulk service queue is the FCTL assumption. In this section we adapt the approach taken in [17] as to derive the pgf of the vehicle delay in the FCTL queue.

We tag a vehicle that arrives in time slot $m$ during the green period, and introduce $U_{[m]}^g$ as the number of vehicles that depart before the tagged vehicle, counted from the beginning of the green period and given that the vehicle is delayed. It then holds that

\[(4.1) \quad U_{[m]}^g = X_{m-1}X_{m-1} > 0 + Z_1 + m - 1,\]

where $X_{m-1} \mid X_{m-1} > 0$ denotes the vehicles present at the end of time slot $m - 1$ (cannot be zero because the tagged vehicle is delayed), $Z_1$ the number of vehicles that arrive in the same time slot as the tagged vehicle but before it, and $m - 1$ vehicles because we start counting from the beginning of the green period. The pgf of $Z_1$ is given by (see Bruneel & Kim [7], p.20)

\[(4.2) \quad Z_1(z) = \frac{1 - Y(z)}{(1 - z)\mu_Y},\]

and the pgf of $U_{[m]}^g$ thus satisfies (since $Z_1$ and $X_{m-1}$ are independent)

\[(4.3) \quad U_{[m]}^g(z) = \left(\frac{X_{m-1}(z) - q_{m-1}}{1 - q_{m-1}}\right) \left(\frac{1 - Y(z)}{(1 - z)\mu_Y}\right) z^{m-1}.\]

We express $U_{[m]}^g$ in terms of two integer random variables $F_{[m]}$ and $R_{[m]}$:

\[(4.4) \quad U_{[m]}^g = gF_{[m]} + R_{[m]}, \quad F_{[m]} \geq 0, \quad 0 \leq R_{[m]} \leq g - 1,\]

where $F_{[m]}$ denotes the number of complete cycles enclosed in the tagged vehicle’s delay, and $R_{[m]}$ the number of vehicles that will depart during the same green period as the
tagged vehicle, but before it. Let $D_{[m]}$ denote the delay of the tagged vehicle. It holds (for
$m = 1, \ldots, g$

\begin{equation}
D_{[m]} = \begin{cases}
cF_{[m]} + R_{[m]} - m + 1, & \text{w.p. } 1 - q_{m-1}, \\
0, & \text{w.p. } q_{m-1}.
\end{cases}
\end{equation}

The delay of a delayed tagged vehicle consists of $F_{[m]}$ cycles, $R_{[m]}$ time slots in the green
period of departure, minus $m$ time slots since we started counting at the beginning of a
green period, plus one time slot that the tagged vehicle itself needs to depart. The pgf of
$D_{[m]}$ then reads (for $m = 1, \ldots, g$

\begin{equation}
D_{[m]}(z) = \sum_{i=0}^{\infty} \mathbb{P}(D_{[m]} = i) z^i
= q_{m-1} + (1 - q_{m-1}) z^{-m+1} \sum_{j=0}^{\infty} \sum_{k=0}^{g-1} \mathbb{P}(F_{[m]} = j, R_{[m]} = k) z^{c_j+k}
\end{equation}

From (4.6) it follows that (for $m = 1, \ldots, g$

\begin{equation}
D_{[m]}(z^g) = q_{m-1} + (1 - q_{m-1}) z^{-(m-1)g} \sum_{k=0}^{g-1} z^{gk} \vartheta_{mk}(z),
\end{equation}

where the functions $\vartheta_{mk}(z)$ are defined as

\begin{equation}
\vartheta_{mk}(z) = \sum_{j=0}^{\infty} \mathbb{P}(U_{[m]}^g = gj+k) z^{gcj}.
\end{equation}

The problem now is that (4.8) does not allow a direct substitution of the pgf of $U_{[m]}$. To
circumvent this, we use a basic approach that can be found in e.g. Bruneel & Kim [7]
or Kang & Steyaert [16], and which has also been applied in van Leeuwaarden et al. [17].
Substituting $l = gj+k$ in (4.8) yields

\begin{equation}
\vartheta_{mk}(z) = \sum_{l=0}^{\infty} \mathbb{P}(U_{[m]}^g = l) z^{(l-k)c} \sum_{j=\infty}^{\infty} \delta(l - gj - k),
\end{equation}

with $\delta(n)$ the Kronecker delta function, which equals 1 for $n = 0$ and 0 for all other $n$.

**Property 1.**

\begin{equation}
\frac{1}{g} \sum_{l=0}^{g-1} a^m = \sum_{j=-\infty}^{\infty} \delta(m - jg),
\end{equation}

where $a = \exp(2\pi i/g)$, $i = \sqrt{-1}$, and $m$ and $g$ integer values. The sum on the left-hand
side is zero unless $m$ is a multiple of $g$.

Using Property 1 we obtain

\begin{equation}
\vartheta_{mk}(z) = \sum_{l=0}^{\infty} \mathbb{P}(U_{[m]}^g = l) z^{(l-k)c} \frac{1}{g} \sum_{l=0}^{g-1} a^{(l-k)}
\end{equation}
Then get (for \(m\) the time slot in which the tagged vehicle arrives (the vehicle is by definition delayed). We as the number of vehicles that depart before the tagged vehicle, counted from the end of tedious calculations, to be equal to (3.17).

\[
U(z) = \sum_{i=0}^{\infty} a^{-i} P(U[m] = i) z^i a^i.
\]

(4.11)

Substituting (4.11) into (4.7) yields (for \(m = 1, \ldots, g\))

\[
D[m](z) = q_{m-1} + (1 - q_{m-1})z^{-(m-1)g} \sum_{k=0}^{g-1} z^k \sum_{i=0}^{k} a^{-i} U[m] z^i a^i.
\]

(4.12)

Now tag a vehicle that arrives in time slot \(m\) during the red period, and introduce \(U^r[m]\) as the number of vehicles that depart before the tagged vehicle, counted from the end of the time slot in which the tagged vehicle arrives (the vehicle is by definition delayed). We then get (for \(m = g + 1, \ldots, c\))

\[
U^r[m] = X_g + \sum_{i=1}^{m-g-1} Y_i + Z_1, \quad Y_i \sim Y \text{ i.i.d.}
\]

(4.13)

That is, \(U^r[m]\) consists of the overflow queue, the vehicles that arrives in the first \(m - g - 1\) time slots of the red period and \(Z_1\). Because these elements are independent, the pgf of \(U^r[m]\) thus satisfies (for \(m = g + 1, \ldots, c\))

\[
U^r[m](z) = X_g(z)Y(z)^{m-g-1} \frac{1 - Y(z)}{(1 - z)\mu_Y}.
\]

(4.14)

The delay of the tagged vehicle consists of (for \(m = g + 1, \ldots, c\))

\[
D[m] = c - m + cF[m] + R[m] + 1,
\]

(4.15)

that is \(c - m\) time slots till the beginning of the next green period, \(F[m]\) cycles, \(R[m]\) time slots of those vehicles that depart within the same green period but before the tagged vehicle, and one time slot that the tagged vehicle needs itself to depart. Using the same approach as for deriving (4.12) we obtain (for \(m = g + 1, \ldots, c\))

\[
D[m](z) = z^{(c-m+1)g} \sum_{i=0}^{g-1} U^r[m](a^i z^i) \frac{1 - (z^{-r} a^{-i})^g}{1 - z^{-r} a^{-i}}.
\]

(4.16)

From (4.12) and (4.16) we obtain the pgf of the vehicle delay as

\[
D(z) = \frac{1}{c} \sum_{m=1}^{c} D[m](z).
\]

(4.17)

The mean delay of a vehicle follows from \(\mathbb{E}D = \frac{1}{g} \frac{d}{dz} D(z)|_{z=1}\), which can be shown, after tedious calculations, to be equal to (3.17).
5. Numerical issues

5.1. Back to the roots. The applicability of the theory presented in Secs. 3 and 4 of this paper indisputably depends on finding the roots of \( z^g = Y(z)^c \) on and inside the unit circle, because these are needed to determine the unknowns \( q_0, q_1, \ldots, q_{g-1} \) in the pgf of the overflow queue \( X_g \), see (3.8). Because this issue of root-finding goes a long way back in queueing theory, it has often been addressed, both from analytical and numerical perspectives. We now give a short overview of this root-finding for the Poisson case \( Y(z) = \exp(\lambda(z - 1)) \), \( \lambda < g/c \), and point out where extensions can be made to other distributions of \( Y \).

The easiest way to determine the roots in the Poisson case is to apply successive substitution to a fixed-point equation. We know that the \( g \) roots of \( z^g = Y(z)^c \) in \(|z| \leq 1\) satisfy
\[
(5.1) \quad z = wY(z)^{c/g} = w \exp(c\lambda(z - 1)/g),
\]
where \( w = 1 \). For each feasible \( w \), (5.1) can be shown to have one unique root in \(|z| \leq 1\). Moreover, the equations can be solved by successive substitutions as
\[
(5.2) \quad z_k^{(n+1)} = w_k Y(z_k^{(n)})^{c/g}, \quad k = 0, 1, ..., g - 1,
\]
where \( w_k = \exp(2\pi ik/g) \), \( i = \sqrt{-1} \), and starting values \( z_k^{(0)} = 0 \). It can be shown that the fixed-point equations (5.2) converge to the desired roots. Adan & Zhao [2] distinguish a class of compound Poisson distributions for which the method works. For more general discrete distributions, the method is further investigated in [15].

For the Poisson case, an exact description of the roots can be obtained as well. In [15] it is shown, using the Lagrange inversion theorem, that the roots are given by (with \( \theta = c\lambda/g \))
\[
(5.3) \quad z_k = \sum_{l=1}^{\infty} e^{-i\theta} \frac{(i\theta)^{l-1}}{l!} w_k^l, \quad k = 0, 1, ..., g - 1.
\]
One could truncate the infinite series over \( l \) in (5.3) to determine the roots. For a large class of discrete distributions, exact expressions for the roots, similar to (5.3), are derived in [15].

Although the class of distributions of \( Y \) for which one can derive an exact expression such as (5.3) is far larger than the class for which the method of successive substitutions (5.2) works, see [15], neither method works for all distributions. Therefore, the most general method relies on numerical techniques. Chaudhry et al. [8] have developed an application to solve root-finding problems in queueing theory numerically. In our experience, this application works for almost all distributions.

5.2. Inversion of a pgf. For the inversion of a pgf we use a technique of Abate and Whitt [1] that relies on the Fourier series method. A distribution \( p_0, p_1, \ldots \), can be retrieved from its pgf \( P(z) = \sum_{k=0}^{\infty} p_k z^k \) via
\[
(5.4) \quad p_k = \frac{1}{2\pi i} \oint_{C_r} \frac{P(z)}{z^{k+1}} dz,
\]
where \( i = \sqrt{-1} \) and \( C_r \) is a circle about the origin of radius \( r \), \( 0 < r < 1 \). Abate and Whitt [1] approximate (5.4) by
\[
(5.5) \quad \hat{p}_k = \frac{1}{2 kr^n} \sum_{j=1}^{2k} (-1)^j \text{Re}(P(re^{ij\pi/k})),
\]
and derive for $0 < r < 1, \ k \geq 1$ the following error bound
\begin{equation}
|p_k - \tilde{p}_k| \leq \frac{r^{2k}}{1 - r^{2k}}.
\end{equation}

For practical purposes one can think of the error bound as $r^{2k}$, because $r^{2k}/(1 - r^{2k}) \approx r^{2k}$ for $r^{2k}$ small. To have accuracy up to the $\gamma$th decimal, we let $r = 10^{-\gamma/2k}$. In the upcoming numerical examples, we set $\gamma$ equal to 7.

6. Examples

We now give some examples of performance characteristics that can be derived using the framework presented in Secs. 3, 4, and 5. We consider two arrival distributions: the Poisson distribution
\begin{equation}
P(Y = j) = e^{-\lambda \frac{\lambda^j}{j!}}, \quad j = 0, 1, \ldots,
\end{equation}
with $\mu_Y = \sigma_Y^2 = \lambda < g/c$, $Y(z) = \exp(\lambda(z-1))$, and the geometric distribution
\begin{equation}
P(Y = j) = (1 - p)p^j, \quad j = 0, 1, \ldots,
\end{equation}
with $\mu_Y = p/(1 - p) < g/c$, $\sigma_Y^2 = p/(1 - p)^2$, $Y(z) = (1 - p)/(1 - pz)$. For the Poisson distribution, the $g$ roots of $z^g = Y(z)^g$ on and within the unit circle are determined with the fixed point iteration (5.2), using the stopping criterion $|z_k^{n+1} - z_k^n| < 1.0 \cdot 10^{-13}$. For the geometric distribution, the roots are determined with the software program QROOT (see [8]). All results presented in this section have been checked to be in accordance with the results of extensive simulations.

6.1. Queue length distribution. The pgf of the overflow queue $X_g$ is given by (3.8), still containing $g$ unknowns $q_0, q_1, \ldots, q_{g-1}$. As explained, these can be determined from (3.11) using the $g$ roots of $z^g = Y(z)^g$ on and within the unit circle as input. The mean overflow queue can then be calculated from (3.13). Furthermore, we can apply the inversion formula (5.5) to retrieve the whole distribution of $X_g$.

From the probability distribution, we can calculate characteristics such as the variance of the overflow queue, denoted by $\text{Var}X_g$, and the tail of the overflow queue distribution, i.e. $P(X_g \geq m)$ for some value $m \geq 0$. In Table 1 we display some of the characteristics of the overflow queue distribution for $g = r = 5$ and several values of $\mu_Y$. Observe that the geometric distribution, which has a much larger coefficient of variation than the Poisson distribution, results in a larger mean and variance of the overflow queue. Also, the percentiles

<table>
<thead>
<tr>
<th>$\mu_Y$</th>
<th>$\mathbb{E}X_g$</th>
<th>$\text{Var}X_g$</th>
<th>$P(X_g \geq 10)$</th>
<th>$P(X_g \geq 20)$</th>
<th>$P(X_g \geq 30)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson</td>
<td>0.30</td>
<td>0.1800</td>
<td>0.4285</td>
<td>2.92·10^{-9}</td>
<td>2.25·10^{-9}</td>
</tr>
<tr>
<td>0.40</td>
<td>1.0971</td>
<td>4.1807</td>
<td>8.41·10^{-3}</td>
<td>1.13·10^{-4}</td>
<td>1.52·10^{-6}</td>
</tr>
<tr>
<td>0.45</td>
<td>3.3998</td>
<td>21.7546</td>
<td>9.99·10^{-2}</td>
<td>1.26·10^{-2}</td>
<td>1.61·10^{-3}</td>
</tr>
<tr>
<td>0.49</td>
<td>23.2249</td>
<td>442.6453</td>
<td>6.22·10^{-1}</td>
<td>4.10·10^{-1}</td>
<td>2.69·10^{-1}</td>
</tr>
</tbody>
</table>

| geometric | 0.30 | 0.3000 | 0.9509 | 4.69·10^{-4} | 6.19·10^{-7} | 8.69·10^{-10} |
| 0.40 | 1.7088 | 9.1760 | 3.23·10^{-2} | 1.71·10^{-3} | 9.04·10^{-5} |
| 0.45 | 5.1807 | 48.1236 | 1.94·10^{-1} | 4.89·10^{-2} | 1.17·10^{-2} |
| 0.49 | 34.9317 | 1203.3224 | 7.24·10^{-1} | 5.52·10^{-1} | 4.21·10^{-1} |
of the distribution, which give the probabilities of exceeding certain thresholds, are much higher for the geometric distribution, which implies that the queue-length distribution has a thicker tail in the case of geometric arrivals. One may thus conclude that the arrival distribution has a substantial impact on the characteristics of the overflow queue.

In a similar way as for the length of the overflow queue, one can obtain the queue-length distribution at the end of each time slot within a cycle from the expressions (3.14) and (3.15).

6.2. Delay distribution. From the mean overflow queue (3.13) we obtain the mean delay through (3.17). However, thanks to the analysis presented in Sec. 4, we can go some steps further. The pgf of the delay is given by (4.17) and relies on the results (4.12) and (4.16). Then, from applying the inversion formula (5.5) we can obtain the whole distribution of $D$. From the distribution we can again obtain $\mathbb{E}D$, which should be equal to (3.17).

Further, as for the overflow queue, we can calculate delay characteristics such as the variance and percentiles. The pgf (4.17) is a bit typical, though, because its argument is $z^g$ instead of $z$. This turns out to be necessary, due to the following phenomenon. Tag a vehicle that joins the queue in the first time slot of a green period, and assume that it meets a batch of more than $g-1$ delayed vehicles in front of him. Then, for the next $g-1$ time slots (remainder of the green period), the batch will diminish by one vehicle per time slot. For the next $r$ time slots the batch remains unaltered, after which again, for a period of $g$ time slots, the batch decreases. This pattern repeats itself until the tagged vehicle itself can depart. The tagged vehicle experiences this as if it is served periodically by a server that works during green periods and takes vacations during red periods. To exemplify this, we give in Fig. 1 the delay distribution of a vehicle that arrives in the first slot of the green period, for a situation of $g = r = 5$, and Poisson arrivals with $\mu_Y = 0.45$. With a probability of 0.415, the vehicle meets no delayed vehicles and can pass through without delay. If the vehicle meets fewer than four delayed vehicles, it can depart from the queue within the same green period of its arrival. However, if the vehicle meets more than four delayed vehicles, it
will receive periodic service. This implies, for instance, that the vehicle cannot experience a delay of 5 till 9 slots, because no vehicle departs during the red period.

This phenomenon of periodic service complicates the analysis, see (4.8)-(4.11), and thus leads to a change of argument of the pgf of $D$ from $z$ to $z^g$. $D(z^g)$ is a pgf of a random variable $\hat{D}$ that is related to $D$ in the following way (with $u = z^g$):

$$D(z^g) = \sum_{j=0}^{\infty} \mathbb{P}(D = j)z^{gj} = \sum_{j=0}^{\infty} \mathbb{P}(\hat{D} = j)u^j = \hat{D}(u),$$

which implies that $\mathbb{P}(D = j) = \mathbb{P}(\hat{D} = gj)$ for $j = 0, 1, \ldots$. Hence, to obtain the delay distribution, we can apply the inversion formula (5.5) to determine from $\hat{D}(u)$ the probabilities $\mathbb{P}(\hat{D} = j)$, $j = 0, g, 2g, \ldots$. In Fig. 1 the delay distribution is plotted for a situation of $g = r = 5$ and Poisson arrivals with $\mu_Y = 0.45$.

In Table 2 we display some of the characteristics of the delay distribution for $g = r = 5$ and several values of $\mu_Y$. As for the overflow queue, the arrival distribution has a large impact on the characteristics of the delay distribution.

Table 2. Characteristics of the delay $D$ in the FCTL queue with $g = r = 5$.

<table>
<thead>
<tr>
<th>$\mu_Y$</th>
<th>$ED$</th>
<th>$\text{Var}D$</th>
<th>$\mathbb{P}(D \geq 10)$</th>
<th>$\mathbb{P}(D \geq 20)$</th>
<th>$\mathbb{P}(D \geq 30)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson</td>
<td>0.30</td>
<td>2.7245</td>
<td>1.82·10^{-2}</td>
<td>4.85·10^{-1}</td>
<td>3.40·10^{-1}</td>
</tr>
<tr>
<td></td>
<td>0.40</td>
<td>5.0634</td>
<td>1.47·10^{-1}</td>
<td>1.75·10^{-2}</td>
<td>2.51·10^{-3}</td>
</tr>
<tr>
<td></td>
<td>0.45</td>
<td>9.9675</td>
<td>3.89·10^{-1}</td>
<td>1.38·10^{-1}</td>
<td>4.87·10^{-2}</td>
</tr>
<tr>
<td></td>
<td>0.49</td>
<td>49.8805</td>
<td>8.23·10^{-1}</td>
<td>6.44·10^{-1}</td>
<td>5.21·10^{-1}</td>
</tr>
<tr>
<td>geometric</td>
<td>0.30</td>
<td>3.1632</td>
<td>4.69·10^{-4}</td>
<td>6.49·10^{-5}</td>
<td>8.69·10^{-6}</td>
</tr>
<tr>
<td></td>
<td>0.40</td>
<td>6.6154</td>
<td>3.23·10^{-2}</td>
<td>1.71·10^{-3}</td>
<td>9.04·10^{-5}</td>
</tr>
<tr>
<td></td>
<td>0.45</td>
<td>13.9372</td>
<td>1.94·10^{-1}</td>
<td>4.89·10^{-2}</td>
<td>1.17·10^{-2}</td>
</tr>
<tr>
<td></td>
<td>0.49</td>
<td>73.7745</td>
<td>8.74·10^{-1}</td>
<td>7.64·10^{-1}</td>
<td>6.51·10^{-1}</td>
</tr>
</tbody>
</table>
7. Concluding remarks

(i) Many of the results in queueing theory are derived and presented in the form of generating functions or Laplace transforms. Although mathematically elegant, these types of results stumble across some scepticism, as obtaining performance characteristics often requires inversion of the transform or root-finding (see e.g. [22]). One might say that Darroch’s solution to the FCTL queue, presented in Sec. 3, belongs to this category. To this day, Darroch’s formula for the mean overflow queue (3.13) is considered to be not of practical importance, because it is numerically involved (see e.g. [24], p. 4). We tend to disagree. Whether or not the evaluation of some expression is numerically involved is highly subjective and very much dependent on the spirit of the times. In the early days of queueing theory (1950’s, 1960’s), root-finding and inverting a transform were rightfully considered to be difficult and perhaps even prohibitive. However, due to improved numerical algorithms and the increase in computational power, both issues can be dealt with adequately nowadays, as discussed in Sec. 5. Also, the roots of $z^g = Y(z)^c$ on and within the unit circle can be given explicitly for a large class of distributions of $Y$, among them the Poisson distribution, see (5.3).

(ii) As mentioned in the introduction, many approximations have been derived for the mean delay in the FCTL queue. For Poisson arrivals, Webster’s formula is quite accurate. Approximations for more general arrival distributions are mostly based on the assumption that the mean delay is not very sensitive to detailed stochastic properties (see [20, 22]). However, the results presented in Sec. 6 show that the differences in terms of performance characteristics between Poisson and geometric arrivals are considerable. From this we conclude that it is of importance to consider the stochastic properties of the arrival distribution. Our model can incorporate almost every discrete arrival process and provides, as such, a high level of generality.

(iii) The framework presented in this paper allows for the evaluation of characteristics other than the mean delay. The transform solutions to the queue length and delay can be inverted by means of the procedures discussed in Sec. 5, which gives the probability distributions of the queue length and delay. From the probability distribution one can obtain characteristics such as the variance and percentiles of the distribution. These types of characteristics can be very useful for the design and performance evaluation of signalized intersections.

(iv) Finally, let us mention some possible model extensions. It is our belief that the same type of modeling presented in this paper can be applied to a two-dimensional model for two opposed traffic streams at an intersection. In that case, we model two queues simultaneously, and green for queue 1 means red for queue 2 and vice versa. For this two-dimensional model, we expect that we can handle, next to a fixed cycle control, other types of control as well. One could think of priority for either one of the queues, and switching from green to red whenever a queue idles. These models are interesting subjects for future research.
References


Eurandom, 5612 AZ Eindhoven, The Netherlands
E-mail address: vanleeuwaarden@eurandom.tue.nl