Directional Derivatives of Oblique Reflection Maps

Avishai Mandelbaum
School of Industrial Engineering and Management
Technion,
Haifa, Israel

and

Kavita Ramanan
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA

Abstract

Given an oblique reflection map Γ and functions $\psi, \chi \in \mathcal{D}_{\lim}$ (the space of functions that have left and right limits at every point), the directional derivative $\nabla_\chi \Gamma(\psi)$ of Γ along χ, evaluated at ψ, is defined to be the pointwise limit (as $\varepsilon \downarrow 0$) of the family of functions

$$\nabla_\chi^\varepsilon \Gamma(\psi) \doteq \frac{1}{\varepsilon}\left[\Gamma(\psi + \varepsilon\chi) - \Gamma(\psi)\right].$$

Directional derivatives are shown to exist and lie in $\mathcal{D}_{\lim}$ for oblique reflection maps of the so-called Harrison-Reiman class. When ψ and χ are continuous, the convergence of $\nabla_\chi^\varepsilon \Gamma(\psi)$ to $\nabla_\chi \Gamma(\psi)$ is shown to be uniform on compact subsets of continuity points of the limit $\nabla_\chi \Gamma(\psi)$ and the derivative $\nabla_\chi \Gamma(\psi)$ is shown to have an autonomous characterization as the unique fixed point of an associated map. Motivation for the study of directional derivatives arises from the fact that they characterize functional central limits of non-stationary queueing networks. This work also shows how the various types of discontinuities of the derivative $\nabla_\chi \Gamma(\psi)$ are related to the topology of the network as well as to the states (of underloading, overloading or criticality) of the various queues in the network. The latter classification is necessary for proving a stronger form of convergence of the functions $\nabla_\chi^\varepsilon \Gamma(\psi)$ to the limit $\nabla_\chi \Gamma(\psi)$, which is in turn useful for obtaining functional central limit theorems for non-stationary queueing networks.

Key words. Reflection map, Skorokhod Map, directional derivatives, queueing networks, non-stationarity, time-varying rates, periodic queues, fluid limits, diffusion approximations, heavy traffic, directional derivatives, overloading, underloading, criticality.

Contents

1 Introduction

Most real-world queueing systems evolve according to laws that vary with time. However, the majority of queueing research has been devoted to time-homogeneous models. While such models may provide reasonable approximations for slowly varying systems, they completely fail to capture many important phenomena such as surges in demand, sudden node failures and periodicity. The explicit analysis of even stationary networks is usually intractable. Instead one usually resorts to asymptotic approximations, which often capture the essential features of network behaviour. A commonly used asymptotic approximation is the conventional heavy-traffic approximation, where arrival and service rates are scaled proportionately but the number of servers are kept constant [34, 37]. In recent years, much progress has been made on this kind of approximations for stationary networks, under reasonably general assumptions on the arrival, service and routing processes (see, for example, the survey [47] and papers cited therein). In contrast, the analysis of time-dependent networks remains challenging even in a Markovian setting. In particular, there has been relatively little work done on non-stationary queueing networks in the conventional heavy-traffic regime. Non-stationary networks for which the conventional asymptotic regime is relevant arise frequently in models of transportation, telecommunication and computer systems [9, 20, 31].

The single queue with time-varying arrival and service rates has been studied by various authors under different assumptions [15, 22, 28, 29, 38, 39]. The detailed asymptotic analysis carried out in [22] shows that the so-called fluid limit of a time-dependent Markovian queue alternates between phases of overloading, critical loading and underloading, and also shows that the functional central limit exhibits different characteristics in each of the three different phases of loading. The analysis in [22] is pathwise and uses strong approximations to represent the functional central limit of the time-dependent queue in terms of directional derivatives of the one-dimensional reflection map Γ. (See also [10] for an insightful discussion of the use of strong approximations in queueing theory.) The characterization of the directional derivative in [22] relies heavily on the following explicit form of the one-dimensional reflection map $\Gamma : \mathcal{C}([0,\infty) : \mathbb{R}) \to \mathcal{C}([0,\infty) : \mathbb{R}_+)$ due to [42]:

$$\Gamma(\psi)(t) = \psi(t) + \theta(t), \tag{1.1}$$

where the constraining term θ that keeps $\Gamma(\psi)$ in $\mathbb{R}_+$ is given by

$$\theta(t) = \left[\sup_{0\leq s\leq t} -\psi(s)\right] \vee 0. \tag{1.2}$$

In contrast, in the multi-dimensional or network setting there is no explicit expression for the reflection map, making characterization of its directional derivatives considerably more involved. In fact derivatives of reflection maps associated with even feedforward tandem networks cannot always be expressed simply as a composition of directional derivatives of one-dimensional reflection maps (see Section 4.2.1 for further discussion of this fact). The network setting also introduces additional complications due to dependence on network topology and leads to interesting new questions about when and how effects propagate through the network. Consequently new techniques need to be developed to analyze the network setting.

The main objectives of this work are to introduce and characterize properties of directional derivatives of a class of multi-dimensional reflection maps associated with single-class open queueing networks, and to illustrate the practical insights that can be obtained from such an analysis. In particular, existence of directional derivatives for the class of so-called Harrison-Reiman reflection maps (see Theorem 1.1) is established and the dependence of the behaviour of directional derivatives on both the states of loading in various queues in the network as well as on the topology of the network is described (see Theorem 1.2 and the illustrative examples in Section 2). As motivation for this work, a heuristic explanation of the connection between directional derivatives of multi-dimensional reflection maps and functional central limit theorems for stationary and non-stationary queueing networks is provided in Section 1.1. Properties of the derivative established in this work are used to make this connection rigorous in [25]. Features of the asymptotic analysis that are special to non-stationary networks are discussed in Section 1.2, while the main results of the paper are summarized in Section 1.3. Some basic notation and a description of the organization of the rest of the paper is provided in Section 1.4.

1.1 Directional Derivatives and Functional Central Limits

Conventional diffusion approximations of queueing networks are often obtained by the following general procedure [5, 6, 24, 36, 37, 45]. Consider a sequence of queueing networks defined in terms of their primitives (i.e. the random processes defined on some probability space $(\Omega, \mathcal{F}, P)$ that describe arrivals, services and routing, as well as the scheduling rules). To each queueing network in the sequence one constructs from the primitives a certain netput process $\tilde{X}^n$ such that the evolution of $\tilde{X}^n$ coincides with the evolution of the queue length process $\tilde{Z}^n$ only when all queues are non-empty. In general the queue length process is a complicated functional of the netput process:

$$\tilde{Z}^n = F^n(\tilde{X}^n).$$

The sequence of netput processes $\{\tilde{X}^n\}$ is assumed to satisfy a functional strong law of large numbers (FSLLN) and functional central limit theorem (FCLT). Specifically, if X^n is defined by

$$X^n(t) \doteq \tilde{X}^n(nt) \qquad \text{for } t \in [0, \infty), \tag{1.3}$$

and

$$\bar{X}^n \doteq \frac{1}{n} X^n, \tag{1.4}$$

then the FSLLN, which characterizes the long-term mean behaviour of the netput process, is given by

$$\bar{X}^n \to \bar{X} \qquad \text{as } n \to \infty, \tag{1.5}$$

where the limit is with respect to the topology of P-a.s. convergence. Similarly, the FCLT for the netput process takes the form

$$\hat{X}^n \to \hat{X} \qquad \text{as } n \to \infty, \tag{1.6}$$

where the limit is with respect to the topology of weak convergence with respect to an appropriate topology on path space (e.g. uniform convergence on compact sets), and

$$\hat{X}^n \doteq \sqrt{n}\, [\bar{X}^n - \bar{X}] \tag{1.7}$$

is a rescaled centered version of the netput process that captures the fluctuations around its mean. In order to obtain a corresponding FSLLN and FCLT for the queue length process, Z^n

and $\bar{Z}^n$ are defined as in (1.3) and (1.4) respectively, with X replaced by Z. Homogeneity of the functionals F^n with respect to space and time are first used to infer that

$$\bar{Z}^n = F^n(\bar{X}^n). \tag{1.8}$$

The FSLLN for the queue length process is then obtained by establishing the convergence

$$\bar{Z}^n = F^n(\bar{X}^n) \to F(\bar{X}) \qquad \text{as } n \to \infty, \tag{1.9}$$

where $\bar{X}$ is the P-a.s. FSLLN limit of the netput process obtained in (1.5), F is the multi-dimensional oblique reflection mapping associated with the queueing network (see Section 3.1 for a precise definition) and

$$\bar{Z} \doteq F(\bar{X}) \tag{1.10}$$

is referred to as the fluid limit. Note that when $F^n \equiv F$, as is the case for a single station or when the routing is deterministic (see [6] for details), and F can be shown to be continuous, then the P-a.s. FSLLN (1.9) for the queue length process is a direct consequence of the FSLLN (1.5) for the netput process.

In order to characterize the fluctuations of the queue lengths around the fluid limit, we analyze limits of the centered sequence $\{\hat{Z}^n\}$ of queue lengths defined by

$$\hat{Z}^n \doteq \sqrt{n}\left[\bar{Z}^n - \bar{Z}\right]. \tag{1.11}$$

Substituting relations (1.7), (1.8) and (1.10) into the above display yields the relations

$$\begin{aligned}
\hat{Z}^n &= \sqrt{n}\left[F^n(\bar{X}^n) - F(\bar{X})\right] \\
&= \sqrt{n}\left[F^n(\bar{X}^n) - F(\bar{X}^n)\right] + \sqrt{n}\left[F(\bar{X}^n) - F(\bar{X})\right] \\
&= \sqrt{n}\left[F^n(\bar{X}^n) - F(\bar{X}^n)\right] + \sqrt{n}\left[F\left(\bar{X} + \frac{1}{\sqrt{n}}\hat{X}^n\right) - F(\bar{X})\right].
\end{aligned}$$

In many cases it is possible to show that with respect to a suitable topology on path space

$$\sqrt{n}\left[F^n(\bar{X}^n) - F(\bar{X}^n)\right] \to 0 \qquad \text{as } n \to \infty \tag{1.12}$$

$$\left[F\left(\bar{X} + \frac{1}{\sqrt{n}}\hat{X}^n\right) - F\left(\bar{X} + \frac{1}{\sqrt{n}}\hat{X}\right)\right] \to 0 \qquad \text{as } n \to \infty \tag{1.13}$$

and the limit

$$\nabla_{\hat{X}} F(\bar{X}) \doteq \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon}\left[F\left(\bar{X} + \varepsilon\hat{X}\right) - F(\bar{X})\right] \tag{1.14}$$

exists. In such a situation, by setting $\varepsilon \doteq 1/\sqrt{n}$ and $\hat{Z}^\varepsilon \doteq \hat{Z}^n$, the last four displays can be combined to obtain

$$\hat{Z} \doteq \lim_{\varepsilon \downarrow 0} \hat{Z}^\varepsilon = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon}\left[F\left(\bar{X} + \varepsilon\hat{X}\right) - F(\bar{X})\right] = \nabla_{\hat{X}} F(\bar{X}). \tag{1.15}$$

The notation in (1.14) reflects the fact that $\nabla_{\hat{X}} F(\bar{X})$ represents the directional derivative of the reflection map F in the direction $\hat{X}$, evaluated at $\bar{X}$. Note that, strictly speaking, the claim that (1.15) follows from (1.12)-(1.14) depends on the topology in which the convergence in (1.12)-(1.15) take place. It clearly holds if all limits takes place with respect to the topology of uniform convergence on compact sets (u.o.c.). However, if the convergence in (1.15) takes place with respect to another topology on path space, as is indeed the case for non-stationary networks, then a more careful analysis is required. We omit a discussion of these subtleties in this heuristic motivational exposition, and instead refer the reader to [25] for more details.

In summary, under suitable assumptions, in the conventional heavy traffic asymptotic regime the fluid limit and functional central limits of the queue length process are given by

$$\bar{Z} = F(\bar{X}) \quad \text{and} \quad \hat{Z} = \nabla_{\hat{X}} F(\bar{X}), \tag{1.16}$$

where F is the multi-dimensional oblique reflection map and $\nabla_{\hat{X}} F(\bar{X})$ is the directional derivative of F, as defined in (1.14). Thus we have indicated how *directional derivatives of the oblique reflection map* arise naturally when establishing functional central limit theorems for queueing networks in the conventional heavy traffic regime. The general philosophy of interpreting functional central limit theorems as directional derivatives of a suitable mapping is also useful in other settings. In particular, this interpretation is also valid in the so-called Halfin-Whitt asymptotic regime [11] (see Section 7 for a brief discussion of this issue).

1.2 Stationary vs. Non-stationary Queueing Networks

As indicated in the last section, functional central limits of queueing networks in the conventional heavy traffic regime often have an interpretation as directional derivatives of corresponding reflection maps [25]. For stationary networks, the primitives describing the queueing networks are time-independent. In this case, under so-called heavy-traffic conditions, the fluid limit $\bar{X}$ of the netput process is trivial (i.e. $\bar{X} \equiv 0$, where 0 represents the function that is identically zero). Since the reflection map F is homogeneous (i.e. $F(\alpha g) = \alpha F(g)$ for all $\alpha > 0$ and g in the domain of F), it follows that $F(0) = 0$ (where the latter 0 is again the zero function) and hence the directional derivative $\nabla_{\hat{X}} F$ evaluated at $\bar{X} = 0$ is simply the function F evaluated at $\hat{X}$. As a result the representation for fluid and functional central limits for the queueing network in (1.16) takes the simpler, more familiar form

$$\bar{Z} \equiv 0 \quad \text{and} \quad \hat{Z} = F(\hat{X}).$$

Thus in the stationary setting, the heavy traffic diffusion limit is simply $F(\hat{X})$. In constrast, in most interesting cases the fluid limit of a non-stationary queueing network is not trivial, since each node in the network varies dynamically with time between periods of overloading, criticality and underloading. As a consequence the functional central limit is not in general equal to the image of $\hat{X}$ under the reflection map, and hence a complete characterization of the directional derivatives of the multi-dimensional reflection map is required.

Another difference between the stationary and non-stationary settings is that the asymptotic scalings X^n defined in (1.3) to approximate stationary networks are not adequate in the non-stationary setting. This is because the stationary scaling approximates the present behaviour of the system by its behaviour at infinity, consequently erasing all the dynamic behaviour that is of interest in the presence of non-stationarity (see [1, 2, 13, 19] for some discussion of this issue in the context of non-stationary queues with periodic evolution). Building on earlier work in [17] and [31], a time-inhomogeneous analogue to steady state analysis involving the perturbation of transition probabilities was proposed in [28], where it was referred to as *uniform acceleration*. The perturbation analysis was then extended to the sample-path level using the theory of strong approximations in [22]. This approach involves scaling all the average instantaneous transition rates of the Markovian model by a factor of $1/\varepsilon$. As $\varepsilon \downarrow 0$ each rate increases, or *is accelerated*, in absolute terms, while the ratio of any two rates relative to each other is held fixed. Uniform acceleration enables a *dynamic* asymptotic analysis of non-stationary queueing models, and it coincides with the usual (heavy traffic or steady state) approximations when applied to stationary models. As a result, in [25] the characterization of functional central limits for non-stationary networks in terms of directional derivatives of reflection maps is carried out within the asymptotic framework of uniform acceleration.

1.3 Main Results of the Paper

As mentioned in Section 1.1, given $\psi, \chi \in \mathcal{D}_{\lim}$, the derivative $\nabla_\chi \Gamma(\psi)$ is defined to be the unique limit of the family of functions

$$\nabla_\chi^\varepsilon \Gamma(\psi) \doteq \varepsilon^{-1} \left[\Gamma\left(\psi + \varepsilon \chi\right) - \Gamma\left(\psi\right) \right], \tag{1.17}$$

where Γ is the oblique reflection mapping. In [22] it was shown that in the one-dimensional case where Γ is given by (1.1), if ψ, χ are continuous, then

$$\nabla_\chi \Gamma(\psi)(t) = \chi(t) + \sup_{s \in \Phi(t)} [-\chi(s)] \vee 0,$$

where

$$\Phi(t) \doteq \{s \in [0,t] : \Gamma(\psi)(s) = 0 \text{ and } \theta(s) = \theta(t)\},$$

and θ is defined by (1.2). (In reality, this was shown in [22] under the additional restriction that $\nabla_\chi \Gamma(\psi)$ has only a finite number of discontinuities in any compact interval, but as shown in Theorem 4.7 and [46, Theorem 9.3.1], this condition can be relaxed.) When ψ and χ are the fluid and functional central limits of the netput process associated with a time-varying queue, then $\nabla_\chi \Gamma(\psi)$ characterizes the functional central limit of the time-varying queue. In this case $\Gamma(\psi)$ has an interpretation as the fluid limit of the queue and θ as the corresponding cumulative potential outflow lost (due to idleness of the server) during the period $[0,t]$. Thus $\Phi(t)$ represents the set of all times s in the interval $[0,t]$ when the fluid queue was zero, but the server was fully utilized in the interval $[s,t]$.

In the multi-dimensional setting there exist analogous quantities $\theta^i(t)$ (see Definition 3.1 for a rigorous definition), which represents the cumulative potential outflow lost from the ith queue during $[0,t]$, and sets

$$\Phi^i(t) \doteq \left\{ s \in [0,t] : \Gamma(\psi)^i(s) = 0 \text{ and } \theta^i(s) = \theta^i(t) \right\}, \tag{1.18}$$

of the times $s \in [0,t]$ at which the ith fluid queue is zero but the ith server is fully utilized during $[s,t]$. The first main result of this work, Theorem 1.1, characterizes the multi-dimensional derivative in terms of these quantities. The H-R class of constraint matrices mentioned in the statement of the theorem is defined in Section 3.1 and includes constraint matrices that model Jackson networks with single-class customers in which arrivals and services are allowed to be time-varying. Throughout the paper $\mathcal{C}$ denotes the class of $\mathbb{R}^K$-valued continuous functions on $[0,\infty)$ and $\mathcal{D}_{\lim}$ is the space of $\mathbb{R}^K$-valued functions on $[0,\infty)$ that have left and right limits at every point $t \in [0,\infty)$.

Theorem 1.1 *Let R be an H-R constraint matrix in $\mathbb{R}^{K \times K}$, let $P \doteq I - R$ and let Γ be the associated oblique reflection map. Given $\psi, \chi \in \mathcal{D}_{\lim}$, the family of functions $\nabla_\chi^\varepsilon(\psi)$ defined in*

(1.17) converges pointwise to a limit $\nabla_\chi \Gamma(\psi)$*, which lies in* $\mathcal{D}_{\lim}$*. Moreover when* $\psi, \chi \in \mathcal{C}$*, the convergence in (1.17) is uniform on compact subsets of continuity points of* $\nabla_\chi \Gamma(\psi)$ *and*

$$\nabla_\chi \Gamma(\psi) = \chi + R\gamma(\psi, \chi),$$

where $\gamma \doteq \gamma(\psi, \chi)$ *is the unique solution to the system of equations*

$$\gamma^i(t) = \begin{cases} 0 & \text{if } t \in (0, t_l^i) \\ \sup_{s \in [t_l^i, t]} \left[-\chi^i(s) + [P\gamma]^i(s)\right] \vee 0 & \text{if } t \in [t_l^i, t_u^i) \\ \sup_{s \in \Phi^i(t)} \left[-\chi^i(s) + [P\gamma]^i(s)\right] & \text{if } t \in [t_u^i, \infty) \end{cases} \tag{1.19}$$

for $i = 1, \ldots, K$*, where* Φ^i *is defined by (1.18) and*

$$\begin{aligned} t_l^i &\doteq \inf\{t > 0 : \Gamma(\psi)^i(t) = 0\} \\ t_u^i &\doteq \sup\{t > 0 : \dot{\theta}^i(t+) > 0\} \end{aligned}$$

are the first times that queue i *becomes empty and underloaded respectively. In particular the derivative* $\nabla_\chi \Gamma(\psi)$ *is Lipschitz in* χ *and also satisfies for* $\alpha, \beta > 0$

$$\Gamma_{\alpha\chi}(\beta\psi) = \alpha \Gamma_\chi(\psi).$$

Remark. Note that if all queues are initially empty and are subsequently always under heavy-traffic (i.e. they are always critically loaded with $\Gamma(\psi)^i(t) = 0$ and $\dot{\theta}^i(t+) = \dot{\theta}^i(t-) = 0$), then $t_l^i = 0$, $t_u^i = \infty$ and $\Phi^i(t) = [0, t]$ for every $i = 1, \ldots, K$ and $t \in [0, \infty)$. In this case γ is the unique solution to the system of equations

$$\gamma^i(t) = \sup_{s \in [0,t]} \left[-\chi^i(s) + [P\gamma]^i(s)\right] \vee 0 \qquad \text{for } i = 1, \ldots, K.$$

By Theorem 3.2 (see also [14]) this implies that the derivative is simply the reflected or constrained version of χ:

$$\nabla_\chi \Gamma(\psi) = \Gamma(\chi),$$

which is consistent with the well-known reflected Brownian motion characterization of heavy-traffic limits of stationary open single-class queueing networks [14, 37].

A slightly more general version of Theorem 1.1 is stated as Theorem 4.10 and proved in Section 4.2. The representation for Φ^i given in (1.18) follows from Theorem 4.10, relation (4.21) and Lemma 5.1(1). Since even when $\psi, \chi \in \mathcal{C}$ the convergence in Theorem 1.1 is uniform only on compact subsets of continuity points of the derivative $\nabla_\chi \Gamma(\psi)$, in order to establish convergence with respect to stronger topologies it is necessary to understand the structure of the discontinuities of $\nabla_\chi \Gamma(\psi)$. The next main result of the paper, Theorem 1.2, describes the various types of discontinuities exhibited by the derivative, and identifies necessary conditions for a discontinuity to take place. The examples in Section 2 demonstrate that each of the different types of discontinuities mentioned in the theorem does indeed occur for some oblique reflection problem associated with a non-stationary Jackson network. The various regimes referred to in the theorem are introduced in Sections 5.1 and 5.2. Roughly speaking, i is said to be overloaded, critical or underloaded if the ith queue in the network is respectively non-empty, empty but with a fully utilized server or empty with a server having spare capacity. (A precise definition of these terms is given in Definition 5.1.) Likewise, a critical (sub-critical) chain is said to precede i if there is either a loop of empty buffers that contains i, or there exists a sequence of empty buffers with the property that the first buffer is at the end of overloading (respectively start of underloading), the last buffer is i and each buffer in the sequence (when non-empty) routes its contents to the next buffer in the sequence with positive probability. (See Definition 5.2 for a precise description of these chains.) We now state Theorem 1.2, whose proof is presented in Section 5.2.

Theorem 1.2 *Given a constraint matrix R satisfying the H-R condition with associated reflection map Γ, and given functions $\psi, \chi \in \mathcal{C}$, the following are necessary conditions for the existence of discontinuities in the directional derivative $\nabla\Gamma \doteq \nabla_\chi \Gamma(\psi)$ at $t \in [0,\infty)$.*

(L) If $\nabla\Gamma^i$ has a left discontinuity then either
(a) i is at the end of overloading, in which case

$$\nabla\Gamma^i(t-) < \nabla\Gamma^i(t) = 0; \tag{1.20}$$

or
(b) i is not underloaded and a critical chain precedes i, in which case

$$\nabla\Gamma^i(t-) > \nabla\Gamma^i(t). \tag{1.21}$$

(R) If $\nabla\Gamma^i$ has a right discontinuity then either
(a) i is at the start of underloading, in which case

$$\nabla\Gamma^i(t) > \nabla\Gamma^i(t+) = 0; \tag{1.22}$$

or
(b) i is not underloaded and a sub-critical chain precedes i, in which case

$$\nabla\Gamma^i(t) < \nabla\Gamma^i(t+). \tag{1.23}$$

(LR) If $\nabla\Gamma^i$ is neither right nor left discontinuous, then either
(a) i is at the end of overloading and a sub-critical chain precedes i, in which case

$$\nabla\Gamma^i(t-) < \nabla\Gamma^i(t) = 0 < \nabla\Gamma^i(t+);$$

or
(b) i is at the start of underloading and a critical chain precedes i, in which case

$$0 = \nabla\Gamma^i(t-) > \nabla\Gamma^i(t) > \nabla\Gamma^i(t+);$$

or
(c) i is not underloaded and there exist both critical and sub-critical chains preceding i, in which case the discontinuity is a separated discontinuity of the form

$$\nabla\Gamma^i(t) < \min\left[\nabla\Gamma^i(t-), \nabla\Gamma^i(t+)\right]. \tag{1.24}$$

Finally, if i is underloaded at t then $\nabla\Gamma^i(t-) = \nabla\Gamma^i(t) = \nabla\Gamma^i(t+) = 0$ and if i is overloaded at t then

$$\nabla\Gamma^i(t) \leq \min\left[\nabla\Gamma^i(t-), \nabla\Gamma^i(t+)\right]. \tag{1.25}$$

Henceforth, we will use the notation "discontinuities of type (La), (Ra) etc." to refer to the different types of discontinuities described in Theorem 1.2 above.

As mentioned earlier, in [25] functional central limits for non-stationary queueing networks are established and characterized in terms of the directional derivative of the reflection map. This involves proving pathwise convergence of the pre-limit processes (1.17) to the derivative in a topology that is strong enough to imply convergence of the corresponding stochastic processes. As shown in [22], even in the one-dimensional setting this requires significant effort – in particular convergence to the derivative can be established only in the Skorokhod M_1 topology [33, 41]. The derivative of the one-dimensional reflection map is upper semicontinuous and lies in $\mathcal{D}_{l,r}$, the space of functions that are either right continuous or left continuous at every point. Thus it exhibits discontinuities only of type (La) and (Ra), which happen only when the queue is empty and is either at the end of overloading or at the start of underloading. In the multi-dimensional setting, discontinuities in the ith component of the derivative can occur even when queue i is overloaded. Thus the multi-dimensional setting is considerably more complicated and requires the consideration of more general topologies. The properties derived in Theorems 1.2 and 5.2 of this paper are used in [25] to prove fundamental convergence results that are required to establish functional central limit theorems for non-stationary queueing networks.

1.4 Organization of the Paper and Some Basic Notation

The organization of the rest of the paper is as follows. In Section 2 we provide several illustrative examples to demonstrate some interesting features of directional derivatives associated with various queueing networks. In Section 3 we first recall the definition and properties of the multi-dimensional oblique reflection map, and then introduce the notion of a directional derivative of the reflection map (Definition 3.3). In Section 4 we characterize the derivative for the so-called Harrison-Reiman class of oblique reflection problems, which arise from open single-class queueing networks [14, 37]. In Section 5 we derive properties of the derivative when $\psi, \chi \in \mathcal{C}$. Section 6 contains proofs of the main theorems, while Section 7 contains some concluding observations. Section A of the Appendix contains proofs of some auxiliary results. When considering functions with discontinuities, the function space of right continuous functions with left limits ($\mathcal{D}_r$) has been most commonly used in the literature. However, we will have cause to use more general function spaces, whose definitions are given in Section B of the Appendix as a convenient reference. In addition, each function space is defined for the first time it is used in the paper.

We close this section with some other common notation used throughout the paper. For $a, b \in \mathbb{R}$ let $a \vee b = \max(a, b)$ and $a \wedge b = \min(a, b)$. Given a vector $x \in \mathbb{R}^K$, x^i or $[x]^i$ will be used to denote the ith component of the vector. For $a \in I\!\!R^K$, we define the norm

$$|a| \doteq \max_{i=1,\ldots,K} |a^i| \tag{1.26}$$

where, for $a^i \in I\!\!R$, $|a^i|$ denotes the usual Euclidan norm. Given a $K \times K$ matrix R, R^T denotes its transpose, $\sigma(R)$ its spectral radius and R_{ij} represents the entry in the jth column and ith row of R. The matrix I represents the $K \times K$ identity matrix, and $\{e_i,\ i = 1, \ldots, K\}$ is the standard orthonormal basis in $\mathbb{R}^K$. Inequalities of vectors and matrices should be interpreted componentwise. Vectors will always be expressed as column vectors. We will deal with many limits in this paper. The notation $\uparrow$ ($\downarrow$) will be used to denote monotone nondecreasing (nonincreasing) convergence to a limit with respect to a particular topology, which will be made clear in each context.

Given a function f on $[0,\infty)$ that takes values in $\mathbb{R}^K$, f^i denotes the ith coordinate function. Moreover, for $T < \infty$ $||f||_T$ denotes the supremum norm:

$$||f||_T \doteq \sup_{s\in[0,T]} |f(s)|,$$

where $|\cdot|$ is the norm defined above in (1.26). Given a real-valued function f, the notation $\overline{f}$ is used to denote the supremum function:

$$\overline{f}(t) \doteq \sup_{s\in[0,t]} f(s).$$

Finally for functions f of bounded variation $|f|_t$ denotes the total variation norm on $[0,t]$ (with respect to the norm $|\cdot|$ defined in (1.26) on $\mathbb{R}^K$).

2 Illustrative Examples

In this section we characterize the directional derivatives associated with two concrete time-dependent queueing networks that arise in applications

2.1 Periodic Arrivals to a Tandem Queue

There are many examples of transportation and computer systems whose mean arrival rates vary periodically with time, e.g. depending on the time of day [31]. Approximations to these systems in terms of average arrival rates over the whole period are clearly inadequate, and it is often of particular interest to determine how the system behaves around points of sharp transitions in the arrival rates. Here we examine a tandem queueing network with a periodic time-dependent arrival rate and identify the associated directional derivatives.

2.1.1 The Tandem Queueing Model

Consider the tandem queueing system illustrated in Figure 3. The associated reflection map, which is depicted in Figure 2, has routing matrix P^T and reflection matrix R given by

$$P^T \doteq \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad \text{and} \qquad R \doteq \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}. \tag{2.1}$$

It is easy to verify that R satisfies the Harrison-Reiman condition (see Definition 3.2), and let Γ denote the associated reflection map.

Note that there are $K = 2$ queues in the network. Suppose there are no exogeneous arrivals to queue 2, and that the arrivals to queue 1 are governed by a non-homogeneous Poisson arrival process with a time-dependent mean rate λ that is periodic with period $4T$ and defined as follows.

$$\lambda^1(t) \doteq \begin{cases} 4t+4T & \text{for } t \in [0,T) \\ 8T-4(t-T) & \text{for } t \in [T,3T) \\ 4(t-3T) & \text{for } t \in [3T,4T], \end{cases} \tag{2.2}$$

and $\lambda^1(t) = \lambda^1(t \bmod 4T)$ for $t \in (4T,\infty)$ (see Figure 1). Moreover, suppose the service rate is constant and equal to $\mu^1 = 5T$ at queue 1 and equal to $\mu^2 = 4.5T$ at queue 2. Note that the average arrival rate $\overline{\lambda}^1$ to queue 1 over the period T is 4.

2.1.2 Fluid Limit

We now define $\psi \in \mathcal{C}$ to be the average netput process, so that $\psi^i(t)$ represents the total cumulative arrivals to minus the total service that would have taken place had both queues been non-empty throughout the interval $[0,t]$. Thus for $i=1,2$

$$\dot{\psi}^1(t) \doteq \lambda(t) - \mu^1, \qquad \text{and} \qquad \dot{\psi}^2(t) \doteq (\mu^1 - \mu^2)$$

and let (ϕ,θ) be the solution to the oblique reflection problem defined by R with input ψ. Then it can be easily verified that

$$\dot{\phi}^1(t) \doteq \begin{cases} 0 & \text{for } t \in (0,T/4) \\ 4(t-T/4) & \text{for } t \in (T/4,T) \\ 3T-4(t-T) & \text{for } t \in (T,(1+\alpha)T) \\ 0 & \text{for } t \in ((1+\alpha)T,4T), \end{cases}$$

and

$$\phi^1(t) \doteq \begin{cases} 0 & \text{for } t \in [0,T/4) \\ 2\left(t-\dfrac{T}{4}\right)^2 & \text{for } t \in [T/4,T) \\ \dfrac{9T^2}{8} + 3T(t-T) - 2(t-T)^2 & \text{for } t \in [T,(1+\alpha)T) \\ 0 & \text{for } t \in [(1+\alpha)T,4T), \end{cases}$$

where $\alpha \doteq \frac{3}{4}(1+\sqrt{2}) < 2$. The departure rate $\dot{\nu}^1$ from the first queue is therefore

$$\dot{\nu}^1(t) \doteq \begin{cases} 4t+4T & \text{for } t \in [0,T/4) \\ 5T & \text{for } t \in [T/4,(1+\alpha)T) \\ 8T-4(t-T) & \text{for } t \in [(1+\alpha)T,3T) \\ 4(t-3T) & \text{for } t \in [3T,4T]. \end{cases}$$

Since the service rate at queue 2 is $\mu^2 = 4.5T$ this implies that

$$\dot{\phi}^2(t) \doteq \begin{cases} 0 & \text{for } t \in [0,T/8) \\ 4\left(t-\dfrac{T}{8}\right) & \text{for } t \in [T/8,T/4) \\ 0.5T & \text{for } t \in [T/4,(1+\alpha)T) \\ 3.5T-4(t-T) = \dfrac{(1-6\sqrt{2})T}{2} - 4(t-(1+\alpha)T) & \text{for } t \in [(1+\alpha)T,3T) \\ 4(t-3T)-4.5T & \text{for } t \in [3T,(1+\beta)T) \\ 0 & \text{for } t \in [(3+\beta)T,4T] \end{cases}$$

and $\phi^2(t)$ is equal to

$$\begin{cases} 0 & \text{for } t \in [0, T/8) \\ 2\left(t - \frac{T}{8}\right)^2 & \text{for } t \in [T/8, T/4) \\ \frac{T^2}{32} + \frac{T}{2}\left(t - \frac{T}{4}\right) & \text{for } t \in [T/4, (1+\alpha)T) \\ \frac{(25+12\sqrt{2})T^2}{32} + \frac{(1-6\sqrt{2})T(t-(1+\alpha)T)}{2} - 2(t-(1+\alpha)T)^2 & \text{for } t \in [(1+\alpha)T, 3T) \\ \frac{17T^2}{32} + 2(t-3T)^2 - 4.5T(t-3T) & \text{for } t \in [3T, (3+\beta)T] \\ 0 & \text{for } t \in [(3+\beta)T, 4T], \end{cases}$$

where $\beta = 1/8 < 1$. Note that the fluid arrival rate λ^1 to queue 1 peaks at time $t = T$, while the mean queue 1 content peaks at time $t = 7T/4$ and the mean queue 2 content peaks at time $t = (1+\alpha)T$. Thus the time of peak congestion is delayed by an amount $\delta_1 = 3T/4$ at the first queue and by an additional amount $\delta_2 = 3\sqrt{2}T/4 > \delta_1$ at the second queue. Figure 1 shows plots of the queue lengths and arrival rates with time, and Figure 3 illustrates the different states of the fluid limit of the tandem queue at various times during the interval $[0, 4T]$. In contrast, if the system were approximated by the time-averaged arrival rate over the period, which is equal to $4T$, the traffic intensities in both queues would be less than 1, and hence both queues would be empty throughout the interval.

2.1.3 Directional Derivatives for the Tandem Queue

From the characterization of the directional derivatives given in Theorem 1.1 it follows that for the oblique reflection problem associated with the H-R constraint matrix R specified in (2.1),

$$\begin{aligned} \nabla\Gamma^1(t) &= \chi^1(t) + \sup_{s\in\Phi^1(t)} \left[-\chi^1(s)\right] \\ \nabla\Gamma^2(t) &= \chi^2(t) - \sup_{s\in\Phi_{-\xi^1}(t)} \left[-\chi^1(s)\right] + \sup_{s\in\Phi^2(t)} \left[-\chi^2(s) + \sup_{r\in\Phi^1(s)} \left[-\chi^1(r)\right]\right]. \end{aligned}$$

The representation (5.2) for $\Phi^i(t)$, combined with the explicit expression for ϕ given above, shows that

$$\Phi^1(t) \doteq \begin{cases} \{t\} & \text{for } t \in [0, T/4) \\ \left\{\frac{T}{4}\right\} & \text{for } t \in [T/4, (1+\alpha)T) \\ \left\{\frac{T}{4}, (1+\alpha)T\right\} & \text{for } t = (1+\alpha)T \\ \{t\} & \text{for } t \in ((1+\alpha)T, 4T], \end{cases}$$

$$\Phi^2(t) \doteq \begin{cases} \{t\} & \text{for } t \in [0, T/8) \\ \left\{\frac{T}{8}\right\} & \text{for } t \in [T/8, (3+\beta)T) \\ \left\{\frac{T}{8}, (3+\beta)TT\right\} & \text{for } t = (3+\beta)T \\ \{t\} & \text{for } t \in ((3+\beta)T, 4T]. \end{cases}$$

Together the last two displays show that $\nabla\Gamma^1$ is equal to

$$\begin{cases} 0 & \text{for } t \in [0, T/4) \\ \chi^1(t) - \chi^1(T/4) & \text{for } t \in [T/4, (1+\alpha)T) \\ \chi^1((1+\alpha)T) + [-\chi^1(T/4) \vee -\chi^1((1+\alpha)T)] & \text{for } t = (1+\alpha)T \\ 0 & \text{for } t \in ((1+\alpha)T, 4T]. \end{cases}$$

and $\nabla\Gamma^2$ is equal to

$$\begin{cases} 0 & \text{for } t \in [0, T/8) \\ \chi^2(t) + \chi^1(t) - \chi^2(T/8) - \chi^1(T/8) & \text{for } t \in [T/8, T/4) \\ \chi^2(t) + \chi^1(T/4) - \chi^2(T/8) - \chi^1(T/8) & \text{for } t \in [T/4, (1+\alpha)T) \\ \chi^2(t) + [\chi^1(T/4) \wedge \chi^1((1+\alpha)T)] - \chi^2(T/8) - \chi^1(T/8) & \text{for } t = (1+\alpha)T \\ \chi^2(t) + \chi^1(t) - \chi^2(T/8) - \chi^1(T/8) & \text{for } t \in ((1+\alpha)T, (3+\beta)T) \\ 0 \vee [\chi^2((3+\beta)T) + \chi^1((3+\beta)T) - \chi^2(T/8) - \chi^1(T/8)] & \text{for } t = (3+\beta)T \\ 0 & \text{for } t \in ((3+\beta)T, 4T]. \end{cases}$$

We will refer to the general classification of discontinuities presented in Theorem 1.2 in order to describe the discontinuities in the derivative of the particular tandem model considered in this section. Note that queue 1 is both at the end of overloading and at the start of underloading at time $t = (1+\alpha)T$. Thus at that time queue 2, though overloaded, has both a critical and sub-critical chain preceding it (see Definition 5.2 for a precise description of chains). From the above explicit expressions it is straightforward to see that $\nabla\Gamma^1$ has a left discontinuous upward jump of type (La) at $t = (1+\alpha)T$ if

$$\chi^1(T/4) > \chi^1((1+\alpha)T) \tag{2.3}$$

and exhibits a right continuous jump downward of type (Ra) if the opposite inequality holds. At the same time, $\nabla\Gamma^1$ exhibits a left discontinuous jump downward at $t = (1+\alpha)T$ of type (Lb) if the inequality (2.3) holds and a right discontinuous upward jump of type (Rb) if the reverse inequality is satisfied. Moreover, note that queue 2 is at the end of overloading and at the start of underloading at $t = (3+\beta)T$. Thus at $t = (3+\beta)T$, $\nabla\Gamma^2$ has a left discontinuous downward jump of type (Lb) if the inequality

$$\chi^1((3+\beta)T) + \chi^2((3+\beta)T) < \chi^1(T/8) - \chi^2(T/8)$$

is satisfied, and a right discontinuous jump upward of type (Rb) if the reverse inequality is satisfied.

2.2 The Merge

2.2.1 Description of the Merge Model

We now consider a scenario in which two upstream queues feed into a common buffer (see Figure 4). The upstream queues experience a surge in arrival rate for an initial period, which then subsides to a lower rate. However, just as the surge ends, the server at queue 2 undergoes a partial failure, resulting in the queue maintaining criticality. We show that in such a scenario there can be a discontinuity in the derivative of the downstream queue at the time congestion ends in the upstream queues.

We assume that queues 1 and 3 have constant service rates $\mu^1(t) = \mu^3(t) = 1$ for $t \in [0, 2]$, while queue 2 has service rate

$$\mu^2(t) = \begin{cases} 1 & \text{for } t \in [0,1] \\ 1/3 & \text{for } t \in [1,2] \end{cases} \tag{2.4}$$

The departures from queues 1 and 2 feed into queue 3. (see Figure 4). The exogenous arrivals to queues are modeled by non-homogeneous Poisson processes with mean arrival rates $\lambda \in \mathcal{C}$, where

$$\begin{aligned} \lambda^1(t) &= \begin{cases} 1 & \text{for } t \in [0,1] \\ 1/2 & \text{for } t \in (1,2]. \end{cases} \\ \lambda^2(t) &= \begin{cases} 3/2 & \text{for } t \in [0,1/2) \\ 1/2 & \text{for } t \in [1/2,1) \\ 1/3 & \text{for } t \in [1,2], \end{cases} \\ \lambda^3(t) &= 0 \text{ for } t \in [0,2] \end{aligned}$$

The routing and constraint matrices P' and R respectively are given by

$$P' = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad \text{and} \qquad R = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

It is trivial to verify that P is a H-R matrix. Let Γ denote the associated reflection map.

2.2.2 The Fluid Limit

Let the netput process ψ be constructed in the usual way:

$$\psi^i(t) = \int_0^t \left(\lambda^i(s) - \mu^i\right) ds \quad \text{for } i = 1, 2$$

and

$$\psi^3(t) = \int_0^t \left(\mu^1 + \mu^2(s) - \mu^3\right) ds.$$

Let $\phi = \Gamma(\psi)$ be the fluid queue. Then it is easy to verify that

$$\begin{aligned} \phi^1(t) &= 0 \text{ for } t \in [0,2] \\ \theta^1(t) &= \begin{cases} 0 & \text{for } t \in [0,1) \\ 1 + \frac{1}{2}(t-1) & \text{for } t \in [1,2]. \end{cases} \\ \phi^2(t) &= \begin{cases} \frac{1}{2}t & \text{for } t \in [0,1/2) \\ \frac{1}{4} - \frac{1}{2}(t-1/2) & \text{for } t \in [1/2,1) \\ 0 & \text{for } t \in [1,2]. \end{cases} \\ \theta^2(t) &= 0 \text{ for } t \in [0,2] \\ \phi^3(t) &= \begin{cases} t & \text{for } t \in [0,1) \\ 1 - \frac{1}{6}(t-1) & \text{for } t \in [1,2] \end{cases} \end{aligned}$$

See Figure 5 for an illustration of the fluid limit of the three queues.

2.2.3 Directional Derivatives for the Merge Model

It is easy to see that

$$\begin{aligned}
\Phi^1(t) &= \begin{cases} [0,t] & \text{for } t \in [0,1] \\ \{t\} & \text{for } t \in (1,2] \end{cases} \\
\Phi^2(t) &= \begin{cases} \{0\} & \text{for } t \in [0,1) \\ \{0\} \cup [1,t] & \text{for } t \in [1,2] \end{cases} \\
\Phi^3(t) &= \{0\} \quad \text{for } t \in [0,2]
\end{aligned}$$

Moreover, from Theorem 1.1 it follows that

$$\begin{aligned}
\nabla\Gamma^1(t) &= \chi^1(t) + \sup_{s\in\Phi^1(t)} \left[-\chi^1(s)\right]. \\
\nabla\Gamma^2(t) &= \chi^2(t) + \sup_{s\in\Phi^2(t)} \left[-\chi^2(s)\right] \\
\nabla\Gamma^3(t) &= \chi^3(t) - \sup_{s\in\Phi^2(t)} \left[-\chi^2(s)\right] - \sup_{s\in\Phi^1(t)} \left[-\chi^1(s)\right] \\
&\quad + \sup_{s\in\Phi^3(t)} \left[-\chi^3(s) + \sup_{r\in\Phi^2(s)} \left[-\chi^2(r)\right] + \sup_{r\in\Phi^1(s)} \left[-\chi^1(r)\right]\right]
\end{aligned}$$

Hence it follows that if

$$\chi^2(1) < \chi^2(0) \quad \text{and} \quad \sup_{s\in[0,1]} \left[-\chi^1(s)\right] > -\chi^1(1)$$

then

$$\begin{aligned}
\nabla\Gamma^3(1) - \nabla\Gamma^3(1-) = \chi^2(1) - \chi^2(0) &< 0 \\
\nabla\Gamma^3(1) - \nabla\Gamma^3(1+) = -\sup_{s\in[0,1]} \left[-\chi^1(s)\right] - \chi^1(1) &< 0,
\end{aligned}$$

which implies $\nabla\Gamma^1$ has a separated discontinuity (i.e. it is neither right continuous or left continuous) at $t = 1$, and moreover observe that

$$\nabla\Gamma^3(1) < \nabla\Gamma^3(1-) \wedge \nabla\Gamma^3(1+)$$

as expected from (LRc) of Theorem 1.2. It is significant that the separated discontinuity arises only in the multi-dimensional setting, and not in the one-dimensional setting. This has important ramifications for the convergence of $\nabla^\varepsilon_\chi(\psi)$ to $\nabla_\chi\Gamma(\psi)$ when ψ, χ are continuous. Specifically, as remarked earlier, it was shown in [22] that the convergence of $\nabla^\varepsilon_\chi(\psi)$ to $\nabla_\chi\Gamma(\psi)$ takes place in the M_1 topology [45]. When ψ, χ are continuous, $\nabla^\varepsilon_\chi(\psi)$ is also continuous for every $\varepsilon > 0$. Since the space $\mathcal{D}_{l,r}$, the space of functions that are either left or right continuous at every point, is complete under the M_1 topology [45], and continuous functions clearly lie in $\mathcal{D}_{l,r}$, while functions with separated discontinuities do not lie in $\mathcal{D}_{l,r}$, it is clear that one cannot in general expect M_1 convergence in the multi-dimensional case. Indeed, as shown in [25], in the network setting under suitable assumptions convergence takes place in the (weak) M_2 topology, but not in general in the M_1 topology.

3 The Multi-dimensional Reflection Map

In this section we give a precise definition of directional derivatives of the oblique reflection map. In Section 3.1 we recall the definition of the oblique reflection problem and the reflection map, in Section 3.2 we summarize known properties of the so-called Harrison-Reiman class of oblique reflection maps and in Section 3.3 we introduce directional derivatives.

3.1 The Oblique Reflection Problem

Let $I\!\!R^K_+$ be the positive K-dimensional orthant given by

$$I\!\!R^K_+ \doteq \{x \in I\!\!R^K : x^i \geq 0 \text{ for every } i = 1, \ldots, K\}.$$

Let $R \in I\!\!R^{K \times K}$ be a matrix whose ith column is the vector r_i, which represents the constraint direction on the face $F_i = \{x \in I\!\!R^K_+ : x^i = 0\}$ of the boundary of the orthant. Roughly speaking, given a trajectory $\psi \in \mathcal{D}_{\text{lim}}$, the oblique reflection problem (ORP) associated with the constraint matrix R defines a constrained version ϕ of ψ that is restricted to live in $I\!\!R^K_+$ by a constraining term that pushes along the direction r_i only when ϕ lies on the face F_i. The rigorous definition of the ORP is as follows. Here $\mathcal{D}^+_{\text{lim}}$ represents the class of functions ψ in $\mathcal{D}_{\text{lim}}$ that satisfy $\psi(0) \in I\!\!R^K_+$ and $\mathcal{I}_0$ the subspace of functions in $\mathcal{D}_{\text{lim}}$ that have $f(0) = 0$ and each co-ordinate non-decreasing.

Definition 3.1 (Oblique Reflection Problem) *Given $R \in R^{K \times K}$, $\psi \in \mathcal{D}^+_{\text{lim}}$, $(\phi, \theta) \in \mathcal{D}^+_{\text{lim}} \times \mathcal{I}_0$ solve the oblique reflection problem associated with the constraint matrix R for ψ if $\phi(0) = \psi(0)$, and if for all $t \in [0, \infty)$*

1. *$\phi(t) \in I\!\!R^K_+$;*
2. *$\phi(t) = \psi(t) + R\theta(t)$, where for every $i = 1, \ldots, K$*

$$\int_0^t 1_{(0,\infty)}(\phi^i(s))d\theta^i(s) = 0. \tag{3.1}$$

Note that the condition (3.1) simply states that the constraining term θ^i can increase at time t only if $\phi^i(t) = 0$. From the definition above it is clear that one can without loss of generality assume that $R_{ii} = 1$ for $i = 1, \ldots, K$. Indeed, we shall assume this normalization throughout the rest of the paper.

In this work we exclusively consider oblique reflection problems (ORPs) that satisfy Assumption 3.1 introduced in Section 3.2, which in particular guarantees that whenever a solution to the ORP exists for some $\psi \in \mathcal{D}^+_{\text{lim}}$, it is unique. We refer to the mapping $\Gamma : \psi \to \phi$ as the *reflection map* (RM), and use $\Theta : \psi \to \theta$ to denote the mapping that takes ψ to the corresponding constraining term θ.

Remark 3.1 The ORP was introduced in [14] to characterize functional central limits of single-class open queueing networks (see Figure 3). Single-class open queueing networks with K queues in which a fraction p'_{ij} of the departures from queue i are sent to queue j, and a fraction $1 - \sum_{j=1}^K p'_{ij}$ of the departures from queue i exit the network give rise to ORPs with a $I\!\!R^{K \times K}$ constraint matrix R given by

$$R_{ij} \doteq \begin{cases} -p'_{ji} & \textit{for } j \neq i \\ 1 & \textit{otherwise.} \end{cases}$$

3.2 Special Classes of Reflection Maps

It is well known that a necessary and sufficient condition for there to be solutions to the ORP is that the constraint matrix R satisfy the so-called completely-$\mathcal{S}$ condition [3, 27]. However, the derivative cannot be expected to exist if the RM is not even continuous. Thus all RMs considered in this paper are assumed to satisfy the following regularity condition.

Assumption 3.1 *(Lipschitz continuity of the RM) The matrix R of the ORP is completely-$\mathcal{S}$ and is such that there exists $L < \infty$ for which given solutions (ϕ_i, θ_i) to the ORP for $\psi_i \in \mathcal{D}^+_{\lim}$, $i = 1, 2$,*

$$||\phi_1 - \phi_2|| \leq L||\psi_1 - \psi_2|| \qquad \textit{and} \qquad ||\theta_1 - \theta_2|| \leq L||\psi_1 - \psi_2||.$$

The above assumption implies in particular that a unique solution to the ORP exists for every $\psi \in \mathcal{D}^+_{\lim}$, and consequently the RM Γ and the map Θ associated with the ORP satisfy

$$||\Gamma(\psi_1) - \Gamma(\psi_2)|| \leq L||\psi_1 - \psi_2|| \quad \text{and} \quad ||\Theta(\psi_1) - \Theta(\psi_2)|| \leq L||\psi_1 - \psi_2||. \tag{3.2}$$

We largely focus on a subset of ORPs that satisfy the additional regularity conditions described below in Definition 3.2. These ORPs have an additional monotonicity property that proves useful for the characterization of directional derivatives of their RMs. We refer to this class of RMs as the Harrison-Reiman class since they were introduced by Harrison and Reiman in [14]. We also refer to a slight generalization of the H-R class as gH-R (or generalized Harrison-Reiman). Recall that I is the $K \times K$ identity matrix and by convention $R_{ii} = 1$ for $i = 1, \ldots, K$.

Definition 3.2 *(H-R and gH-R condition) A constraint matrix $R \in R^{K\times K}$ is said to satisfy the gH-R condition if the spectral radius of the matrix $P = |I - R|$ is less than one. Moreover, R is said to satisfy the H-R condition if it satisfies the gH-R condition, and in addition $P = I - R \geq 0$.*

We now state a minor generalization of a well-known result about ORPs in the gH-R class. Recall the notation $\overline{f}(t) = \sup_{s\in[0,t]} f(s)$.

Theorem 3.2 (Harrison & Reiman [14]) *Consider an ORP with constraint matrix R. If R satisfies the H-R or gH-R condition, then it also satisfies Assumption 3.1. Moreover, if $P \doteq I - R$, then given $\psi \in \mathcal{D}^+_{\lim}$, $\theta = \Theta(\psi)$ if and only if for $i = 1, \ldots, K$ and $t \in [0, \infty)$*

$$\theta^i(t) = \left[\overline{-\psi^i + [P\theta]^i}\ (t)\right] \vee 0\,. \tag{3.3}$$

In other words, θ is the unique fixed point of the map $F(\psi, \cdot) : \mathcal{I}_0 \to \mathcal{I}_0$ given by

$$F^i(\psi, \theta)(t) \doteq \left[\overline{-\psi^i + [P\theta]^i}(t)\right] \vee 0 \qquad \textit{for } i = 1, \ldots, K. \tag{3.4}$$

Proof. It is easy to verify that $F(\psi, \theta) \in \mathcal{I}_0$ whenever $(\psi, \theta) \in \mathcal{D}^+_{\lim} \times \mathcal{I}_0$. Since $\mathcal{D}_{\lim}$ is complete with respect to the sup norm, the same argument as that used in [14] shows that $F(\psi, \cdot)$ is a contraction mapping that maps $\mathcal{I}_0$ into $\mathcal{I}_0$, and thus has a unique fixed point. It only remains to show that θ is a fixed point of $F(\psi, \cdot)$ if and only if $\theta = \Theta(\psi)$. This also follows from a straightforward generalization of the argument used in [14] for functions in $\mathcal{C}$ to functions in $\mathcal{D}^+_{\lim}$, and is thus omitted. ■

The following simple fact will be used in the sequel without explicit reference. Note that for any $(\psi, \theta) \in \mathcal{D}^+_{\lim} \times \mathcal{I}_0$, $F(\psi, \theta) = F(\psi - P\theta, 0)$. Therefore given the fixed point θ that satisfies $F(\psi, \theta) = \theta$, one can also write

$$F(\psi - P\theta, 0) = F(\psi, \theta) = \theta\,.$$

3.3 Directional Derivatives of the Reflection Map

Given a RM Γ that satisfies Assumption 3.1, we define directional derivatives of Γ as follows.

Definition 3.3 *Consider an oblique reflection problem whose reflection map is Γ. Given paths $\psi \in \mathcal{D}^{+}_{\lim}, \chi \in \mathcal{D}_{\lim}$, the derivative of Γ along χ evaluated at ψ is the pointwise limit of the sequence $\{\nabla^{\varepsilon}_{\chi}\Gamma(\psi)\}$, as $\varepsilon \downarrow 0$, where $\nabla^{\varepsilon}_{\chi}\Gamma(\psi)$ is defined by*

$$\nabla^{\varepsilon}_{\chi}\Gamma(\psi) \doteq \frac{1}{\varepsilon}\left[\Gamma(\psi + \varepsilon\chi) - \Gamma(\psi)\right]. \tag{3.5}$$

4 Characterization of the Derivative

In Section 4.1 we show that the derivative introduced in Definition 3.3 exists and is uniquely defined for any $\psi \in \mathcal{D}^{+}_{\lim}, \chi \in \mathcal{D}_{\lim}$ when the ORP is of H-R type (see Theorem 4.6). Under additional regularity conditions on ψ and χ, in Section 4.2 we derive autonomous characterizations of the derivative.

4.1 Existence of the derivative $\nabla_{\chi}\Gamma(\psi)$

In order to show the existence of the derivative or, equivalently, to show the existence of a unique pointwise limit of the sequence $\nabla^{\varepsilon}_{\chi}(\psi) \in \mathcal{D}_{\lim}$ as $\varepsilon \downarrow 0$, it turns out to be more convenient to work with a closely related sequence $\{\gamma_\varepsilon(\psi,\chi)\}$. This sequence is introduced in Section 4.1.1 and is shown to have a unique pointwise limit $\gamma(\psi,\chi)$ in Section 4.1.2. In Section 4.1.3 the limit $\gamma(\psi,\chi)$ and the derivative $\nabla_{\chi}\Gamma(\psi)$ are shown to lie in $\mathcal{D}_{\lim}$ if the ORP is of H-R type.

4.1.1 A related sequence $\{\gamma_\varepsilon\}$

Given an ORP with constraint matrix R recall the definition of the mapping Θ following Definition 3.1 in Section 3.1 and let

$$\gamma_\varepsilon(\psi,\chi) \doteq \varepsilon^{-1}\left[\Theta(\psi + \varepsilon\chi) - \Theta(\psi)\right]. \tag{4.1}$$

Using the fact that $\Gamma(\psi) = \psi + R\Theta(\psi)$ for $\psi \in \mathcal{D}_{\lim}$, along with definition (3.5) of the sequence $\{\nabla^{\varepsilon}_{\chi}\Gamma(\psi)\}$ one obtains the relation

$$\nabla^{\varepsilon}_{\chi}\Gamma(\psi) = \chi + R\gamma_\varepsilon(\psi,\chi). \tag{4.2}$$

Thus, in order to establish the existence of and characterize the derivative, it clearly suffices to examine the limiting behaviour of $\gamma_\varepsilon(\psi,\chi)$ as $\varepsilon \downarrow 0$.

Now suppose that the ORP is of gH-R type and let $P \doteq |I - R|$. Then, as stated in Theorem 3.2, Assumption 3.1 is satisfied and so the associated RM Γ is well-defined on $\mathcal{D}_{\lim}$ and is Lipschitz continuous. Fix $\psi, \chi \in \mathcal{D}_{\lim}$. For conciseness let $\theta \doteq \Theta(\psi)$ and for $\varepsilon > 0$, let $\theta_\varepsilon \doteq \Theta(\psi + \varepsilon\chi)$ and $\gamma_\varepsilon \doteq \gamma_\varepsilon(\psi,\chi)$. Then from (3.3) it follows that

$$\theta^i = \left[\overline{-\psi^i + [P\theta]^i}\right] \vee 0, \tag{4.3}$$

and likewise

$$\theta^i_\varepsilon = \left[\overline{-\psi^i - \varepsilon\chi^i + [P\theta_\varepsilon]^i}\right] \vee 0.$$

Adding and subtracting $[P\theta]^i$ under the supremum on the right side of the above display, and introducing the function

$$\xi^i \doteq \psi^i - [P\theta]^i\,, \tag{4.4}$$

one can rewrite

$$\theta^i_\varepsilon = \overline{\left[-\xi^i - \varepsilon\chi^i + [P(\theta_\varepsilon - \theta)]^i\right]} \vee 0. \tag{4.5}$$

Multiplying the difference between (4.5) and (4.3) by ε^{-1}, and using the definitions of F and γ_ε in (3.4) and (4.1) respectively, one infers that for $i = 1, \ldots, K$

$$\begin{array}{rcl} \gamma^i_\varepsilon = \gamma^i_\varepsilon(\psi,\chi) & = & \overline{-\varepsilon^{-1}\xi^i - \chi^i + [P\gamma_\varepsilon]^i} \vee 0 - \overline{-\varepsilon^{-1}\xi^i} \vee 0 \\ & = & F^i\left(\varepsilon^{-1}\xi + \chi, \gamma_\varepsilon\right) - F^i\left(\varepsilon^{-1}\xi, 0\right) \\ & = & F^i\left(\varepsilon^{-1}\xi + \chi - P\gamma_\varepsilon, 0\right) - F^i\left(\varepsilon^{-1}\xi, 0\right). \end{array} \tag{4.6}$$

4.1.2 Pointwise convergence of $\{\gamma_\varepsilon\}$ for H-R ORPs

In this section some basic properties of the sequences $\{\gamma_\varepsilon\}$ and $\{\nabla^\varepsilon_\chi\Gamma(\psi)\}$ are established. Lemma 4.1 proves the uniform boundedness of these two sequences for ORPs that satisfy Assumption 3.1, and Lemma 4.2 proves a crucial monotonicity property of the sequence $\{\gamma_\varepsilon\}$ for ORPs in the H-R class.

Lemma 4.1 (Uniform Boundedness) *Given an ORP that satisfies Assumption 3.1 and has an invertible constraint matrix R let $\nabla^\varepsilon_\chi\Gamma(\psi)$ and $\gamma_\varepsilon(\psi,\chi)$ be defined by (3.5) and (4.1) respectively. Then there exists $L < \infty$ such that for any $\xi, \chi_1, \chi_2 \in \mathcal{D}_{\lim}$ and $T < \infty$*

$$\sup_{\varepsilon>0} ||\nabla^\varepsilon_{\chi_1}\Gamma(\psi) - \nabla^\varepsilon_{\chi_2}\Gamma(\psi)||_T \leq L||\chi_1 - \chi_2||_T,$$

and

$$\sup_{\varepsilon>0} ||\nabla^\varepsilon_\chi\Gamma(\psi)||_T \leq L||\chi||_T.$$

In addition $L < \infty$ can be chosen to also satisfy

$$\sup_{\varepsilon>0} ||\gamma_\varepsilon(\psi,\chi_1) - \gamma_\varepsilon(\psi,\chi_2)||_T \leq L||\chi_1 - \chi_2||_T\,,$$

$$\sup_{\varepsilon>0} ||\gamma_\varepsilon(\psi,\chi)||_T \leq L||\chi||_T\,. \tag{4.7}$$

Proof. The first inequality follows directly from the Lipschitz continuity of the RM stated in Assumption 3.1 and the definiton of $\nabla_\chi\Gamma(\psi)$. The third inequality follows from the first inequality and the fact that the norm of the matrix R is bounded away from zero since R is invertible. The second and fourth bounds follow simply by choosing $\chi_1 = \chi$ and $\chi_2 = 0$ in the first and third bounds respectively and noting that $\nabla^\varepsilon_0\Gamma(\psi) = \gamma_\varepsilon(\psi, 0) = 0$, where the zeros here represent the function that is identically zero. The last inequality is a direct consequence of the third. ■

The uniform boundedness property proved in Lemma 4.1 shows that for every $t > 0$ the sequence $\{\gamma_\varepsilon(t)\}$ must have a convergent subsequence. In the next lemma we establish an additional monotonicity property that holds for H-R ORPs, which leads to the conclusion in Corollary 4.3 that the sequence $\{\gamma_\varepsilon(\psi,\chi)\}$ has a unique pointwise limit as $\varepsilon \downarrow 0$.

Lemma 4.2 (Monotonicity) *Given an H-R ORP and $\psi, \chi \in \mathcal{D}_{\lim}$, let $\gamma_\varepsilon \doteq \gamma_\varepsilon(\psi, \chi)$ be defined by (4.1). Then each coordinate function γ_ε^i is monotonically nonincreasing as $\varepsilon \downarrow 0$. In other words for $i = 1, \ldots, K$,*

$$0 < \varepsilon_1 \leq \varepsilon_2 \qquad \textit{implies} \qquad \gamma_{\varepsilon_1}^i(s) - \gamma_{\varepsilon_2}^i(s) \leq 0 \quad \textit{for } s \in [0, \infty). \tag{4.8}$$

Proof. Let $0 < \varepsilon_1 \leq \varepsilon_2$ and fix $i \in \{1, \ldots, K\}$ and $s \in [0, \infty)$. Using the representation (4.6) for γ_ε^i and making repeated use of Lemma A.1, it follows that for $0 \leq \varepsilon_1 \leq \varepsilon_2$

$$\begin{aligned}
\gamma_{\varepsilon_1}^i(t) - \gamma_{\varepsilon_2}^i(t) &= \overline{-\varepsilon_1^{-1}\xi^i - \chi^i + [P\gamma_{\varepsilon_1}]^i}(t) \vee 0 - \overline{\varepsilon_1^{-1}\xi^i}(t) \vee 0 \\
&\qquad - \overline{-\varepsilon_2^{-1}\xi^i - \chi^i + [P\gamma_{\varepsilon_2}]^i}(t) \vee 0 + \overline{\varepsilon_2^{-1}\xi^i}(t) \vee 0 \\
&= \overline{-\varepsilon_1^{-1}\xi^i - \chi^i + [P\gamma_{\varepsilon_1}]^i}(t) \vee 0 - \overline{-\varepsilon_2^{-1}\xi^i - \chi^i + [P\gamma_{\varepsilon_2}]^i}(t) \vee 0 \\
&\qquad -(\varepsilon_1^{-1} - \varepsilon_2^{-1}) \left[\overline{-\xi^i}(t) \vee 0\right] \\
&\leq \overline{-(\varepsilon_1^{-1} - \varepsilon_2^{-1})\xi^i + [P\gamma_{\varepsilon_1}]^i - [P\gamma_{\varepsilon_2}]^i}(t) \vee 0 - \overline{-(\varepsilon_1^{-1} - \varepsilon_2^{-1})\xi^i} \vee 0 \\
&\leq \overline{[P\gamma_{\varepsilon_1}]^i - [P\gamma_{\varepsilon_2}]^i}(t) \vee 0.
\end{aligned}$$

Hence, using the nonnegativity of P, observe that for $t \in [0, s]$

$$\begin{aligned}
\gamma_{\varepsilon_1}^i(t) - \gamma_{\varepsilon_2}^i(t) \qquad &\leq \left[\sum_{j=1}^K P_{ij} \overline{\gamma_{\varepsilon_1}^j - \gamma_{\varepsilon_2}^j}(t)\right] \vee 0 \\
&\leq \left[\sum_{j=1}^K P_{ij} \max_{k=1,\ldots,K} \overline{\gamma_{\varepsilon_1}^k - \gamma_{\varepsilon_2}^k}(s)\right] \vee 0.
\end{aligned}$$

First consider the case when P is substochastic, so that there exists $\delta > 0$ that satisfies $\max_{i=1,\ldots,K} \sum_{j=1}^K P_{ij} \leq 1 - \delta$. If $\max_{k=1,\ldots,K} \overline{\gamma_{\varepsilon_1}^k - \gamma_{\varepsilon_2}^k}(s) \leq 0$, the above display automatically implies that $\gamma_{\varepsilon_1}(t) = \gamma_{\varepsilon_2}(t)$ for every $t \in [0, s]$ and the lemma holds. Therefore we suppose that

$$\max_{k=1,\ldots,K} \overline{\gamma_{\varepsilon_1}^k - \gamma_{\varepsilon_2}^k}(s) > 0. \tag{4.9}$$

Then taking the supremum over $t \in [0, s]$ on the left side and then the maximum over i on both sides of the last display, results in the inequality

$$\max_{i=1,\ldots,K} \overline{\gamma_{\varepsilon_1}^i - \gamma_{\varepsilon_2}^i}(s) \leq (1 - \delta) \max_{i=1,\ldots,K} \overline{\gamma_{\varepsilon_1}^i - \gamma_{\varepsilon_2}^i}(s),$$

which implies that

$$\max_{i=1,\ldots,K} \overline{\gamma_{\varepsilon_1}^i - \gamma_{\varepsilon_2}^i}(s) \leq 0,$$

and therefore contradicts the assumption (4.9). Thus it must be that

$$\max_{k=1,\ldots,K} \overline{\gamma_{\varepsilon_1}^k - \gamma_{\varepsilon_2}^k}(s) \leq 0,$$

and the lemma is established for the case when P is substochastic.

Since the H-R condition implies that P has spectral radius less than one, P is similar to a substochastic matrix through a strictly positive diagonal transformation [44], and so the proof for the general H-R case can be obtained from the substochastic case using diagonal similarity transforms (in a manner analogous to the proof of Lemma 4.4). ■

In the following corollary, $\mathcal{D}_{usc}$ denotes the space of upper semicontinuous functions in $\mathcal{D}_{\lim}$.

Corollary 4.3 *Given a H-R ORP and $\psi, \chi \in \mathcal{D}_{\lim}$, there exists a unique pointwise limit γ of the sequence $\{\gamma_\varepsilon\}$, which is a bounded $\mathbb{R}^K$-valued function. Moreover, if $\xi, \chi \in \mathcal{C}$, then $\gamma \in \mathcal{D}_{usc}$ and the convergence of γ_ε to γ is uniform over compact subsets of the set of continuity points of γ. In particular, given any continuity point t of γ, $\lim_{\varepsilon \downarrow 0} \gamma_\varepsilon(t_\varepsilon) = \gamma(t)$ for any sequence $t_\varepsilon \to t$.*

Proof. The uniform boundedness of the sequence $\{\gamma_\varepsilon\}$ proved in Lemma 4.1 shows that for each $s \in [0, \infty)$, there exists a subsequence (which could depend on s) of $\{\gamma_\varepsilon(s)\}$ that converges to a limit. The monotonicity property shows that this limit is independent of the subsequence, and by (4.7) of Lemma 4.1 the limit is finite. If $\xi, \chi \in \mathcal{C}$, then for each $i = 1, \ldots, K$, γ^i is the limit of a non-increasing sequence of continuous functions and hence must be upper semicontinuous [40, p. 196]. The last two statements of the corollary follow from Dini's Theorem [40, p. 195], which states that the monotone convergence of continuous functions to a limit must be uniform on compact subsets of continuity points of the limit. ■

4.1.3 The limit γ of the sequence $\{\gamma_\varepsilon\}$ lies in $\mathcal{D}_{\lim}$

Corollary 4.3 established the existence of a pointwise limit γ for the sequence $\{\gamma_\varepsilon\}$ for H-R ORPs and $\psi, \chi \in \mathcal{D}_{\lim}$. However, even if ψ and χ are assumed to be continuous (which would in turn imply that the functions $\gamma_\varepsilon = \gamma_\varepsilon(\psi, \chi), \varepsilon > 0$, are continuous), it is not a priori obvious that the pointwise limit γ of the sequence $\{\gamma_\varepsilon\}$ would lie in $\mathcal{D}_{\lim}$. In fact, as the example below illustrates, the limit of a monotone non-increasing sequence of real-valued continuous functions $\{f_n\}$ need not in general lie in $\mathcal{D}_{\lim}$. However, Lemma 4.4 and Corollary 4.5 exploit the special structure that the sequence $\{\gamma_\varepsilon\}$ possesses by virtue of the fact that it is defined via an ORP in order to prove that γ must lie in $\mathcal{D}_{\lim}$. As shown in Theorem 4.6, this leads to the existence of a well-defined derivative for H-R ORPs.
Example. Consider the sequence of functions

$$f_n(x) = \begin{cases} \sin \frac{1}{|x|} & \text{if } |x| \in [1/(2n\pi + \pi/2), \infty), \\ 1 & \text{if } |x| \in [0, 1/(2n\pi + \pi/2). \end{cases}$$

It is easy to verify that $\{f_n\}$ forms a monotone decreasing sequence of continuous functions, whose pointwise limit is

$$f(x) = \begin{cases} \sin \frac{1}{|x|} & \text{if } x \neq 0, \\ 1 & \text{if } x = 0. \end{cases}$$

The pointwise limit f clearly does not lie in $\mathcal{D}_{\lim}$ since it has neither a left nor right limit at $x = 0$.

The property proved in the next lemma, unlike the monotonicity property established in Lemma 4.2, is not restricted to H-R ORPs, but holds for all gH-R ORPs. Recall that $|f|_T$ denotes the total variation of the function f on the interval $[0, T]$ with respect to the norm $|\cdot|$ on $\mathbb{R}^K$ defined in (1.26).

Lemma 4.4 (Uniformly BV) *Given a gH-R ORP, $\psi \in \mathcal{D}_{\lim}$ and $\chi \in \mathcal{BV}$, let $\gamma_\varepsilon \doteq \gamma_\varepsilon(\psi, \chi)$ be defined by (4.1). Then for every $T \in [0, \infty)$*

$$\sup_{\varepsilon > 0} |\gamma_\varepsilon|_T < \infty. \tag{4.10}$$

Proof. Fix $T < \infty$. Using the representation (4.6) for γ_ε^i and Lemma A.1 it follows that for $\varepsilon > 0$

$$\begin{aligned} \left|\gamma_\varepsilon^i\right|_{\rm T} &= \left|\overline{-\varepsilon^{-1}\xi^i - \chi^i + [P\gamma_\varepsilon]^i} \vee 0 - \overline{-\varepsilon^{-1}\xi^i} \vee 0\right|_{\rm T} \\ &\leq \left|-\chi^i + [P\gamma_\varepsilon]^i\right|_{\rm T} \\ &\leq \left|\chi^i\right|_{\rm T} + \textstyle\sum_{j=1}^K |P_{ij}| \left|\gamma_\varepsilon^j\right|_{\rm T}. \end{aligned}$$

Since the ORP is gH-R, the spectral radius of $|P|$ is less than one. Thus there exists a strictly positive diagonal matrix A (with $a_i \doteq A_{ii} > 0$) such that $\tilde{P} \doteq A^{-1}|P|A$ is a nonnegative substochastic matrix, i.e. there exists $\delta > 0$ such that $\max_{i=1,\ldots,K} \sum_{j=1}^K \tilde{P}_{ij} \leq 1 - \delta$. Multiplying both sides of the last display by a_i and substituting for $|P|$ in terms of $\tilde{P}$ (note that $a_j \tilde{P}_{ij} = a_i |P_{ij}|$), yields the inequality

$$a_i \left|\gamma_\varepsilon^i\right|_{\rm T} \leq a_i \left|\chi^i\right|_{\rm T} + \sum_{j=1}^K \tilde{P}_{ij} a_j \left|\gamma_\varepsilon^j\right|_{\rm T},$$

which implies that

$$\max_i a_i \left|\gamma_\varepsilon^i\right|_{\rm T} \leq \max_i a_i \left|\chi^i\right|_{\rm T} + (1-\delta) \max_i a_i \left|\gamma_\varepsilon^i\right|_{\rm T},$$

which in turn can be rearranged to obtain

$$\max_i a_i |\gamma_\varepsilon^i|_{\rm T} \leq \frac{\max_i a_i |\chi^i|_{\rm T}}{\delta}.$$

Thus

$$|\gamma_\varepsilon|_{\rm T} \leq K \max_i |\gamma_\varepsilon^i|_{\rm T} \leq \frac{K \max_i a_i}{\delta \min_i a_i} |\chi|_{\rm T} < \infty,$$

where the last inequality follows because of the assumption that $\chi \in \mathcal{BV}$. ■

Corollary 4.5 *Consider a gH-R ORP, $\psi, \chi \in \mathcal{D}_{\rm lim}$ and $\gamma^\varepsilon \doteq \gamma^\varepsilon(\psi, \chi)$ defined by (4.1). If there exists a subsequence of $\{\gamma_\varepsilon\}$ that converges pointwise to a limit function γ, then γ lies in $\mathcal{D}_{\rm lim}$.*

Proof. Fix $\psi \in \mathcal{D}_{\rm lim}$ and first consider the case when $\chi \in \mathcal{BV}$. By Lemma 4.4 it is clear that for any gH-R ORP and $T \in [0, \infty)$ the total variations of the functions γ_ε, $\varepsilon > 0$, over the interval $[0, T]$ are uniformly bounded. The fact that any limit γ of a subsequence of $\{\gamma_\varepsilon\}$ is uniformly bounded follows from (4.7) of Lemma 4.1. When combined with Lemma A.2, this shows that $\gamma \in \mathcal{D}_{\rm lim}$. Now let $\chi \in \mathcal{D}_{\rm lim}$. Since χ can be approximated in the uniform norm by a sequence $\{\chi_n\} \subset \mathcal{BV}$, from Lemma 4.1 it follows that $\gamma(\psi, \chi_n)$ converges to $\gamma(\psi, \chi)$ in the uniform norm. Since $\mathcal{D}_{\rm lim}$ is complete with respect to the uniform norm and, as just proved above, $\gamma(\psi, \chi_n) \in \mathcal{D}_{\rm lim}$ for every n, clearly $\gamma(\psi, \chi) \in \mathcal{D}_{\rm lim}$. ■

Corollaries 4.3 and 4.5 and the relation (4.2) together yield the following theorem.

Theorem 4.6 (Existence of the Derivative for H-R ORPs) *Given a H-R ORP with RM Γ, constraint matrix R and $\psi, \chi \in \mathcal{D}_{\rm lim}$, there exists a unique pointwise limit $\gamma(\psi, \chi)$ of the sequence $\{\gamma_\varepsilon(\psi, \chi)\}$ defined in (4.1). In addition, $\nabla_\chi \Gamma(\psi) \doteq \chi + R\gamma(\psi, \chi)$ is the directional derivative of Γ along χ, evaluated at the point ψ, and $\nabla_\chi \Gamma(\psi) \in \mathcal{D}_{\rm lim}$. Furthermore, if $\psi, \chi \in \mathcal{C}$ then the convergence of $\nabla_\chi^\varepsilon \Gamma(\psi)$ to $\nabla_\chi \Gamma(\psi)$ is uniform over compact subsets of continuity points of $\nabla_\chi \Gamma(\psi)$ and $\nabla_\chi \Gamma(\psi) \in \mathcal{D}_{usc}$.*

4.2 Useful representations of the derivative

Theorem 4.6 established that for H-R ORPs and $\psi, \chi \in \mathcal{D}_{\lim}$ the derivative has the form $\nabla_\chi \Gamma(\psi) = \chi + R\gamma$, where $\gamma \doteq \gamma(\psi, \chi) \in \mathcal{D}_{\lim}$ is the pointwise limit of the monotonically non-increasing sequence $\{\gamma_\varepsilon(\psi, \chi)\}$. From the expression (4.6) for $\gamma_\varepsilon^i \doteq \gamma_\varepsilon^i(\psi, \chi)$, it is clear that for every $t \in [0, \infty)$

$$\begin{aligned} \gamma^i(t) &= \lim_{\varepsilon \downarrow 0} \left[\overline{-\varepsilon^{-1}\xi^i - \chi^i + [P\gamma_\varepsilon]^i}(t) \vee 0 - \overline{-\varepsilon^{-1}\xi^i}(t) \vee 0 \right] \\ &= \lim_{\varepsilon \downarrow 0} \left[F^i\left(\varepsilon^{-1}\xi + \chi - P\gamma_\varepsilon, 0\right)(t) - F^i\left(\varepsilon^{-1}\xi, 0\right)(t) \right]. \end{aligned} \tag{4.11}$$

Thus γ^i has a representation as a one-dimensional limit of the form

$$\lim_{\varepsilon \downarrow 0} \left[F^i(\varepsilon^{-1} f + g_\varepsilon, 0) - F^i(\varepsilon^{-1} f, 0) \right], \tag{4.12}$$

where g_ε monotonically converges pointwise down to a function $g \in \mathcal{D}_{\lim}$ as $\varepsilon \downarrow 0$. If for every $\varepsilon > 0$, $g_\varepsilon \equiv g$ is independent of ε, then (4.12) reduces to

$$\lim_{\varepsilon \downarrow 0} \left[F^i(\varepsilon^{-1} f + g, 0) - F^i(\varepsilon^{-1} f, 0) \right]. \tag{4.13}$$

Limits of the form (4.13) were first analyzed in [22] for the case when $f, g \in \mathcal{C}$ and the limit has a finite number of discontinuities on any compact interval, and later generalized in [46, Theorem 9.3.1] to include the case when $f, g \in \mathcal{D}_r$, the subspace of right continuous functions in $\mathcal{D}_{\lim}$. One may be tempted to conjecture that the limit in (4.12) is equal to the limit in (4.13) with $g \in \mathcal{D}_{\lim}$ equal to the pointwise limit of $\{g_\varepsilon\}$. If that were true, then the limit in (4.12) could be identified simply by generalizing the results in [22, 46] to include the case when $g \in \mathcal{D}_{\lim}$. However, it turns out that the topology of pointwise convergence $g_\varepsilon \downarrow g$ is too weak for such a conjecture to hold in general (see Remark 4.8(3) for examples when the two limits fail to coincide). Thus a more careful analysis is required in order to determine the correct limit in (4.12). This is carried out in Section 4.2.1. Fortunately, it turns out that the conjecture is true for the special case when $f \in \mathcal{C}$ and $g_\varepsilon \in \mathcal{C}$ for all $\varepsilon > 0$, and in this case the one-dimensional limit takes a rather nice form (see Theorem 4.7).

When the reflection map is one-dimensional, $P \equiv 0$ in (4.11), and so in this case characterization of the limit on the right-hand side is sufficient to determine $\gamma(\psi, \chi)$. However in the multi-dimensional case, $P \neq 0$ and so (4.11) leads to a finite system of coupled equations that implicitly determine γ. The additional justification required to establish that this system of equations uniquely identifies γ is provided in Section 4.2.2.

4.2.1 Generalization of the one-dimensional derivative

In order to describe the limit in (4.12) we need to first introduce some definitions. Recall that the left and right regularizations g_l and g_r of any function $g \in \mathcal{D}_{\lim}$ are defined by

$$g_l(s) \doteq g(s-) \qquad \text{and} \qquad g_r(s) \doteq g(s+) \tag{4.14}$$

for $s \in [0, \infty)$. It is easy to see that $g_l(s-) = g_l(s) = g(s-)$ and $g_l(s+) = g(s+)$, and likewise $g_r(s+) = g_r(s) = g(s+)$ and $g_r(s-) = g(s-)$. Thus $g_l \in \mathcal{D}_l$, $g_r \in \mathcal{D}_r$ and

$$g \in \mathcal{D}_l \Rightarrow g_l = g, \qquad \text{and} \qquad g \in \mathcal{D}_r \Rightarrow g_r = g. \tag{4.15}$$

Also, for $f, g, g_1, g_2 \in \mathcal{D}_{\lim}$, define

$$M(f, g, g_1, g_2)(t) \doteq \begin{cases} 0 & \text{for } t \in (0, t_\ell), \\ S(f, g, g_1, g_2)(t) \vee 0 & \text{for } t \in [t_l, t_u], \\ S(f, g, g_1, g_2)(t) & \text{for } t \in (t_u, \infty), \end{cases} \tag{4.16}$$

where

$$t_\ell \doteq t_\ell(f) = \inf\{t > 0 : \bar{f}(t) = 0\}, \tag{4.17}$$

$$t_u \doteq t_u(f) = \sup\{t > 0 : \bar{f}(t) = 0\}, \tag{4.18}$$

$$S(f, g, g_1, g_2)(t) \doteq \sup_{s \in \Phi_f^L(t)} \{g_1(s)\} \vee \sup_{s \in \Phi_f(t)} \{g(s)\} \vee \sup_{s \in \tilde{\Phi}_f^R(t)} \{g_2(s)\}, \tag{4.19}$$

$$\Phi_f^L(t) \doteq \{s \in [0, t] : f(s-) = \bar{f}(t)\}, \tag{4.20}$$

$$\Phi_f(t) \doteq \{s \in [0, t] : f(s) = \bar{f}(t)\}, \tag{4.21}$$

$$\tilde{\Phi}_f^R(t) \doteq \{s \in [0, t) : f(s+) = \bar{f}(t)\} \tag{4.22}$$

Note that when $f \in \mathcal{C}$ then

$$S(f, g, g_1, g_2) = \sup_{s \in \Phi_f(t)} [g_1(s) \vee g(s)] \vee \sup_{s \in \Phi_f(t) \setminus \{t\}} [g_2(s)].$$

Let

$$S_1(f, g) \doteq S(f, g, g, g), \tag{4.23}$$

and note that

$$S_1(f, g) = \sup_{s \in \Phi_f^L(t) \cup \Phi_f(t) \cup \tilde{\Phi}^R(t)} [g(s)]. \tag{4.24}$$

Also, let

$$S_2(f, g) \doteq S(f, g, g_l, g_r), \tag{4.25}$$

and note that

$$S_2(f, g) = \sup_{s \in \Phi_f^L(t)} [g(s-)] \vee \sup_{s \in \Phi_f(t)} [g(s)] \vee \sup_{s \in \tilde{\Phi}_f^R(t)} [g(s+)]. \tag{4.26}$$

Moreover, for $j = 1, 2$, let $M_j(f, g)$ be defined as in (4.16) with $S(f, g, g_1, g_2)$ replaced by $S_j(f, g)$. In the above definitions, we use the convention that $\inf \emptyset = -\infty$. It is easy to see that for $f \in \mathcal{D}_{\lim}$ and $t \in [0, \infty)$, $\Phi_f^L(t) \cup \Phi_f(t) \cup \tilde{\Phi}_f^R(t) \neq \emptyset$ and hence $S(f, g, g_1, g_2)$, $S_1(f, g)$ and $S_2(f, g)$ are always finite. Theorem 4.7 below characterizes the generalized one-dimensional derivative. Here $\mathcal{D}_c$ and $\mathcal{D}_l$ are the subspaces of piecewise constant and left continuous functions, respectively, in $\mathcal{D}_{\lim}$.

Theorem 4.7 (Generalization of the one-dimensional derivative) *Consider a sequence* $\{g_\varepsilon\} \subseteq \mathcal{D}_{\lim}(\mathbb{R})$ *such that*

$$\sup_{\varepsilon > 0} \|g_\varepsilon\|_T < \infty \quad \textit{for every } \; T \in [0, \infty),$$

and for every $s \in [0, \infty)$

$$\varepsilon_1 \leq \varepsilon_2 \quad \Rightarrow \quad g_{\varepsilon_1}(s) \leq g_{\varepsilon_2}(s),$$

and let $g, g_l^*, g_r^* \in \mathcal{D}_{\lim}(\mathbb{R})$ be such that $g_\varepsilon \downarrow g \in \mathcal{D}_{\lim}(\mathbb{R})$, $g_{\varepsilon,l} \downarrow g_l^*$ and $g_{\varepsilon,r} \downarrow g_r^*$ pointwise as $\varepsilon \downarrow 0$, where $g_{\varepsilon,l}$ and $g_{\varepsilon,r}$ are respectively the left and right regularizations of g_ε, as defined in (4.14). For $f \in \mathcal{D}_{\lim}(\mathbb{R})$, if

$$\tilde{\gamma}_\varepsilon \doteq \overline{\varepsilon^{-1} f + g_\varepsilon} \vee 0 - \overline{\varepsilon^{-1} f} \vee 0 \tag{4.27}$$

then $\tilde{\gamma}_\varepsilon \to \tilde{\gamma} \in \mathcal{D}_{\lim}(\mathbb{R})$ pointwise as $\varepsilon \downarrow 0$, where

$$\tilde{\gamma} \doteq M(f, g, g_l^*, g_r^*), \tag{4.28}$$

and M is given by (4.16). Moreover, when $\{g_\varepsilon, \varepsilon > 0\} \subset \mathcal{C}$, then the generalized derivative takes the simpler form

$$\tilde{\gamma} = M_1(f, g), \tag{4.29}$$

and if in addition $f \in \mathcal{C}$, then

$$\tilde{\gamma} = M_1(f,g) = \begin{cases} 0 & \text{if } t \in (0, t_l), \\ \sup_{s \in \Phi_f(t)} [g(s)] \vee 0 & \text{if } t \in [t_l, t_u), \\ \sup_{s \in \Phi_f(t)} [g(s)] & \text{if } t \in [t_u, \infty), \end{cases} \tag{4.30}$$

and

$$M_1(f, g) = M_2(f, g). \tag{4.31}$$

Lastly, if $f \in \mathcal{D}_c$ and $g \in \mathcal{D}_{\lim}$ then

$$\tilde{\gamma} = M_2(f, g). \tag{4.32}$$

The proof of Theorem 4.7 is given in Section 6.1. Here we make some observations on the theorem.

Remark 4.8 *(The generalized one-dimensional derivative)*

1. Note that the limit in (4.13) is given by $S_2(f,g)$ for $f, g \in \mathcal{D}_{\lim}$, where $S_2(f,g)$ is defined by (4.26). Indeed, when $g_\varepsilon = g$ is independent of ε, then clearly $g_l^* = g_l$ and $g_r^* = g_r$ (see Lemma 6.1(1)), and so in this case by (4.25)

   $$S(f, g, g_l^*, g_r^*) = S(f, g, g_l, g_r) = S_2(f, g).$$

   If in addition $f, g \in \mathcal{C}$, then $\Phi_f^L(t) \cup \Phi_f(t) \cup \Phi_f^R(t) = \Phi_f(t)$ and $g(s-) = g(s) = g(s+)$, so that

   $$S_2(f,g)(t) \doteq \sup_{s \in \Phi_f(t)} g(s). \tag{4.33}$$

   Thus Theorem 4.7 contains as a special case the results in [22, Lemma 5.2] and [46, Theorem 9.3.1] (with Φ_f^R replaced by $\tilde{\Phi}_f^R$ in the latter result, as mentioned in greater detail in Remark 4.8.2 below).

2. Note that the notation $\tilde{\Phi}_f^R$ rather than Φ_f^R is used in the definitions of S, S_1 and S_2 in order to emphasize that $t \notin \tilde{\Phi}_f^R(t)$, in contrast with the sets $\Phi_f^L(t)$ and $\Phi_f(t)$, which could contain t. In the definition for $S_2(f,g)$ in [46, Theorem 9.3.1], however, the set $\tilde{\Phi}_f^R$ is replaced by the set

   $$\Phi_f^R(t) \doteq \{s \in [0,t] : f(s+) = \overline{f}(t)\}, \tag{4.34}$$

which could contain t. The following example illustrates that if modified in this manner, $S_2(f,g)$ no longer identifies the correct limit in (4.13) when $g \in \mathcal{D}_{\lim}$, even if $f \in \mathcal{C}$. Thus the correct definition of S_2 is with $\tilde{\Phi}_f^R$ rather than with Φ_f^R.

Example 1. Let $f(s) \doteq s1_{[0,1)}(s) + 1_{[1,2)}(s)$ for $s \in [0,2]$ and for every $\varepsilon > 0$ let $g_\varepsilon(s) = g(s) \doteq 1_{(1,2]}$ for $s \in [0,2]$. Then f is continuous and g is left continuous. Moreover, from the definition of f it follows that $\Phi_f^L(1) = \Phi_f(1) = \{1\}$ and $\tilde{\Phi}_f^R(1) = \emptyset$, while $\Phi_f^R(1) = \{1\}$. By (4.26) we have $S_2(f,g)(1) = g(1-) \vee g(1) = 0$, while for the modified case (i.e. with $\tilde{\Phi}_f^R$ replaced by Φ_f^R in the definition of S_2) we see that $S_2(f,g)(1) = g(1-) \vee g(1) \vee g(1+) = 1$. However, by direct verification it is easy to see in this simple example that

$$\lim_{\varepsilon \downarrow 0} \left[\overline{\varepsilon^{-1} f + g_\varepsilon}(1) - \overline{\varepsilon^{-1} f}(1) \right] = \lim_{\varepsilon \downarrow 0} \left[\overline{\varepsilon^{-1} f + g}(1) - \overline{\varepsilon^{-1} f}(1) \right] = g(1) = 0.$$

So clearly for this example $\tilde{\Phi}_f^R(t)$ is the right set to be used in the definition of S_2 in order to obtain the correct limit in (4.13). An analysis of the proof of Theorem 4.7 reveals that this is true in general.

3. By Theorem 4.7 it follows that when both f and $\{g_\varepsilon, \varepsilon > 0\}$ are continuous, $\tilde{\gamma} = M_1(f,g) = M_2(f,g)$. Since the limit in (4.12) is given by $\tilde{\gamma}$ and by Remark 4.8.1 above the limit in (4.13) is given by $M_2(f,g)$, we see that the two limits in (4.12) and (4.13) coincide in this special case. However, the following two examples demonstrate these two limits need not be equal for general $f, g, \{g_\varepsilon\} \in \mathcal{D}_{\lim}$. Example 2 demonstrates the necessity of having g_ε continuous, while the Example 3 shows why f must be continuous.

Example 2. Let $f(s) \doteq s$ and $g(s) \doteq 1$ for $s \in [0,2]$. Also, for $\varepsilon > 0$ let

$$g_\varepsilon \doteq \begin{cases} 1 & \text{for } t \in [0, 1-\varepsilon) \\ 2 & \text{for } t \in [1-\varepsilon, 1) \\ 1 & \text{for } t \in [1,2] \end{cases}$$

Then clearly f and g are continuous and each g_ε is right continuous. Moreover, $\Phi_f^L(1) = \Phi_f(1) = 1$, $\tilde{\Phi}_f^R(1) = \emptyset$ and the fact that $g_\varepsilon(1-) = 2$ for every $\varepsilon > 0$ implies $g_l^*(1) = 2$. By Theorem 4.7 the limit in (4.12) is equal to $S(f, g, g_l^*, g_r^*) = g_l^*(1) \vee g(1) = 2$, while by Remark 4.8.1 above the limit in (4.13) is equal to $S_2(f,g) = g(1-) \vee g(1) = 1$, which is clearly not equal to 2.

Example 3. On the interval $[0,2]$ define the functions $f(s) \doteq s1_{[0,1)}$, $g \doteq 1_{[1,2]}$ and

$$g_\varepsilon(s) \doteq \begin{cases} 0 & \text{for } s \in [0, 1-\varepsilon) \\ \dfrac{s-(1-\varepsilon)}{\varepsilon} & \text{for } s \in [1-\varepsilon, 1) \\ 1 & \text{for } s \in [1,2] \end{cases}$$

Then clearly $\{g_\varepsilon, \varepsilon > 0\}$ is a sequence of continuous functions that converges pointwise monotonically down to g, which is right continuous. Moreover f is also right continuous, $\Phi_f^L(1) = \{1\}$ and $\Phi_f(1) = \tilde{\Phi}^R(1) = \emptyset$. By Remark 4.8.1 above the limit in (4.13) is given by $S_2(f,g)(1) = g(1-) = 0$. On the other hand since g_ε are continuous, by (4.29) the limit in (4.12) is equal to $S_1(f,g) = g(1) = 1$, which is clearly not equal to 0.

4.2.2 Autonomous Characterizations of γ

In this section we show that for H-R ORPs Theorems 4.6 and 4.7 along with the relation (4.1) uniquely characterize the limit γ when $\psi, \chi \in \mathcal{C}$ or when $\psi \in \mathcal{D}_c$ and $\chi \in \mathcal{D}_{\lim}$. Fix $\psi, \chi \in \mathcal{D}_{\lim}$, and as usual let γ_ε and ξ be defined via (4.1) and (4.4) respectively. Also, for $i = 1, \ldots, K$ let χ_l^i and χ_r^i be the left and right regularizations of χ^i as defined in (4.14), let γ_l^* and γ_r^* be the limits of the left and right regularized sequences $\{\gamma_{\varepsilon,l}\}$ and $\{\gamma_{\varepsilon,r}\}$ respectively, and let γ be the unique pointwise limit of $\{\gamma_\varepsilon\}$ which exists by Corollary 4.3. Moreover, let the functions M and M_j, $j = 1, 2$ be defined as in Section 4.2.1. Given the relation (4.11) and Theorem 4.7 it follows that for $i = 1, \ldots, K$

$$\gamma^i(\psi, \chi)(t) = M\left(-\xi^i, -\chi^i + [P\gamma]^i, -\chi_l^i + [P\gamma_l^*]^i, -\chi_r^i + [P\gamma_r^*]^i\right)(t). \tag{4.35}$$

Since in general γ_l^* and γ_r^* are defined in terms of the sequence $\{\gamma_\varepsilon(\psi, \chi\}$, and are not uniquely determined by $\gamma(\psi, \chi)$ this does not lead to an autonomous characterization of $\gamma(\psi, \chi)$. However, Theorem 4.7 identifies cases in which γ_l^* and γ_r^* are uniquely determined by $\gamma(\psi, \chi)$. Specifically $\gamma \doteq \gamma(\psi, \chi)$ satisfies

$$\gamma^i = M_j\left(-\xi^i, -\chi^i + [P\gamma]^i\right) \qquad \text{for} \quad i = 1, \ldots, K \tag{4.36}$$

with $j = 1$ when $\psi, \chi \in \mathcal{C}$ and $j = 2$ when $\psi \in \mathcal{D}_c$ and $\chi \in \mathcal{D}_{\lim}$.

Lemma 4.9 *Given a matrix P whose spectral radius is less than one, let $\psi, \chi \in \mathcal{D}_{\lim}$. Then for $j = 1, 2$ the system of equations (4.36) has a unique solution $\gamma_{(j)} \doteq \gamma_{(j)}(\psi, \chi) \in \mathcal{D}_{\lim}$. Moreover, for $j = 1, 2$ given any $\gamma_{0,j} \in \mathcal{D}_{\lim}$, if the sequence $\{\gamma_{n,j}\}$ is defined recursively by*

$$\gamma_{n+1,j} \doteq M_j\left(-\xi^i, -\chi^i + [P\gamma_{n,j}]^i\right)$$

then for every $T < \infty$ $||\gamma_{(j)} - \gamma_{n,j}||_T \to 0$ as $n \uparrow \infty$.

Proof. Fix $\psi, \chi \in \mathcal{D}_{\lim}$, and recall from Lemma 4.1 and Corollary 4.3 that $\gamma(\psi, \chi)$ is uniformly bounded in $\mathcal{D}_{\lim}$ (with respect to the sup norm). From the definition of M it is easy to see that M maps bounded sets to bounded sets. We show below that M_1 is a contraction mapping (with respect to the sup norm topology) on $\mathcal{D}_{\lim}$. Since $\mathcal{D}_{\lim}$ endowed with the sup norm metric is a complete metric space, the existence of a unique fixed point for M_1 then follows from standard theorems [43, Theorems 5.2.1 and 5.2.3]. To establish the contraction property we first consider the case when the maximum row sum of the matrix P is equal to $\delta < 1$. The general case can then be handled in the usual way using diagonal similarity transforms (see, for example, [44] or the proof of Lemma 4.4). Let $\gamma_1, \gamma_2 \in \mathcal{D}_{\lim}$ and fix $T \in [0, \infty]$. Then the definition of M_1 along with Lemma A.1 yields

$$\begin{aligned}
\max_{i=1,\ldots,K} \sup_{t\in[0,T]} |M_1^i\left(-\xi^i, -\chi^i + [P\gamma_1]^i\right) - M_1^i\left(-\xi^i, -\chi^i + [P\gamma_2]^i\right)| \\
\le \sup_{s\in[0,T]} \max_{i=1,\ldots,K} |[P\gamma_1]^i(s) - [P\gamma_2]^i(s)| \\
\le \sup_{s\in[0,T]} \max_{i=1,j\ldots,K} |\textstyle\sum_{k=1}^K P_{ik}||\gamma_1^j(s) - \gamma_2^j(s)| \\
\le \delta \max_{j=1,\ldots,K} \sup_{s\in[0,T]} |\gamma_1^j(s) - \gamma_2^j(s)|,
\end{aligned}$$

which proves the contraction property since $\delta < 1$. The proof for M_2 follows analogously and is thus omitted. ■

When combined with Corollary 4.3 and Theorem 4.7, Lemma 4.9 yields the following autonomous characterization of the derivative for H-R SPs.

Theorem 4.10 (Autonomous Characterizations of the Derivative) *Given a H-R ORP with constraint matrix R, RM Γ, and $\psi, \chi \in \mathcal{D}_{\lim}$, let ξ be defined as in (4.4) and for $j = 1, 2$ let $\gamma_j(\psi, \chi) \in \mathcal{D}_{\lim}$ be the unique solution to the system of equations (4.36). Then the directional derivative $\nabla_\chi \Gamma(\psi)$ lies in $\mathcal{D}_{\lim}$ and satisfies*

$$\nabla_\chi \Gamma(\psi) = \begin{cases} \chi + R\gamma_{(1)}(\psi, \chi) & \text{if } \psi, \chi \in \mathcal{C}, \\ \chi + R\gamma_{(2)}(\psi, \chi) & \text{if } \psi \in \mathcal{D}_c \text{ and } \chi \in \mathcal{D}_{\lim}. \end{cases} \tag{4.37}$$

Moreover, for $\psi, \chi \in \mathcal{C}$, $\nabla^\varepsilon_\chi \Gamma(\psi)$ converges to $\nabla_\chi \Gamma(\psi)$ uniformly on compact subsets of the continuity points of $\nabla_\chi \Gamma(\psi)$.

5 Properties of $\nabla_\chi \Gamma(\psi)$ when $\psi, \chi \in \mathcal{C}$

In this section we derive properties of the directional derivative $\nabla_\chi \Gamma(\psi)$ for H-R ORPs when ψ and χ are continuous. The main result of this section is the proof of Theorem 1.2, which classifies the discontinuities of $\nabla_\chi \Gamma(\psi) \in \mathcal{D}_{\lim}$ when $\psi, \chi \in \mathcal{C}$. It turns out that there is a connection between the points at which these discontinuities occur and points at which there is a change in regime for the solution (ϕ, θ) to the ORP with input ψ. When the ORP and the input trajectory ψ arise from a non-stationary network, this corresponds to a change in regime of the fluid limit of the network. These different regimes are introduced in Section 5.1 and the proof of Theorem 1.2 is presented in Section 5.2. The proof relies on the classification of the discontinuities of $\gamma_{(1)}(\psi, \chi)$ presented in Theorem 5.2. This classification is also used in [25] to establish convergence of $\nabla^\varepsilon_\chi \Gamma(\psi)$ to $\nabla_\chi \Gamma(\psi)$ in topologies stronger than that of pointwise convergence, and to prove functional central limits for non-stationary networks.

5.1 Regimes of the Fluid Limit (ϕ, θ)

As mentioned in the introduction, the study of directional derivatives is largely motivated by its use in describing the behaviour of non-stationary queueing networks. In this setting, ψ has the interpretation as the functional strong law of large numbers (FSLLN) limit of the so-called netput process of a stochastic network, and χ represents the sample path of (a possibly time-changed version of) the functional central limit of the netput process. In this context $\phi = \Gamma(\psi)$ represents the fluid approximation to the queues in the network. The terminology of this section reflects this interpretation.

A change in regime represents a transition at any one node from one of the states of under-loading, criticality or overloading to another state. For single-class stationary networks these three regimes correspond to the regions in which the traffic load parameter ρ (which is obtained from the solution to the traffic equations of the queueing network [12]) is less than one, greater than one, and equal to one. Since for stationary networks the traffic load parameter ρ is constant, the regimes are independent of time. In particular, heavy-traffic diffusion approximations of stationary networks correspond to the study of the critical regime, where $\rho = 1$. The situation is considerably more complicated for non-stationary queueing networks. As shown in [22], for the non-stationary $M_t/M_t/1$ queue with average arrival and service rates of $\lambda(t)$ and $\mu(t)$ respectively at time t, the instantaneous effective load parameter

$$\rho(t) \doteq \frac{\int_0^t \lambda(s)ds}{\int_0^t \mu(s)ds},$$

can be used to determine in which regime the system lies at time t. Although the corresponding regimes can vary with time, in the single-queue case the regimes of underloading, overloading and criticality (at time t) still correspond to when $\rho(t)$ is less than one, greater than one and equal to one. For the multi-dimensional network with time-dependent rates it is more convenient to define the regimes in terms of the "fluid limit" of the queue length process. In particular, given an input trajectory ψ and an ORP satisfying Assumption 3.1, let (ϕ, θ) be the solution to the ORP for the input ψ. Then for $t \in [0, \infty)$ we define the following sets:

$$\begin{array}{rcl} Over(t) & \doteq & \{i \in \mathcal{I} : \phi^i(t) > 0\} \\ Crit(t) & \doteq & \{i \in \mathcal{I} : \phi^i(t) = 0, \dot{\theta}^i(t-) = 0\} \\ Under(t) & \doteq & \{i \in \mathcal{I} : \phi^i(t) = 0, \dot{\theta}^i(t-) > 0\}. \\ Over(t-) & \doteq & \cup_{\varepsilon>0} \cap_{s \in (t-\varepsilon,t)} Over(s) \\ Under(t+) & \doteq & \cup_{\varepsilon>0} \cap_{s \in (t,t+\varepsilon)} Under(s) \end{array} \tag{5.1}$$

Note that i lies in $Over(t-)$ (respectively $Under(t+)$) if and only if there exists $\varepsilon > 0$ such that $i \in Over(s)$ for all $s \in (t-\varepsilon, t)$ (respectively $i \in Under(s)$ for all $s \in (t, t+\varepsilon)$). We now define various regimes of the fluid limit.

Definition 5.1 *(Regimes of the Fluid Limit) Given an ORP satisfying Assumption 3.1, let (ϕ, θ) be the solution to the ORP for a given input trajectory ψ. Then $i \in \mathcal{I}$ is said to be overloaded (respectively critical, underloaded) at time t if and only if $i \in Over(t)$ (respectively $i \in Crit(t)$, $i \in Under(t)$). Moreover, at time t i is said to be at the end of overloading if and only if $i \in Crit(t) \cap Over(t-)$ and to be at the start of underloading if and only if $i \in Crit(t) \cap Under(t+)$.*

The regimes above are defined in terms of physical characteristics of the fluid limit. On the other hand, analysis of the discontinuities of the derivative $\nabla_\chi \Gamma(\psi)$ lead naturally to conditions involving the sets $\Phi_{-\xi^i}(t)$ defined in Section 4.2.1. Lemma 5.1 links these two sets of conditions by providing equivalent definitions for the regimes introduced in Definition 5.1 in terms of the sets $\Phi_{-\xi^i}(t)$. The proof of the lemma is given in Section 6.2.1.

Lemma 5.1 *Given an ORP satisfying Assumption 3.1 and $\psi \in \mathcal{C}$ let ϕ, θ and ξ be as defined in Section 4.1.1 and for $i \in \mathcal{I}$, let $\Phi_{-\xi^i}(t)$ be given by (4.34). Then the following equivalences are satisfied.*

1. *$t_l^i = \inf\{t > 0 : \phi^i(t) = 0\}$ and $t_u^i = \inf\{t > 0 : \dot{\theta}^i(t+) > 0\}$;*
2. *$\Phi_{-\xi^i}(t) = \Phi^i(t)$ for $t \in [t_u^i, \infty)$, where*
$$\Phi^i(t) \doteq \left\{ s \in [0, t] : \phi^i(s) = 0 \quad \text{and } \theta^i(s) = \theta^i(t) \right\}; \tag{5.2}$$
3. **(Overloaded)** *i is overloaded at t if and only if either $t \in [0, t_l^i)$ or $t \notin \Phi_{-\xi^i}(t)$;*
4. **(Critical)** *i is critical at t if and only if either $t \in \Phi_{-\xi^i}(t) \neq \{t\}$ or $t \in [t_l^i, t_u^i]$;*
5. **(Underloaded)** *i is underloaded at t if and only if $t \in (t_u^i, \infty)$ and $\{t\} = \Phi_{-\xi^i}(t)$;*
6. **(End of Overloading)** *i is at the end of overloading at t if and only if either $t = t_l^i$ or $t \in (t_u^i, \infty)$, $t \in \Phi_{-\xi^i}(t) \neq \{t\}$ and t is isolated in $\Phi_{-\xi^i}(t)$ (i.e. there exists $\varepsilon > 0$ such that $\Phi_{-\xi^i}(t) \cap (t-\varepsilon, t) = \emptyset$);*

7. **(Start of Underloading)** *i is at the start of underloading at t if and only if either $t = t_u^i$ or $t \in (t_u^i, \infty)$, $t \in \Phi_{-\xi^i}(t) \neq \{t\}$ and $\Phi_{-\xi^i}(s) \subset (s,t]$ for all $s > t$.*

5.2 Classification of the discontinuities of $\nabla_\chi \Gamma(\psi)$ when $\psi, \chi \in \mathcal{C}$

In this section the discontinuities of $\gamma_{(1)}$ are classified in terms of various conditions on the sets Φ^i (see Theorem 5.2). In order to allow a physical interpretation of the conditions in a stochastic network setting, the statement of the theorem also provides equivalent descriptions of these conditions (as established in Lemma 5.1) in terms of the fluid limit regimes defined in the last section. In Lemma 5.3 necessary conditions for the existence of discontinuities in $\gamma_{(1)}^i$ are shown to have an interpretation in terms of so-called critical and underloaded chains, which are introduced in Definition 5.2. This interpretation is used to prove Theorem 1.2, which states necessary conditions for the derivative $\nabla_\chi \Gamma(\psi)$ to have various types of discontinuities.

Theorem 5.2 (Discontinuities of $\gamma_{(1)}$) *Given a H-R ORP with constraint matrix $R \in \mathbb{R}^{K \times K}$, and $\psi, \chi \in \mathcal{C}$, let $\gamma \doteq \gamma_{(1)}(\psi, \chi)$ be the unique solution to the equation (4.36) with $j = 1$. Moreover, let $\nabla \Gamma \doteq \nabla_\chi \Gamma(\psi)$ and let $t_u^i \doteq t_u(-\xi^i)$ be defined as in (4.18) with ξ^i defined by (4.4). Then $\gamma^i(t) = 0$ for $t \in [0, t_l^i)$. Moreover, $t \in [t_l^i, \infty)$ is a point of left discontinuity for γ^i if and only if one of the following properties hold.*

L0. **(End of Overloading)** *$t = t_l^i$ and*

$$-\chi(t_l^i) + [P\gamma]^i(t_l^i) > 0, \tag{5.3}$$

in which case $\nabla \Gamma^i(t_l^i) = 0$. Moreover, $\nabla \Gamma^i(t_l^i -) < \nabla \Gamma^i(t_l^i)$ if $[P\gamma]^i$ is continuous at t_l^i.

L1. **(End of Overloading)** *$t \in \Phi^i(t) \neq \{t\}$, t is isolated in $\Phi^i(t)$ (i.e. there exists $\delta > 0$ such that $(t - \delta, t) \cap \Phi^i(t) = \emptyset$),*

$$\sup_{s \in \Phi^i(t) \setminus \{t\}} [-\chi^i(s) + [P\gamma]^i(s)] < -\chi^i(t) + [P\gamma]^i(t). \tag{5.4}$$

and $[P\gamma]^i$ is left continuous at t. In this case $\nabla \Gamma^i(t-) < \nabla \Gamma^i(t) = 0$.

L2. **(End of Overloading)** *$t \in \Phi^i(t) \neq \{t\}$, t is isolated in $\Phi^i(t)$, (5.4) holds and $[P\gamma]^i$ is left discontinuous. In this case $\nabla \Gamma^i(t) = 0$.*

L3. **(Critical)** *$t \in \Phi^i(t) \neq \{t\}$, t is not isolated in $\Phi^i(t)$ (i.e. there exists $\delta > 0$ such that $(t - \delta, t] \subset \Phi^i(t)$), (5.4) holds and $[P\gamma]^i$ is left discontinuous at t. In this case $\nabla \Gamma^i(t-) \geq \nabla \Gamma^i(t) = 0$.*

L4. **(Underloaded)** *$\{t\} = \Phi^i(t)$ and $[P\gamma]^i$ is left discontinuous at t. In this case $\nabla \Gamma^i(t-) = \nabla \Gamma^i(t) = 0$.*

Similarly, γ^i is right continuous on $[0, t_u^i)$ and $t \in [t_u^i, \infty)$ is a point of right discontinuity for γ^i if and only if

R0. **(Start of Underloading)** *$t = t_u^i$ and either $[P\gamma]^i$ is right continuous at t_u^i and*

$$\sup_{s \in [t_l^i, t_u^i]} \left[-\chi^i(s) + [P\gamma]^i(s) \right] < 0,$$

or $[P\gamma]^i$ is right discontinuous at t_u^i. In both cases, $\nabla \Gamma^i(t) > \nabla \Gamma^i(t+) = 0$.

R1. **(Start of Underloading)** *$t \in \Phi^i(t) \neq \{t\}$, $\Phi_{-\xi^i}(s) \subset (t, s]$ for every $s > t$,*

$$\gamma^i(t) > -\chi^i(t) + [P\gamma]^i(t), \tag{5.5}$$

and $[P\gamma]^i$ is right continuous at t. In this case γ^i is left continuous at t and $\nabla\Gamma^i(t) > \nabla\Gamma^i(t+) = 0$.

R2. **(Start of Underloading)** *$t \in \Phi^i(t) \neq \{t\}$, $\Phi_{-\xi^i}(s) \subset (t, s]$ for every $s > t$, (5.5) holds and $[P\gamma]^i$ is right discontinuous at t. In this case γ^i is left continuous at t and $\nabla\Gamma^i(t) > \nabla\Gamma^i(t+) = 0$.*

R3. **(Start of Underloading)** *$t \in \Phi^i(t) \neq \{t\}$, $\Phi^i(s) \subset (t, s]$ for every $s > t$,*

$$\gamma^i(t) = -\chi^i(t) + [P\gamma]^i(t), \tag{5.6}$$

and $[P\gamma]^i$ is right discontinuous at t. In particular, (5.6) holds if γ^i is left discontinuous at t. In this case $\nabla\Gamma^i(t) = \nabla\Gamma^i(t+) = 0$.

R4. **(Underloaded)** *$\{t\} = \Phi^i(t)$ and $[P\gamma]^i$ is right discontinuous at t. In this case $\nabla\Gamma^i(t) = \nabla\Gamma^i(t+) = 0$.*

Finally, $t \in [0, \infty)$ is a point of both left or right discontinuity for γ^i if and only if $t \in (t_u^i, \infty)$ and one of the following holds.

S1. **(End of Overloading and Start of Underloading)** *Either L1 or L2 holds along with R3. In this case $\nabla\Gamma^i(t) = \nabla\Gamma^i(t+) = 0$.*

S2. **(Start of Underloading)** *L3 holds along with R3. In this case $\nabla\Gamma^i(t-) \geq \nabla\Gamma^i(t) = \nabla\Gamma^i(t+) = 0$.*

S3. **(Underloaded)** *L4 holds along with R4. In this case $\nabla\Gamma^i(t-) = \nabla\Gamma^i(t) = \nabla\Gamma^i(t+) = 0$.*

The proof of the theorem is given in Section 6.2.3. We now use the classification of the discontinuities of $\gamma_{(1)}(\psi, \chi)$ in Theorem 5.2 to study continuity properties of the derivative $\nabla_\chi\Gamma(\psi) = \chi + R\gamma_{(1)}(\psi, \chi)$. Before we prove the main theorem, Theorem 1.2, it will be useful to introduce a couple definitions.

Definition 5.2 (Critical and Sub-critical Chains) *Given a constraint matrix $R \in \mathbb{R}^{K\times K}$ satisfying Assumption 3.1, with $P \doteq I - R$ and associated RM Γ and $\psi \in \mathcal{C}$, let $\phi \doteq \Gamma(\psi)$. Then a sequence $i \doteq j_0, j_1, j_2, \ldots, j_m$ with $j_i \in \{1, \ldots, K\}$ for $i = 0, \ldots, m$, is said to be an* **empty chain preceding** *i at t if $P_{j_k j_{k-1}} > 0$ and $\phi^{j_k}(t) = 0$ for $k = 1, \ldots, m$. An empty chain preceding i is said to be* **critical** *at t if either*

i. **(End of Overloading)** *m is at the end of overloading or*

ii. **(Cycle)** *there exists $k < m$ such that $j_k = j_m$.*

Moreover, an empty chain preceding i is said to be **sub-critical** *at t if k is at the start of underloading for every $k = 1, \ldots, m$, and either*

i. **(Start of Underloading)** *m is at the start of underloading*
or

ii. **(Cycle)** *there exists $k < m$ such that $j_k = j_m$.*

The following lemma shows that necessary conditions for $\gamma_{(1)}(\psi, \chi)$ to have a discontinuity can be phrased in terms of critical and sub-critical chains.

Lemma 5.3 (Physical conditions for discontinuities in γ) *Given an ORP with H-R constraint matrix R, and $\psi, \chi \in \mathcal{C}$, and let $\gamma \doteq \gamma_{(1)}(\psi, \chi)$ be the unique solution to the equation (1.19). If γ^i is left discontinuous at $t \in [0, \infty)$ then i is either at the end of overloading or there exists a critical chain preceding i at time t. Likewise, if γ^i is right discontinuous at $t \in [0, \infty)$ then either i is at the start of underloading, or there exists a sub-critical chain preceding i at time t.*

Proof. This can be easily deduced from Theorem 5.2 and Definition 5.2. ■

We are now in a position to prove Theorem 1.2 stated in Section 1.3, which classifies the discontinuities of the derivative $\nabla_\chi \Gamma(\psi)$ when $\psi, \chi \in \mathcal{C}$. The conditions L1, R2 etc. mentioned in the proof below refer to the notation in Theorem 5.2. For notational simplicity we omit the explicit dependence of $\nabla_\chi \Gamma(\psi)$ on χ and ψ.

Proof of Theorem 1.2. A basic observation is that since $\chi \in \mathcal{C}$, due to (6.8) it follows that if $\nabla \Gamma^i$ has a (left) discontinuity, then either γ^i or $[P\gamma]^i$ must have a left (right) discontinuity. Also note that the upper semicontinuity of $[P\gamma]^i$ (which follows from Corollary 4.3 and the fact that $P \geq 0$) ensures that if γ^i is left continuous, then

$$\nabla \Gamma^i(t) - \nabla \Gamma^i(t-) = [P\gamma]^i(t-) - [P\gamma]^i(t) \leq 0, \tag{5.7}$$

while if γ^i is right continuous, then

$$\nabla \Gamma^i(t) - \nabla \Gamma^i(t+) = [P\gamma]^i(t+) - [P\gamma]^i(t) \leq 0. \tag{5.8}$$

For (1.20) to hold, it follows from from (5.7) that γ^i must be left discontinuous and from Theorem 5.2 that either L0, L1 or L2 must hold, from which L(a) follows. Likewise, if (1.21) is satisfied then it follows from (5.7) and Theorem 5.2 that i is not underloaded and either γ^i is left continuous or γ^i is left-discontinuous and L3 holds. In either case $[P\gamma]^i$ must be left-discontinuous, and so there exists γ^j with $p_{ji} > 0$ such that γ^j is left-discontinuous at t. By Lemma 5.3 and Theorem 5.2 this implies that $\phi^j(t) = 0$. and either j is at the end of overloading, or has a critical chain preceding it. Since $\phi^j(t) = 0$ and $p_{ji} > 0$ this implies that there is a critical chain preceding i, which completes the proof of L(b).

The proof for right discontinuities follows in a similar fashion. By (5.8), if (1.22) holds then γ^i must be right continuous, and so by Theorem 5.2 either R0, R1 or R2 must hold, thus establishing R(a). Now observe that Theorem 5.2 implies that (1.23) cannot occur either i is underloaded or when γ^i is right discontinuous, and so $[P\gamma]^i$ must be right discontinuous. In analogy with the argument given for left discontinuities, by Lemma 5.3 this guarantees the existence of a sub-critical chain preceding i, and completes the proof of R(b). The conditions [LR] can be easily deduced from the conditions for [L] and [R] and the observation that when both [La] and [Ra] are satisfied, then $\nabla \Gamma^i$ is right continuous.

To establish the two remaining statements of the theorem, first note that the continuity of $\nabla \Gamma^i$ when i is underloaded follows from L4 and R4 of Theorem 5.2. Secondly, note that if

i is overloaded, then Theorem 5.2 shows that γ^i must be continuous. If in addition $[P\gamma]^i$ is continuous, then $\nabla\Gamma^i$ is continuous and (1.25) follows with equality. On the other hand by the arguments given above it follows that if $\nabla\Gamma^i$ is discontinuous and i is overloaded, then either (Lb), (Rb) or (LRc) must hold, from which (1.25) follows.

Finally the Lipschitz and homogeneity properties of the derivative are a straightforward consequence of the explicit form of the derivative given in (1.19) and the scaling property of the ORM: $\Gamma(\beta\psi) = \beta\Gamma(\psi)$ for $\beta > 0$. This concludes the proof of the theorem. ■

6 Proofs of the Main Theorems

6.1 Proof of Theorem 4.7

In this section we prove the representation for the generalized one-dimensional derivative presented in Section 4.2.1. We first establish a lemma that will be needed for the proof of the theorem. The lemma identifies conditions under which the expression $M(f, g, g_l^*, g_r^*)$ defined in (4.16) can be expressed purely as a function of f and g.

Lemma 6.1 *Let the sequence $\{g_\varepsilon\}$ satisfy the uniform boundedness and monotonicity properties stated in Theorem 4.7, and let g_l^* and g_r^* be defined as in Theorem 4.7. Then the following properties hold.*

1. *$g_l^* \geq g_l$ and $g_r^* \geq g_r$.*
2. *If $g_\varepsilon = g$ is independent of ε, then $g_l^* = g_l$ and $g_r^* = g_r$.*
3. *If $\{g_\varepsilon\} \subset \mathcal{C}$ then $g_l^* = g_r^* = g$.*
4. *If g_ε converges to g in the uniform topology, i.e. for every $T < \infty$*

$$\lim_{\varepsilon \downarrow 0} ||g_\varepsilon - g||_T = 0,$$

   then $g_l^* = g_l$ and $g_r^* = g_r$.

5. *For any s that is a point of left continuity for g,*

$$g_l^*(s) = g_l(s),$$

   and likewise for any s that is a point of right continuity for g

$$g_r^*(s) = g_r(s).$$

Proof. Fix $t \in [0, \infty)$. For every $\varepsilon > 0$ choose $t_\varepsilon \in (t - \varepsilon, t)$ such that $|g_\varepsilon(t_\varepsilon) - g_\varepsilon(t-)| < \varepsilon$. Then $t_\varepsilon \uparrow t$ and the monotonicity of the sequence g_ε dictates that $g(t_\varepsilon) < g_\varepsilon(t_\varepsilon) < g_\varepsilon(t-) + \varepsilon$. Taking limits as $\varepsilon \downarrow 0$ leads to the conclusion that $g_l(t) = g(t-) \leq g_l^*(t)$. An analogous argument yields the inequality $g_r \leq g_r^*$, thus establishing the first property. The next two properties follow directly from the definitions and the assumed monotonicity of the sequence $\{g_\varepsilon\}$.

To prove the fourth property, for $\varepsilon \in (0,1]$ choose t_ε such that $t_\varepsilon \downarrow s$ and $|g_\varepsilon(t_\varepsilon) - g_\varepsilon(s+)| \leq \varepsilon$. Then for any $s \in [0,\infty)$ uniform convergence of g_ε to g on the interval $[0,2s]$ implies that given any $\delta > 0$ there exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ and $n \in I\!N$,

$$g(t_\varepsilon) - \delta \leq g_\varepsilon(t_\varepsilon) \leq g(t_\varepsilon) + \delta,$$

which in turn implies that

$$g(t_\varepsilon) - \delta - \varepsilon \leq g_\varepsilon(s+) \leq g(t_\varepsilon) + \delta + \varepsilon.$$

Taking limits as $\varepsilon \downarrow 0$ yields the inequality

$$g_r(s) - \delta = g(s+) - \delta \leq g_r^*(s) \leq g(s+) + \delta = g_r(s) + \delta,$$

and sending $\delta \downarrow 0$ leads to the conclusion that $g_r(s) = g_r^*(s)$. An analogous argument can be used to prove that $g_l = g_l^*$.

For the fifth property first note that if s is a point of left continuity for g then since $g \in \mathcal{D}_{\lim}$ has only a countable number of discontinuities, there exists a sequence $\{t_n\}$ of continuity points of g such that $t_n \uparrow s$. Then clearly $\{t_n\}$ is a compact subset of the continuity points of g and hence by Corollary 4.3 we know that g_ε converges uniformly to g on the subset $\{t_n\}$, and hence on the subset $\{t_n\} \cup \{s\}$. Using an argument similar to that used to prove the fourth property above, it then follows that $g_l^*(s) = g_l(s) = g(s-)$. The result when s is a point of right continuity follows likewise. ■

Proof of Theorem 4.7. First note that the sequence $\{g_\varepsilon\}$ has a unique pointwise limit g since for each $s \in [0,\infty)$ $\{g_\varepsilon(s)\}$ is uniformly bounded and monotonically non-increasing. By the same token, the fact that the left and right regularized sequences $\{g_{\varepsilon,l}\}$ and $\{g_{\varepsilon,r}\}$ inherit the uniform boundedness and monotonicity properties of $\{g_\varepsilon\}$ establishes the existence of g_l^* and g_r^*.

Let $t_\ell \doteq t_\ell(f)$ and $t_u \doteq t_u(f)$ be defined as in (4.17) and (4.18) respectively. Fix $t > t_u$, so that $\overline{f}(t) > 0$. Since $\{g_\varepsilon\}$ is uniformly bounded, relation (1.4) of Lemma A.3 guarantees the existence of $\varepsilon_0 > 0$ such that $\overline{\varepsilon^{-1} f + g_\varepsilon}(t) > 0$ for all $\varepsilon \in (0,\varepsilon_0)$. Hence for all $\varepsilon \in (0,\varepsilon_0)$ we have

$$\tilde{\gamma}_\varepsilon(t) = \overline{\varepsilon^{-1} f + g_\varepsilon}(t) - \overline{\varepsilon^{-1} f}(t)\,.$$

Let $L = \sup_{\varepsilon > 0} ||g_\varepsilon||_t$, which is finite by assumption. For each $\varepsilon \in (0,\varepsilon_0)$ choose $s_\varepsilon(t) \in [0,t]$ to satisfy

$$\left(\varepsilon^{-1} f + g_\varepsilon\right)(s_\varepsilon(t)) \geq \overline{\varepsilon^{-1} f + g_\varepsilon}(t) - 8L\varepsilon.$$

Then by Lemma A.3(2) and the definition of the supremum, it is clear that $\overline{f}(t) - 8L\varepsilon \leq f(s_\varepsilon(t)) \leq \overline{f}(t)$, and consequently

$$\lim_{\varepsilon \downarrow 0} f(s_\varepsilon(t)) = \overline{f}(t). \tag{6.1}$$

Moreover, clearly

$$\begin{aligned} \tilde{\gamma}_\varepsilon(t) &\leq g_\varepsilon(s_\varepsilon(t)) + 8L\varepsilon + \varepsilon^{-1}\left[f(s_\varepsilon(t)) - \overline{f}(t)\right] \\ &\leq g_\varepsilon(s_\varepsilon(t)) + 8L\varepsilon, \end{aligned}$$

and therefore

$$\limsup_{\varepsilon \downarrow 0} \tilde{\gamma}_\varepsilon(t) \leq \limsup_{\varepsilon \downarrow 0} g_\varepsilon(s_\varepsilon(t)). \tag{6.2}$$

We now show that

$$\limsup_{\varepsilon\downarrow 0} g_\varepsilon(s_\varepsilon(t)) \le \tilde{\gamma}(t). \tag{6.3}$$

Select a sequence $\{\varepsilon_n\}$ with $\varepsilon_n \downarrow 0$ such that

$$\lim_{n\uparrow\infty} g_{\varepsilon_n}(s_{\varepsilon_n}(t)) = \limsup_{\varepsilon\downarrow 0} g_\varepsilon(s_\varepsilon(t)). \tag{6.4}$$

Since $\{s_{\varepsilon_n}(t)\} \subset [0,t]$ is uniformly bounded, it can be assumed without loss of generality (by choosing a subsequence if necessary) that there exists $s_0(t) \in [0,t]$ such that $\lim_{n\to\infty} s_{\varepsilon_n}(t) = s_0(t)$. By choosing a further subsequence if necessary, it can be assumed that either (i) $s_{\varepsilon_n}(t) = s_0(t)$ for all n sufficiently large, or (i) does not hold and either $s_{\varepsilon_n}(t) \uparrow s_0(t)$ or $s_{\varepsilon_n}(t) \downarrow s_0(t)$ as $n \to \infty$ If (i) holds, so that $s_{\varepsilon_n}(t) = s_0(t)$ for all $n \in I\!N$ sufficiently large, then (6.1) implies $f(s_0(t)) = \overline{f}(t)$, so that $s_0(t) \in \Phi_f(t)$. In that case

$$\limsup_{\varepsilon\downarrow 0} g_\varepsilon(s_\varepsilon(t)) = \lim_{n\uparrow\infty} g_{\varepsilon_n}(s_0(t)) = g(s_0(t)) \le \sup_{s\in\Phi_f(t)} g(s) \le \tilde{\gamma}(t),$$

and (6.3) holds. Now suppose that (i) above does not hold, but instead $s_{\varepsilon_n}(t) \uparrow s_0(t)$ as $n \uparrow \infty$. Then $s_0(t) \in \Phi_f^L(t)$ due to (6.1). Fix $\delta > 0$ and given $\varepsilon_m > 0$ choose $N(m) \ge m$ such that for all $n \ge N(m)$

$$g_{\varepsilon_m}(s_{\varepsilon_n}(t)) \le g_{\varepsilon_m}(s_0(t)-) + \delta.$$

The fact that $\{g_{\varepsilon_n}\}$ is a monotone non-increasing sequence as $n \uparrow \infty$ then shows that for all $n \ge N(m)$

$$g_{\varepsilon_n}(s_{\varepsilon_n}(t)) \le g_{\varepsilon_m}(s_0(t)-) + \delta.$$

Take limits as $n \uparrow \infty$ and then $m \uparrow \infty$ to obtain

$$\lim_{n\uparrow\infty} g_{\varepsilon_n}(s_{\varepsilon_n}(t)) \le \lim_{m\uparrow\infty} g_{\varepsilon_m}(s_0(t)-) + \delta.$$

Send $\delta \downarrow 0$ in the above display, and use (6.4) and the definition of g_l^* to conclude that

$$\limsup_{\varepsilon\downarrow 0} g_\varepsilon(s_\varepsilon(t)) \le g_l^*(s_0(t)) \le \sup_{s\in\Phi_f^L(t)} g_l^*(s) \le \tilde{\gamma}(t).$$

Lastly if (i) does not hold but $s_{\varepsilon_n}(t) \downarrow s_0(t)$ as $n \uparrow \infty$, it must be that $s_0(t) \ne t$ (since $s_{\varepsilon_n}(t) \in [0,t]$), and $f(s_0(t)+) = \overline{f}(t)$ due to (6.1). Thus $s_0(t) \in \tilde{\Phi}_f^R(t)$, and arguments similar to those given above yield

$$\limsup_{\varepsilon\downarrow 0} g_\varepsilon(s_\varepsilon(t)) = g_r^*(s_0(t)) \le \lim_{\varepsilon\downarrow 0} g_\varepsilon(s_0(t)+) \le \sup_{s\in\tilde{\Phi}_f^R(t)} g_r^*(s) \le \tilde{\gamma}(t).$$

This establishes (6.3), which when combined with (6.2) shows that

$$\limsup_{\varepsilon\downarrow 0} \tilde{\gamma}_\varepsilon(t) \le \tilde{\gamma}(t). \tag{6.5}$$

In order to establish the reverse inequality, first note that for any $r \in \Phi_f(t)$

$$\tilde{\gamma}_\varepsilon(t) \ge \varepsilon^{-1} f(r) + g_\varepsilon(r) - \varepsilon^{-1}\overline{f}(t) = g_\varepsilon(r).$$

Take limits as $\varepsilon \downarrow 0$ and the supremum over $r \in \Phi_f(t)$ in the last display to obtain

$$\liminf_{\varepsilon\downarrow 0} \tilde{\gamma}_\varepsilon(t) \geq \sup_{r\in\Phi_f(t)} g(r). \tag{6.6}$$

Now let $r \in \Phi_f^L(t)$ and for each $\varepsilon > 0$ choose $r_\varepsilon(t) \in [t-\varepsilon, t]$ such that

$$\varepsilon^{-1}\left[f(r_\varepsilon(t)) - \overline{f}(t)\right] > -\frac{\varepsilon}{2} \qquad \text{and} \qquad |g_\varepsilon(r_\varepsilon(t)) - g_\varepsilon(r-)| < \frac{\varepsilon}{2}.$$

Then

$$\begin{aligned} \tilde{\gamma}_\varepsilon(t) &\geq \varepsilon^{-1} f(r_\varepsilon(t)) + g_\varepsilon(r_\varepsilon(t)) - \varepsilon^{-1}\overline{f}(t) \\ &> g_\varepsilon(r-) - \varepsilon. \end{aligned}$$

Take limits as $\varepsilon \downarrow 0$, recall the definition of g_l^* and take the supremum over $r \in \Phi_f^L(t)$ to show that

$$\liminf_{\varepsilon\downarrow 0} \tilde{\gamma}_\varepsilon(t) \geq \sup_{r\in\Phi_f^L(t)} g_l^*(r).$$

Analogous arguments show that

$$\liminf_{\varepsilon\downarrow 0} \tilde{\gamma}_\varepsilon(t) \geq \sup_{r\in\tilde{\Phi}_f^L(t)} g_r^*(r).$$

The last two displays together with (6.6) yield

$$\liminf_{\varepsilon\downarrow 0} \tilde{\gamma}_\varepsilon(t) \geq \sup_{r\in\Phi_f^L(t)} [g_l^*(r)] \vee \sup_{r\in\Phi_f(t)} [g(r)] \vee \sup_{r\in\tilde{\Phi}_f^R(t)} [g_r^*(r)] = \tilde{\gamma}(t).$$

The last display along with (6.5) shows that for $t > t_u$

$$\lim_{\varepsilon\downarrow 0} \tilde{\gamma}_\varepsilon(t) = \tilde{\gamma}(t).$$

For $t \in [t_\ell, t_u]$, note that $\overline{f}(t) = 0$ and so

$$\tilde{\gamma}_\varepsilon(t) = \overline{\varepsilon^{-1} f + g_\varepsilon}(t) \vee 0\,.$$

Using the same reasoning that was used for the case $t > t_u$, it is possible to show that for $t \in [t_l, t_u]$,

$$\lim_{\varepsilon\downarrow 0} \tilde{\gamma}_\varepsilon(t) = \left[\sup_{s\in\Phi_f^L(t)} \{g_l^*(s)\} \vee \sup_{s\in\Phi_f(t)} \{g(s)\} \vee \sup_{s\in\tilde{\Phi}_f^R(t)} \{g_r^*(s)\}\right] \vee 0 = \tilde{\gamma}(t)\,.$$

Lastly when $t < t_\ell$, then $\overline{\varepsilon^{-1} f}(t) < 0$, and since the family $\{g_\varepsilon\}$ is uniformly bounded, relation (1.5) of Lemma A.3 implies that for all ε sufficiently small $\overline{\varepsilon^{-1} f + g}(t) < 0$. Hence for all sufficiently small $\varepsilon > 0$, $\tilde{\gamma}_\varepsilon(t) = 0$. Since $\tilde{\gamma}(t) = 0$ for $t < t_l$, this completes the proof of (4.28) in Theorem 4.7.

The identity (4.29) follows from (4.28), Lemma 6.1(3) and the definition of S_1 in (4.23), while the relation (4.30) follows directly from the definition of M_1 and the fact that f is continuous.

When f and g_ε are continuous, then $\Phi^L_f(t) = \Phi_f(t)$ and $\tilde{\Phi}^R_f(t) = \Phi_f(t) \setminus \{t\}$ and so from (4.26) it follows that

$$S_2(f,g) = \sup_{s\in\Phi_f(t)} [g(s-) \vee g(s)] \vee \sup_{s\in\Phi_f(t)\setminus\{t\}} [g(s+)] \,.$$

However, since $g_\varepsilon \downarrow g$ and g_ε are continuous, g is upper semicontinuous and so $g(s) \geq g(s+) \vee g(s-)$. Thus

$$S_2(f,g) = \sup_{s\in\Phi_f(t)} g(s) = S_1(f,g),$$

from which (4.31) follows. Finally, uf $f \in \mathcal{D}_c$ it is easy to see that $\gamma_\varepsilon \downarrow \gamma$ in the uniform topology, and so (4.32) is a simple consequence of Lemma 6.1(4), (4.25) and (4.28). ■

6.2 Classification of the discontinuities of $\gamma_{(1)}$

As shown in Theorem 4.10 for ORPs with H-R constraint matrices R, given any $\psi,\chi \in \mathcal{C}$, the derivative $\nabla_\chi\Gamma(\psi)$ satisfies $\nabla_\chi\Gamma(\psi) = \chi + R\gamma_{(1)}(\psi,\chi)$, where $\gamma_{(1)}(\psi,\chi)$ is the unique solution to the system of equations (4.36) with $j = 1$. In this section we prove Theorem 5.2, which provides a classification of the discontinuities of $\gamma_{(1)}$. In Section 6.2.1 we establish a correspondence between the regimes introduced in Definition 5.1 and properties of the sets $\Phi_{-\xi^i}(t)$, stated in Lemma 5.1. In Section 6.2.2 we establish some preliminary lemmas required for the proof of Theorem 5.2, which is presented in Section 6.2.3.

We now introduce some new notation used in this section. For $t \in [0,\infty)$ let $Disc(t)$ (respectively $LDisc(t)$, $RDisc(t)$ and $SDisc(t)$) denote the set of coordinates i such that $\gamma^i_{(1)}$ has a discontinuity (respectively left discontinuity, right discontinuity and separated discontinuity) at t. Clearly

$$Disc(t) = LDisc(t) \cup RDisc(t) \qquad \text{and} \qquad SDisc(t) = LDisc(t) \cap RDisc(t).$$

The other notation used in this section has been made consistent with that used to define $\gamma_{(1)}$ in Section 4.2.1, and thus this notation will not be explicitly redefined here. Finally, to simplify notation, throughout the rest of this section we will use just γ to denote $\gamma_{(1)}$.

6.2.1 Characterization of Regimes

In Lemma 5.1 the various regimes in Definition 5.1 are equivalently characterized in terms of conditions on the set $\Phi^i(t)$. Here we provide a proof of the lemma.

Proof of Lemma 5.1. Fix $i \in \mathcal{I}$. Recall from (4.4) that for $s > t^i_u$ $\xi^i(s) = \psi^i(s) - [P\theta]^i(s)$, $\theta^i(s) = \overline{-\xi^i}(s) \vee 0$ and

$$\phi^i(s) = \xi^i(s) + \theta^i(s) = \xi^i(s) + \overline{-\xi^i}(s) \vee 0.$$

It is clear from the definitions that $-\xi^i(t) < 0$ for $t \in [0,t^i_l)$. This implies that $\theta^i(t) = 0$ and $\phi^i(t) > 0$, and hence that i is overloaded. For $t \in [t^i_l,t^i_u)$, $-\xi^i(t) = \overline{-\xi^i}(t) = 0$, and so $\phi^i(t) = \theta^i(t) = \dot{\theta}_i(t-) = 0$, which shows that i is critical at t.

It only remains to prove Lemma 5.1 for $t > t^i_u$. If $t > t^i_u$ and $s \in \Phi_{-\xi^i}(t)$ then $s > t^i_u$ and $\overline{-\xi^i}(t) = -\xi^i(s) = \overline{-\xi^i}(s) > 0$. Along with the last display this implies $\phi^i(s) = 0$ and $\theta^i(s) = \theta^i(t)$. Conversely, if $\phi^i(s) = 0$ and $\theta^i(s) = \theta^i(t)$, then $-\xi^i(s) = \overline{-\xi^i}(s) = \overline{-\xi^i}(t) \vee 0 = \overline{-\xi^i}(t)$, where the last equality uses the fact that $t \in [t_u^i, \infty)$. Thus $s \in \Phi_{-\xi^i}(t)$ and the proof of the representation (5.2) is complete.

The representation (5.2) implies that $t \in \Phi_{-\xi^i}(t)$ if and only if $\phi^i(t) = 0$, and that $s \in \Phi_{-\xi^i}(t) \cap (0,t)$ implies $\dot{\theta}^i(t-) = 0$, which establishes (3) and the "if" statement in (4). On the other hand, if $\Phi_{-\xi^i}(t) = \{t\}$ and $t > t_u^i$ then $\phi^i(t) = 0$ and $\theta^i(s) = \overline{-\xi^i}(s) < \overline{-\xi^i}(t)$ for every $s < t$, which implies $\dot{\theta}^i(t-) > 0$. This proves the "only if" statement in (4) and also establishes (5).

From (4) it follows that for both the end of overloading and start of underloading $t \in \Phi_{-\xi^i}(t) \neq \{t\}$. Now from (3) we know that if $i \in Over(s)$ for all $s \in (t-\delta, t)$, then $s \notin \Phi_{-\xi^i}(s)$, and therefore $s \notin \Phi_{-\xi^i}(t)$. This implies that there exists $\delta > 0$ such that $(t-\delta, t) \cap \Phi_{-\xi^i}(t) = \emptyset$, so that t is isolated in $\Phi_{-\xi^i}(t)$, which establishes (6). On the other hand, from (5) we know that $i \in Under(s)$ for $s \in (t, t+\delta)$ if and only if $\Phi_{-\xi^i}(s) = \{s\}$ for $s \in (t, t+\delta)$, which in turn implies that $\Phi_{-\xi^i}(s) \subset (t, s]$ for every $s \in (t, t+\delta)$. This proves (7) and hence completes the proof of the lemma. ■

6.2.2 Preliminary Lemmas

In this section we establish some results that are required for the proof of Theorem 5.2. For $i = 1, \ldots, K$ define

$$\tilde{t}_u^i \doteq \inf\{t \geq t_l^i : -\chi^i(t) + [P\gamma]^i(t) > 0 \text{ and } \overline{-\xi^i}(t) = 0\}, \tag{6.7}$$

where we adopt the convention $\inf \emptyset = \infty$.

Lemma 6.2 *Suppose $\gamma_{(1)}$ and $\tilde{t}_u^i$ are defined by (4.36) and (6.7) respectively. For $i = 1, \ldots, K$ $\gamma_{(1)}^i(t) = 0$ for $t \in [0, \tilde{t}_u^i)$, and*

$$\gamma_{(1)}^i(t) = \sup_{s \in \Phi_{-\xi^i}(t)} \left[-\chi^i(s) + [P\gamma]^i(s)\right] \tag{6.8}$$

for $t \in [\tilde{t}_u^i, \infty)$. Moreover, for $t \in [0, \infty)$, if $i \in Disc(t)$ then $t \in \Phi_{-\xi^i}(t)$ and

$$\gamma_{(1)}^i(t) = [-\chi^i(t) + [P\gamma]^i(t)] \vee \gamma_{(1)}^i(t-). \tag{6.9}$$

Furthermore, if $i \in LDisc(t)$ then

$$\gamma_{(1)}^i(t) = -\chi^i(t) + [P\gamma_{(1)}]^i(t) > \gamma_{(1)}^i(t-), \tag{6.10}$$

while if $i \in RDisc(t)$ then $\Phi_{-\xi^i}(s) \subset (t, s]$ for all $s > t$ and

$$\gamma_{(1)}^i(t+) = -\chi^i(t) + [P\gamma_{(1)}]^i(t+). \tag{6.11}$$

Proof. For conciseness $\gamma_{(1)}$ will be denoted simply by γ for the entire proof. The claim that $\gamma^i = 0$ for $t \in [0, \tilde{t}_u^i)$ and the relation in (6.8) follow directly from the representation of γ^i given in (1.19) and the definition of $\tilde{t}_u^i$. Note that γ^i has no points of right discontinuity in $[0, \tilde{t}_u^i]$. If $\tilde{t}_u^i = \infty$ then $\gamma^i(t_u^i) = 0$ and there are no points of left discontinuity either in $[0, \tilde{t}_u^i]$. However, if $\tilde{t}_u^i < \infty$ is a point of left discontinuity for γ^i, then

$$\gamma^i(\tilde{t}_u^i-) = 0 < g(\tilde{t}_u^i) = -\chi^i(\tilde{t}_u^i) + [P\gamma]^i(\tilde{t}_u^i),$$

so that (6.10) is satisfied at $t = \tilde{t}^i_u$.

Thus it only remains to prove the lemma for $t \in [\tilde{t}^i_u, \infty)$, where the representation (6.8) holds. Since $\xi \in \mathcal{C}$, if $t \notin \Phi_{-\xi^i}(t)$ there must exist $\delta > 0$ such that

$$-\xi^i(s) < \overline{-\xi^i}(t) \quad \text{and} \quad \Phi_{-\xi^i}(s) = \Phi_{-\xi^i}(t) \qquad \text{for } s \in [t-\delta, t+\delta].$$

This implies that $\gamma^i(s) = \gamma^i(t)$ for $s \in [t-\delta, t+\delta]$, which establishes the continuity of γ^i at t. Thus a necessary condition for γ^i to have a (right or left) discontinuity at t is that $t \in \Phi_{-\xi^i}(t)$. When combined with the representation (6.8) of γ^i, this implies that $\gamma^i(t) \geq -\chi^i(t) + [P\gamma]^i(t)$. Along with the inequality $\gamma^i(t) \geq \gamma^i(t-)$ for all $t \in [0, \infty)$ (which holds due to the upper semicontinuity of γ^i proved in Theorem 4.6), this implies

$$\gamma^i(t) \geq [-\chi^i(t) + [P\gamma]^i(t)] \vee \gamma^i(t-) \qquad \text{if } i \in Disc(t).$$

If $i \notin LDisc(t)$ then $\gamma^i(t) = \gamma^i(t-)$ which, together with the last display, implies (6.9). If $i \in LDisc(t)$ then either $\{t\} = \Phi_{-\xi^i}(t)$ or $t \in \Phi_{-\xi^i}(t) \neq \{t\}$. In the former case the definition of γ^i directly implies $\gamma^i(t) = -\chi^i(t) + [P\gamma^i](t)$, which in turn implies (6.9). In the latter case $\Phi_{-\xi^i}(r) = \Phi_{-\xi^i}(t) \cap [0, r]$ for all $r < t$ sufficiently close to t, and hence

$$\gamma^i(t-) = \lim_{r \uparrow t} \gamma^i(r) = \lim_{r \uparrow t} \sup_{s \in \Phi_{-\xi^i}(r)} [-\chi^i(s) + [P\gamma]^i(s)] = \sup_{s \in \Phi_{-\xi^i}(t) \setminus \{t\}} [-\chi^i(s) + [P\gamma]^i(s)].$$

Therefore

$$\gamma^i(t) = \sup_{s \in \Phi_{-\xi^i}(t)} [-\chi^i(s) + [P\gamma]^i(s)] = [-\chi^i(t) + [P\gamma]^i(t)] \vee \gamma^i(t-)$$

which, along with the fact that $\gamma^i(t) \geq \gamma^i(t-)$, establishes (6.9) and (6.10).

To prove (6.11), suppose $i \in RDisc(t)$. If there exists $\delta > 0$ such that $-\xi^i(s) \leq -\xi^i(t)$ for $s \in [t, t+\delta]$, then for $s \in (t, t+\delta]$ $\Phi_{-\xi^i}(s) = \Phi_{-\xi^i}(t) \cup A_s$ where $A_s \subset (t, t+\delta]$. This in turn implies that for $s \in (t, t+\delta]$

$$\gamma^i(t) \leq \gamma^i(s) = \gamma^i(t) \vee \sup_{u \in A_s} \left[-\chi^i(u) + [P\gamma]^i(u) \right].$$

Take limits as $s \downarrow t$ to obtain

$$\gamma^i(t) \leq \gamma^i(t+) = \gamma^i(t) \vee \left[-\chi^i(t) + [P\gamma]^i(t+) \right] \leq \gamma^i(t) \vee \left[-\chi^i(t) + [P\gamma]^i(t) \right] = \gamma^i(t),$$

where the inequality in the above display holds because $[P\gamma]^i$ is upper semicontinuous by Corollary (4.3), and the last equality follows from (6.9). This shows that γ^i is right continuous at t whenever $-\xi^i$ is locally non-decreasing to the right of t. Thus a necessary condition for $i \in RDisc(t)$ is that there exist a sequence $s_n \downarrow t$ such that for every $n \in I\!N$

$$\overline{-\xi^i}(s_n) > \overline{-\xi^i}(t),$$

which in turn is equivalent to the condition that $\Phi_{-\xi^i}(s) \subset (t, s]$ for all $s > t$. As a consequence, for any sequence $s_n \downarrow t$

$$\gamma^i(s_n) = \sup_{s \in \Phi_{-\xi^i}(s_n)} [-\chi^i(s) + [P\gamma]^i(s)] = \sup_{s \in \Phi_{-\xi^i}(s_n) \cap (t, s_n]} [-\chi^i(s) + [P\gamma]^i(s)].$$

Send $s_n \downarrow t$ and observe that $-\chi^i$ is continuous at t to obtain (6.11). ■

Lemma 6.3 *Let* $\gamma \doteq \gamma_{(1)}$ *and* $\tilde{t}_u^i$ *be defined by (4.36) and (6.7) respectively. If*

$$\{t\} = \Phi_{-\xi^i}(t) \text{ for } t \in [t_u^i, \infty) \tag{6.12}$$

then

$$\gamma^i(t-) = -\chi^i(t) + [P\gamma]^i(t-), \tag{6.13}$$

and

$$\gamma^i(t) = -\chi^i(t) + [P\gamma]^i(t), \tag{6.14}$$

and $\nabla_\chi \Gamma(\psi)^i$ *is left continuous at* t*. If*

$$\Phi_{-\xi^i}(s) \subset (t, s] \quad \text{for all} \quad s > t, \tag{6.15}$$

then

$$\gamma^i(t+) = -\chi^i(t) + [P\gamma]^i(t+). \tag{6.16}$$

Moreover, if (6.12) holds along with (6.15), then t *is a point of right continuity for* $\nabla_\chi \Gamma(\psi)$*.*

Proof. Suppose $\{t\} = \Phi_{-\xi^i}(t)$ for some $t \in [\tilde{t}_u^i, \infty)$. Let s_n be an increasing sequence such that $s_n \uparrow t$, and let $u_n \in [0, s_n]$ satisfy

$$u_n = \min\{u \in [0, s_n] : \overline{-\xi^i}(u) = \overline{-\xi^i}(s_n)\}.$$

We claim that then $u_n \uparrow t$. Indeed, since u_n is uniformly bounded, there exists a convergent subsequence (which we denote again by u_n) that converges to a limit $u_* \in [0, t]$. Since ξ^i is continuous, clearly $u_* \in \Phi_{-\xi^i}(t)$. Since $\Phi_{-\xi^i}(t) = \{t\}$ by assumption, we conclude that $u_* = t$. Now observe that

$$\gamma^i(s_n) = \max_{s \in \Phi_{-\xi^i}(s_n)} [-\chi^i(s) + [P\gamma]^i(s)] = \max_{s \in \Phi_{-\xi^i}(s_n) \cap [u_n, s_n]} [-\chi^i(s) + [P\gamma]^i(s)].$$

Take limits as $n \uparrow \infty$ on both sides of the above equality and use the fact that $u_n \uparrow t$ to obtain (6.13). The definition of γ^i automatically gives (6.14). Combining (6.13) with (6.14) one observes that $\nabla \Gamma^i = \gamma^i - [P\gamma]^i + \chi^i$ is left continuous at t.

If (6.15) holds then given a sequence $s_n \downarrow t$ it follows that

$$\gamma^i(s_n) = \sup_{s \in \Phi_{-\xi^i}(s_n)} [-\chi^i(s) + [P\gamma]^i(s)] = \sup_{s \in \Phi_{-\xi^i}(s_n) \cap (t, s_n]} [-\chi^i(s) + [P\gamma]^i(s)],$$

from which (6.16) follows. If in addition (6.12) holds, then (6.14) along with (6.16) shows that $\nabla_\chi \Gamma(\psi)$ is right continuous. ■

6.2.3 Proof of Theorem 5.2

In this section we prove Theorem 5.2. The proof makes use of Lemma 5.1(2) to replace the sets Φ^i in the statement of Theorem 5.2 by $\Phi_{-\xi^i}$.

Proof. For simplicity of notation, $\gamma_{(1)}$ is denoted by γ for the rest of the proof. Recall from Lemma 6.2 that $t \in \Phi_{-\xi^i}(t)$ is a necessary condition for γ^i to have a (right or left) discontinuity, and so assume for the rest of this proof that this condition holds. The proof is presented only for the case $t \in [\tilde{t}_u^i, \infty)$, since the proof for $t \in [t_l^i, \tilde{t}_u^i)$ is analogous.

We first analyze the left discontinuities of γ^i. In [22] properties of the one-dimensional derivative were studied for the case when $\psi, \chi \in \mathcal{C}$. Consequently, when $[P\gamma]^i$ is left continuous at t, the fact that γ^i has a point of left discontinuity at t only if $\{t\} \neq \Phi_{-\xi^i}(t)$ and t is isolated in $\Phi_{-\xi^i}(t)$ follows from [22, Lemmas 6.5 and 6.6]. Under this condition one has

$$\gamma^i(t-) = \sup_{s \in \Phi_{-\xi^i}(t) \setminus \{t\}} [-\chi^i(s) + [P\gamma]^i(s)].$$

If (5.4) does not hold then clearly $\gamma^i(t-) = \gamma^i(t)$ so that γ^i is continuous at t, while if (5.4) does hold, then clearly $i \in LDisc(t)$ and

$$\gamma^i(t) = -\chi^i(t) + [P\gamma]^i(t),$$

which also implies that $\nabla\Gamma^i(t) = 0$. The last two displays and the fact that $[P\gamma]^i$ is continuous then show that

$$\nabla\Gamma^i(t-) = \gamma^i(t-) - [P\gamma]^i(t-) + \chi^i(t-) = \gamma^i(t-) - [P\gamma]^i(t) + \chi^i(t) < 0,$$

which completes the justification of condition L1.

Now consider the case when $[P\gamma]^i$ is left discontinuous. Suppose first that $\{t\} = \Phi_{-\xi^i}(t)$. Then (6.13) and (6.14) of Lemma 6.3 show that γ^i is left discontinuous, but $\nabla\Gamma^i$ is left continuous, which establishes L4. Now suppose that $t \in \Phi_{-\xi^i}(t) \neq \{t\}$. Then

$$\gamma^i(t-) = \sup_{s \in \Phi_{-\xi^i}(t) \setminus \{t\}} [-\chi^i(s) - [P\gamma]^i(s)],$$

and it follows that (5.4) is necessary and sufficient for $i \in LDisc(t)$. Furthermore, in that case clearly $\gamma^i(t) = -\chi^i(t) + [P\gamma]^i(t)$, so that $\nabla\Gamma^i(t) = 0$, thus proving L2. If, in addition, t is not isolated in $\Phi_{-\xi^i}(t)$, i.e. there exists $\delta > 0$ such that $[t-\delta, t] \subset \Phi_{-\xi^i}(t)$, then

$$\gamma^i(t-) \geq \sup_{s \in [t-\delta, t)} [-\chi^i(s) - [P\gamma]^i(s)] \geq -\chi^i(t) + [P\gamma]^i(t-),$$

and so $\nabla\Gamma^i(t-) \geq \nabla\Gamma^i(t) = 0$, which establishes L3. Note that when t is isolated in $\Phi_{-\xi^i}(t)$ there exist scenarios where $\nabla\Gamma^i(t) < \nabla\Gamma^i(t-)$ and where $\nabla\Gamma^i(t) \geq \nabla\Gamma^i(t-)$.

We now consider the right discontinuities of γ^i. It follows from Lemma 6.2 that $t \in \Phi_{-\xi^i}(t)$ and $\Phi_{-\xi^i}(s) \subset (t, s]$ are necessary conditions for $i \in RDisc(t)$. The fact that $t \in \Phi_{-\xi^i}(t)$ shows that $\gamma^i(t) \geq -\chi^i(t) + [P\gamma]^i(t)$, the upper semicontinuity of γ^i (which was proved in Corollary 4.3) dictates that $\gamma^i(t) > \gamma^i(t+)$, and (6.11) implies $\gamma^i(t+) = -\chi^i(t) + [P\gamma]^i(t+)$. When combined, these three statements yield

$$\nabla\Gamma^i(t) \geq \nabla\Gamma^i(t+) = 0.$$

If $[P\gamma]^i$ is right continuous at t, then the previous two statements clearly imply that $i \in RDisc(t)$ if and only if (5.5) holds (i.e. if $\gamma^i(t) > -\chi^i(t) + [P\gamma]^i(t)$). If $[P\gamma]^i$ is right discontinuous at t, then the uppersemicontinuity of $[P\gamma]^i$ (which follows from Corollary 4.3 and the non-negativity of P), the fact that $t \in \Phi_{-\xi^i}(t)$, the definition of γ and (6.16) show that

$$\gamma^i(t) \geq -\chi^i(t) + [P\gamma]^i(t) > -\chi^i(t) + [P\gamma]^i(t+) = \gamma^i(t+), \tag{6.17}$$

so that $i \in RDisc(t)$. Now if (5.5) holds then clearly $\{t\} \neq \Phi_{-\xi^i}(t)$ and also $i \notin LDisc(t)$ by (6.10) of Lemma 6.2. Moreover, in this case

$$\nabla\Gamma^i(t) > \nabla\Gamma^i(t+) = 0.$$

This completes the proof of conditions R1 and R2. If $\{t\} = \Phi_{-\xi^i}(t)$ then the first inequality in (6.17) can be replaced by equality, and so $\nabla\Gamma^i(t) = \nabla\Gamma^i(t+) = 0$, which establishes conditions R3 and R4.

Finally, a separated discontinuity can only happen when $i \in LDisc(t) \cap RDisc(t)$. However if conditions R1 or R2 are satisfied, then $i \notin LDisc(t)$, and therefore a separated discontinuity can only happen when the right discontinuity is of type R3 or R4. If R3 holds, then L4 cannot be true, and thus one must have either L1, L2 or L3, which leads to conditions S1 and S2. On the other hand if R4 holds and $i \in LDisc(t)$, then L4 must hold, which results in S3. The statements about $\nabla\Gamma^i$ in this case follow directly by combining the statements about $\nabla\Gamma^i$ made for the relevant left and right discontinuous cases. This completes the proof of the theorem. ■

7 Conclusions

In this work we introduced the notion of the directional derivative of the multi-dimensional oblique reflection map $\nabla_\chi\Gamma(\psi)$. For the class of Harrison-Reiman oblique reflection problems, which arise when analyzing models of open single-class Jackson queueing networks, and continuous ψ, χ, we established an autonomous representation for $\nabla_\chi\Gamma(\psi)$. This representation was then used to identify properties of the derivative. In particular, we showed that the discontinuities of the derivative have an intuitive interpretation in terms of transitions in the regimes of $\Gamma(\psi)$, the image of ψ under the reflection map Γ. We also presented examples to illustrate the fact that the derivative can be explicitly calculated in many situations.

One of our main motivations for introducing the derivative is its connection with diffusion approximations of non-stationary queueing networks and of stationary networks that exhibit transient behaviour [25]. In this context, $\Gamma(\psi)$ corresponds to the fluid limit of the network, and jumps in the diffusion approximation occur at certain transitions of the fluid limit from one state (of underloading, critical loading or overloading at a node) to another. In this work we showed that these jumps are influenced by the topology of the network. In particular, fluctuations in one part of the network can influence another part of the network at time t only if there is either a path of empty buffers connecting one with the other, or there is a cycle in the network. It would be of interest to understand the implications of the characterizations derived in this work for the simulation of time-dependent networks. Another global objective is to use the qualitative insights gained from these diffusion approximations to design appropriate controls for mitigating undesired effects due to non-stationarity.

This work has concentrated on the conventional heavy-traffic regime. Another asymptotic regime that has recently gained considerable attention is the so-called Halfin-Whitt regime [11], where the number of servers is scaled along with the arrival rate while the service rate of each server remains fixed (see [18] for a motivation of this regime). Functional strong laws and functional central limits for a large class of time-dependent Markovian networks in the Halfin-Whitt regime have been obtained in [23]. Notably, asymptotic analysis of the Halfin-Whitt regime has been restricted to service durations that are exponential [11], of phase-type [35] and deterministic [16]. The complete analysis in the case of general service distributions remains an important open problem. We believe that the general methodology of characterizing functional central limits in terms of directional derivatives of an appropriate mapping, as outlined in Section 1.1 for the conventional heavy traffic regime, may also be applicable to the Halfin-Whitt asymptotic regime. The main difference lies only in the definitions of the processes $\bar{X}^n$ and the mapping F. Specifically, while for the conventional heavy traffic regime the mapping F is the oblique reflection mapping, it can be shown (see, for example, [23, 26]) that in the Halfin-Whitt regime with exponential service times, the appropriate mapping F is defined as follows: given non-decreasing functions f_A and f_S, $F(f_A, f_S) = q$, where q is the unique function that solves the integral equation

$$q(t) = q(0) + f_A(t) - \int_0^t [q(r) \wedge 1] \, df_S(r) \qquad \text{for } t \in [0, \infty),$$

in which df_S is the Lebesgue-Stieltjes measure on $[0, \infty)$ induced by f_S. Indeed, when f_A is the functional strong law of large numbers limit of the normalized cumulative arrival process and f_S is the functional strong law of large numbers limit of the normalized sum of the cumulative service processes of all N servers, respectively, then $F(f_A, f_S)$ is the functional strong law of large numbers limit (or fluid) limit of the queue length process [23]. Furthermore, the functional central limit of the queue process can be expressed in terms of directional derivatives of the mapping F (see [23, 26] for details). We expect that, in the case of general service distributions, the corresponding mapping F may be more complicated but that the general approach would still remain valid.

A Auxiliary Results

Lemma A.1 *Given real-valued functions f and g*

$$\overline{f} \vee 0 - \overline{g} \vee 0 \leq \overline{f - g} \vee 0 \, .$$

Moreover, if $f, g \in \mathcal{BV}$ then

$$\left| \overline{f} \vee 0 - \overline{g} \vee 0 \right|_T \leq |f - g|_T \, ,$$

where $| \cdot |_T$ represents the total variation norm.

Proof. This follows trivially from a case-by-case verification, and is thus omitted. ■

Lemma A.2 *Consider a sequence of functions $\{f_n\} \subset \mathcal{D}_{\lim}$ that converges pointwise to a bounded function f. If for every $T < \infty$ $\sup_n |f_n|_T < \infty$, then $f \in \mathcal{D}_{\lim}$.*

Proof. We argue by contradiction to prove the lemma. Suppose $f \notin \mathcal{D}_{\lim}$. Then there exists $t \in (0, \infty)$ such that f does not have either a left limit or a right limit at t. We assume without loss of generality that f does not have a left limit at t (the case when f does not have a right limit clearly follows by a similar argument). Since f is bounded there must exist $\delta > 0$ and sequences $\{s_i\}$ and $\{s'_i\}$ such that $s_i \uparrow t$, $s'_i \uparrow t$ and for all i, i'

$$|f(s_i) - f(s'_{i'})| \geq 4\delta. \tag{1.1}$$

By choosing further subsequences if necessary, we can assume without loss of generality that

$$s_1 < s'_1 < s_2 < s'_2 < \dots$$

Since $f_n \to f$ pointwise, given any $m \in I\!N$ there exists $N < \infty$ such that for all $n \geq N$

$$|f_n(s) - f(s)| < \delta \qquad \text{for all } s \in \{s_i, s'_i, i = 1, \dots, m\}.$$

Combining (1.1) with the last display we conclude that for every $i = 1, \dots, m$ and $n \geq N$

$$|f_n(s_i) - f_n(s'_i)| > \delta,$$

and hence

$$|f_n|_t \geq \sum_{i=1}^{m} |f_n(s_i) - f_n(s'_i)| \geq m\delta.$$

Taking the supremum over all n in the last display, and then letting m to go to infinity, we conclude that $\sup_n |f_n|_t = \infty$, which leads to a contradiction. Thus it must be that $f \in \mathcal{D}_{\lim}$. ■

Lemma A.3 *Suppose $f \in \mathcal{D}_{\lim}$ and $\{g_\varepsilon,\ \varepsilon > 0\} \subseteq \mathcal{D}_{\lim}$ satisfies*

$$L_T \doteq \sup_{\varepsilon > 0} \|g_\varepsilon\|_T < \infty \qquad \text{for every } \ T \in [0, \infty)\,. \tag{1.2}$$

Then the following properties hold for any $t \in [0, \infty)$.

1. *There exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$*

$$\overline{f}(t) < 0 \quad \Rightarrow \quad \overline{\varepsilon^{-1} f + g_\varepsilon}(t) < 0\,, \tag{1.3}$$

   and likewise

$$\overline{f}(t) > 0 \quad \Rightarrow \quad \overline{\varepsilon^{-1} f + g_\varepsilon}(t) > 0\,. \tag{1.4}$$

2. *Given $\delta \in (0, L_t)$, if*

$$\overline{\varepsilon^{-1} f + g_\varepsilon}(t) \leq \varepsilon^{-1} f(s) + g_\varepsilon(s) + \delta \tag{1.5}$$

   for some $s \in [0, t]$ and $\varepsilon \in (0, \delta/6L_t)$, then

$$\overline{f}(t) \leq f(s) + \delta\,. \tag{1.6}$$

Proof. Fix $t \in [0, \infty)$, let $L \doteq L_t$ and choose $\varepsilon_0 = |\overline{f}(t)|/2L$. If $\overline{f}(t) < 0$, then for $\varepsilon \in (0, \varepsilon_0)$

$$\begin{aligned} \overline{\varepsilon^{-1} f + g_\varepsilon} &\leq \varepsilon^{-1} \overline{f}(t) + \overline{g}_\varepsilon(t) \\ &\leq -2L + L \\ &< 0\,, \end{aligned}$$

which establishes (1.3). A similar argument establishes (1.4).

We now argue by contradiction to establish the second property. Suppose there exists $\varepsilon < \delta/6L_t$ and $s \in [0, t]$ that satisfies (1.5), but for which

$$f(s) < \overline{f}(t) - \delta\,.$$

Choose $\tilde{s} \in [0,t]$ such that

$$f(\tilde{s}) > \overline{f}(t) - \delta/2.$$

Then the last two displays together show that $f(\tilde{s}) > f(s) + \delta/2$, which along with (1.5) the fact that $\delta < L_t$ and $\varepsilon < \delta/6L_t$ imply that

$$\begin{aligned} & \varepsilon^{-1} f(\tilde{s}) + g_\varepsilon(\tilde{s}) - \overline{\varepsilon^{-1} f + g_\varepsilon}(t) \\ \geq\ & \varepsilon^{-1} f(\tilde{s}) + g_\varepsilon(\tilde{s}) - \Big(\varepsilon^{-1} f(s) + g_\varepsilon(s)\Big) - \delta \\ \geq\ & \varepsilon^{-1} \left[f(\tilde{s}) - f(s)\right] - 2L_t - \delta \\ >\ & \frac{\varepsilon^{-1}\delta}{2} - 3L_t \\ >\ & 0\,, \end{aligned}$$

which contradicts the definition of the supremum since $\tilde{s} \in [0,t]$. ■

Lemma A.4 *Consider a family of left (respectively right) continuous functions $\{g_\varepsilon\}$ that converges pointwise monotonically down to a function $g \in \mathcal{D}_{\lim}$ as $\varepsilon \downarrow 0$. If s is a point of left (respectively right) continuity for g, then given any sequence $s_\varepsilon \uparrow s$ (respectively $s_\varepsilon \downarrow s$)*

$$\lim_{\varepsilon \downarrow 0} g_\varepsilon(s_\varepsilon) = g(s).$$

Proof. Given any $\delta > 0$ the pointwise convergence of g_κ shows that there exists $\kappa_0 > 0$ such that for all $\kappa \in (0, \kappa_0)$

$$|g_\kappa(s) - g(s)| < \delta/3.$$

Likewise, given any $\kappa > 0$ since either g_κ is left continuous and $s_\varepsilon \uparrow s$, or g_κ is right continuous and $s_\varepsilon \downarrow 0$, there exists $\varepsilon_0(\kappa) \leq \kappa$ such that for all $\varepsilon < \varepsilon_0(\kappa)$

$$|g_\kappa(s_\varepsilon) - g_\kappa(s)| < \delta/3,$$

and

$$|g(s_\varepsilon) - g(s)| < \delta/3.$$

The last three displays, when combined, show that given any $\delta > 0$ there exists $\kappa_0 > 0$ such that for all $\kappa < \kappa_0$ and $\varepsilon < \varepsilon_0(\kappa)$,

$$|g_\kappa(s_\varepsilon) - g(s)| < \delta.$$

When combined with the fact that $\varepsilon < \varepsilon_0(\kappa) < \kappa$ and the pointwise monotonic convergence of g_ε down to g, the last display yields

$$g(s_\varepsilon) \leq g_\varepsilon(s_\varepsilon) \leq g_\kappa(s_\varepsilon) \leq g(s) + \delta.$$

Taking limits, first as $\varepsilon \downarrow 0$ and then as $\kappa \downarrow 0$, and using the left continuity of g and the fact that $s_\varepsilon \uparrow s$ (or the right continuity of g and the fact that $s_\varepsilon \downarrow s$), one concludes that

$$g(s) \leq \liminf_{\varepsilon \downarrow 0} g_\varepsilon(s_\varepsilon) \leq \limsup_{\varepsilon \downarrow 0} g_\varepsilon(s_\varepsilon) \leq g(s) + \delta.$$

Sending $\delta \downarrow 0$ concludes the proof of the lemma. ■

B List of Notations for Function Spaces

Given a real-valued function defined on $[0, \infty)$, a *separated discontinuity* is defined to be a point $t \in [0, \infty)$ at which

$$f(t) \notin [f(t-) \wedge f(t+), f(t-) \vee f(t+)].$$

An $I\!R^K$-valued function f is said to have a separated discontinuity at t if for some $i = 1, \ldots, K$ f^i has a separated discontinuity at t. The set of discontinuity points of a function $f \in \mathcal{D}_{\lim}$ will be denoted $Disc(f)$. Note that $Disc(f) = \cup_{i=1}^K Disc(f^i)$.

$\mathcal{D}_{\lim}$	the space of all functions on $[0, \infty)$ taking values in $\mathbb{R}^K$ that have left and right limits for every $t \in [0, \infty)$.
$\mathcal{D}_r$	the subspace of right continuous functions in $\mathcal{D}_{\lim}$
$\mathcal{D}_{\ell,r}$	the subspace of functions that are either right continuous or left continuous at every $t \in [0, \infty)$
$\mathcal{D}_{usc}$	the subspace of functions in $\mathcal{D}_{\lim}$ such that each coordinate function f^i is uppersemicontinuous (i.e. $f(t) \geq f(t-) \vee f(t+)$ for every $t \in [0, \infty)$.)
$\mathcal{D}_c$	the subspace of piecewise constant functions in $\mathcal{D}_r$ with a finite number of jumps.
$\mathcal{I}_+$	the space of functions in $\mathcal{D}_{\lim}$ taking values in $I\!R_+^K$ such that each coordinate function is non-decreasing.
$\mathcal{C}$	the subspace of continuous functions in $\mathcal{D}_{\lim}$.
$\mathcal{BV}$	the subspace of bounded variation functions in $\mathcal{D}_{\lim}$

Acknowledgements. The authors would like to thank Ward Whitt for his valuable comments, as well as for his encouragement throughout the course of this work. The authors are also grateful to Lillian Bluestein for her help in typing up parts of this manuscript. The second author would also like to thank Rami Atar, Avi Mandelbaum, Adam Shwartz and Ofer Zeitouni for an invitation to the Technion, which provided the leisure necessary to complete this paper.

References

[1] S. Asmussen and H. Thorisson. 1987. A Markov chain approach to periodic queues. *J. Appl. Probab.*, 24:215–225.

[2] N. Bambos and J. Walrand. 1989. On queues with periodic inputs. *J. Appl. Probab.*, 26:381–389.

[3] A. Bernard and A. El Kharroubi. 1991. Regulation de processus dans le premier orthant de $I\!R^n$. *Stoch. and Stoch. Rep.*, 34:149–167.

[4] P. Billingsley. 1968. *Convergence of Probability Measures.* John Wiley, New York.

[5] H. Chen and A. Mandelbaum. 1991. Discrete flow networks: Bottleneck analysis and fluid approximations. *Math. Oper. Res.*, 16:408–446.

[6] H. Chen and A. Mandelbaum. 1991. Discrete flow networks: Diffusion approximations and bottlenecks *Ann. Prob.*, 19:1463–1519.

[7] P. Dupuis and H. Ishii. 1991. On Lipschitz continuity of the solution mapping to the Skorokhod problem, with applications. *Stochastics*, 35:31–62.

[8] P. Dupuis and K. Ramanan. 1999. Convex duality and the Skorokhod Problem, I, II. *Prob. Theor. and Rel. Fields*, 115:153–195 (I); 197–236 (II).

[9] M. Gerla and L. Kleinrock. 1980. Flow control: A comparative survey. *IEEE Trans. Commun. vol. COM-29*, 4:553–574.

[10] P. Glynn. 1998. Strong approximations in queueing theory. *Asymptotic Methods in Probability and Statistics: A Volume in Honour of Mikols Csorgo*, B. Szyszkowicz, ed., Elsevier Press, Netherlands, 135–150.

[11] S. Halfin and W. Whitt. 1981. Heavy-traffic limit theorems for queues with many exponential servers. *Oper. Res.*, 29:567–588.

[12] J.M. Harrison. 1985. *Brownian Motion and Stochastic Flow Systems.* Wiley, New York.

[13] J. M. Harrison and A. J. Lemoine. 1977. Limit theorems for periodic queues, *J. Appl. Prob.*, 14:566–576.

[14] J.M. Harrison and M.I. Reiman. 1981. Reflected Brownian motion on an orthant. *Ann. of Prob.*, 9:302–308.

[15] D. P. Heyman and W. Whitt. 1984. The asymptotic behaviour of queues with time-varying arrival rates. J. Appl. Prob., 21:143–156.

[16] 2003. P. Jelenkovic, A. Mandelbaum and P. Momcilovic. Heavy traffic limits for queues with many deterministic servers. To appear in *Queueing Syst.*.

[17] J. B. Keller. 1982. Time-dependent queues. *SIAM Rev.*, **24**, 401–412.

[18] N. Gans, G. Koole and A. Mandelbaum. 2003. Telephone call centers: Tutorial, review and research projects. Invited review paper by *Manufacturing and Service Operations Management*, **5**.

[19] A. J. Lemoine. 1989. Waiting time and workload in queues with periodic Poisson input. J. Appl. Prob., 26:390–397.

[20] W. P. Lovegrove, J. L. Hammond and D. Tipper. 1990. Simulation methods for studying nonstationary behavior of computer networks. IEEE Journal of Sel. Areas in Comm., 9:1696–1708.

[21] A. Mandelbaum. 1989. The dynamic complementarity problem. *Unpublished manuscript.*

[22] A. Mandelbaum and W. Massey. 1995. Strong approximations for time-dependent queues. *Math. of Oper. Res.*, 20:33–63.

[23] A. Mandelbaum, W. Massey and M. Reiman. 1998. Strong approximations for Markovian service networks. *Queueing Syst.* 30:149–201.

[24] A. Mandelbaum and G. Pats. 1998. State-dependent stochastic networks. PartI: Approximations and applications with continuous diffusion limits. *Ann. of App. Prob.*, 2:569–646.

[25] A. Mandelbaum and K. Ramanan. Functional central limit theorems for non-stationary queueing networks. *In preparation.*

[26] A. Mandelbaum and K. Ramanan. Directional derivatives of the QED map. *In preparation.*

[27] A. Mandelbaum and L. Van der Heyden. 1987. Complementarity and reflection. *Unpublished work.*

[28] W. Massey. 1981. Nonstationary queues. *Ph.D. Thesis*, Stanford University, CA.

[29] W. Massey. 1985. Asymptotic analysis of the time dependent M/M/1 Queue. *Math. of Oper. Res.*, 10:305–327.

[30] W. Massey and W. Whitt. 1993. Networks of infinite-server queues with nonstationary Poisson input. *Queueing Systems*, 13:183–250.

[31] G. F. Newell. 1968. Queues with time-dependent arrival rates I, II, III. *J. Appl. Prob.* 5:436–451 (I); 570–590 (II); 591–606 (III).

[32] G. F. Newell. 1982. *Applications of Queueing Theory.* Chapman and Hall, Second Edition.

[33] J. L. Pomerade. 1976. A unified approach via graphs to Skorokhod's topologies on the function space $\mathcal{D}$. *Ph.D. Thesis*, Dept. of Statistics, Yale University.

[34] W. P. Peterson. 1991. Diffusion approximations for networks of queues with multiple customer types. *Math. of Oper. Res.*, 9:90–118.

[35] A. Puhalskii and M. Reiman. 2000. The multiclass GI/PH/N queue in the Halfin-Whitt regime. *Adv. in App. Prob.* **32**, 564–595.

[36] K. Ramanan and M. Reiman. 2003. Fluid and heavy traffic diffusion limits for a generalized processor sharing model. *Ann. of App. Prob.*, **13** 1:100-139.

[37] M.I. Reiman. 1984. Open queueing networks in heavy traffic. *Math. of Oper. Res.*, 9:441–458.

[38] T. Rolski. 1981. Queues with non-stationary input stream: Ross' conjecture. *Adv. Appl. Prob.*, 13:608–618.

[39] T. Rolski. 1990. Queues with non-stationary inputs. *Queueing Systems*, 5:113–130.

[40] H. L. Royden. 1989. *Real Analysis, Third Edition.* Macmillan Publishing Co., New York.

[41] A. V. Skorokhod. 1956. Limit theorems for stochastic processes. *Theor. Prob. and its Appl.*, 1:261–290.

[42] A. V. Skorokhod. 1961. Stochastic equations for diffusions in a bounded region. *Theor. of Prob. and its Appl.*, 6:264–274.

[43] D. R. Smart. 1974. *Fixed Point Theorems.* Cambridge University Press, Great Britain.

[44] A. F. Veinott Jr. 1969. Discrete dynamic programming with sensitive discount optimality criteria. *Ann. Math. Statist.*, 40:1635–1660.

[45] W. Whitt. 2002. *An Introduction to Stochastic-Process Limits and Their Application to Queues.* Springer-Verlag, New York.

[46] W. Whitt. 2002. *An Introduction to Stochastic-Process Limits and Their Application to Queues.* Internet Supplement, http://www.columbia.edu/ ww2040/supplement.html.

[47] R. Williams. 1998. Reflecting diffusions and queueing networks. *Proceedings of the ICM.*

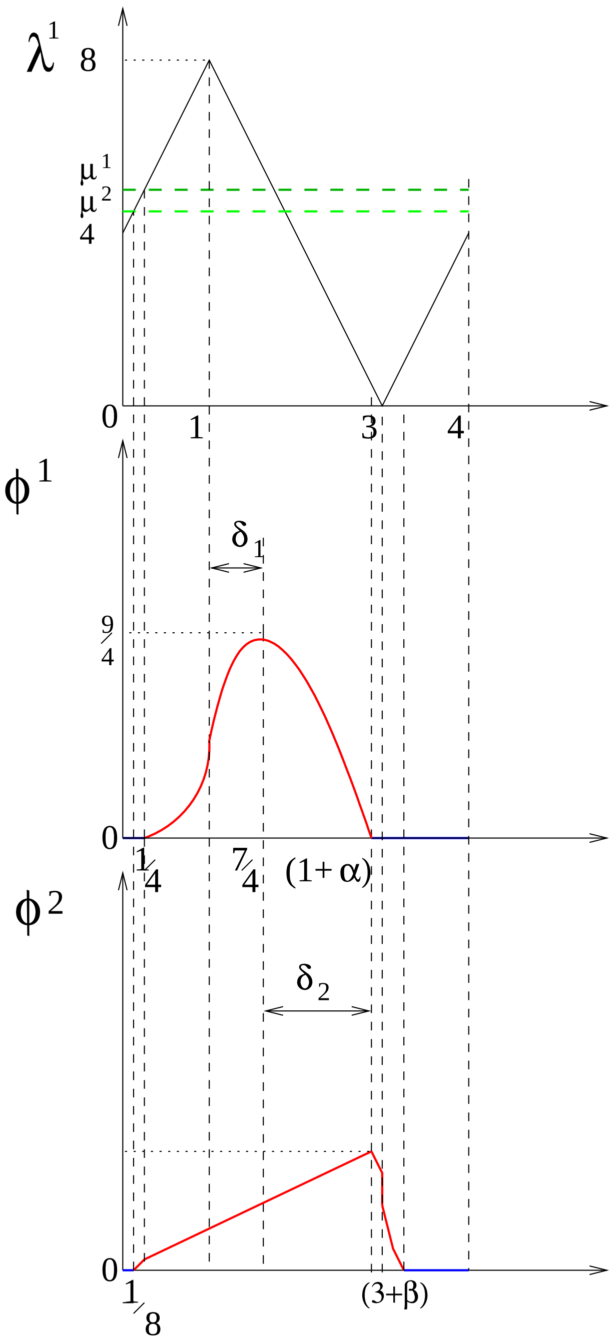

Figure 1: The time-varying arrival rate λ^1, and the contents ϕ^1 and ϕ^2 of the first and second queue in the tandem queueing network of Section 2.1 – note the lags δ_1 and $\delta_1 + \delta_2$ between the time of peak arrival rate and peak congestion at queues 1 and 2 respectively

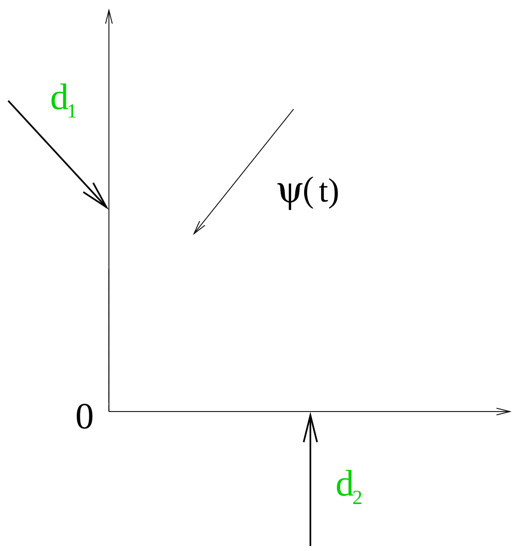

Figure 2: The oblique reflection map associated with a tandem queueing network

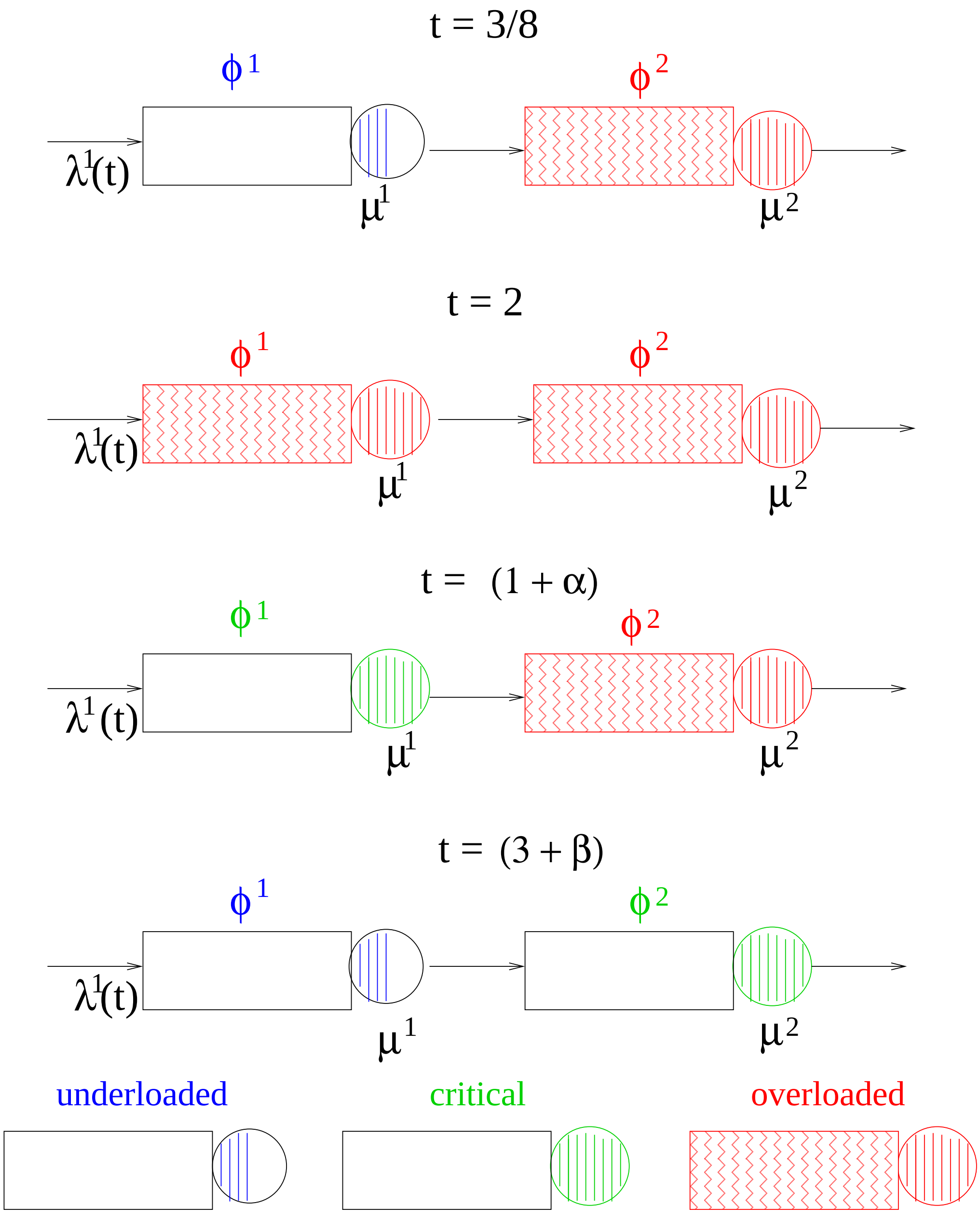

Figure 3: The tandem queueing network with time-varying arrival rates considered in Section 2.1 – each queue alternates between periods of underloading, criticality and overloading

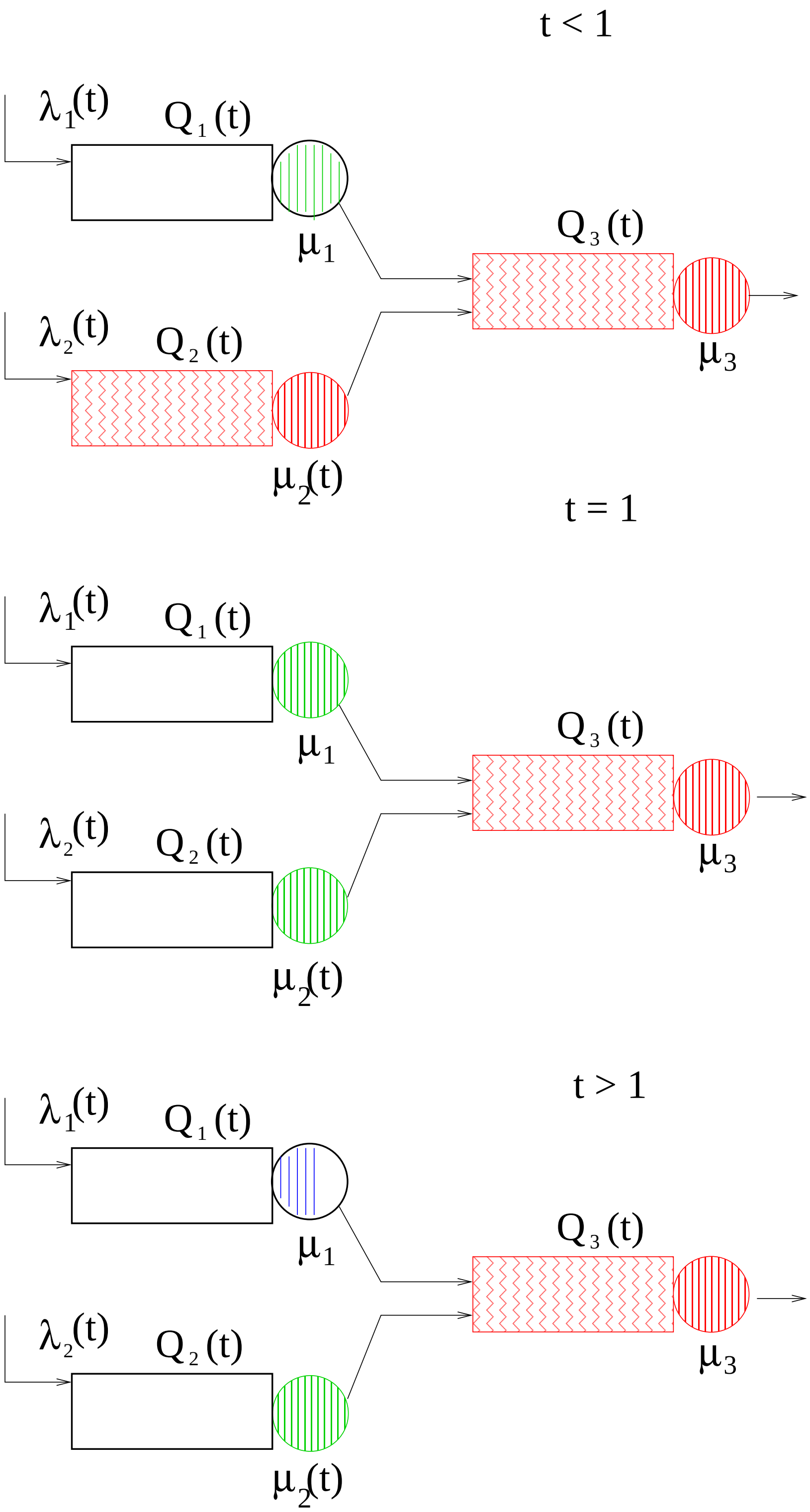

Figure 4: A queueing network with a merge and time-varying arrival and sevice rates giving rise to a separated discontinuity in the directional derivative at $t = 1$

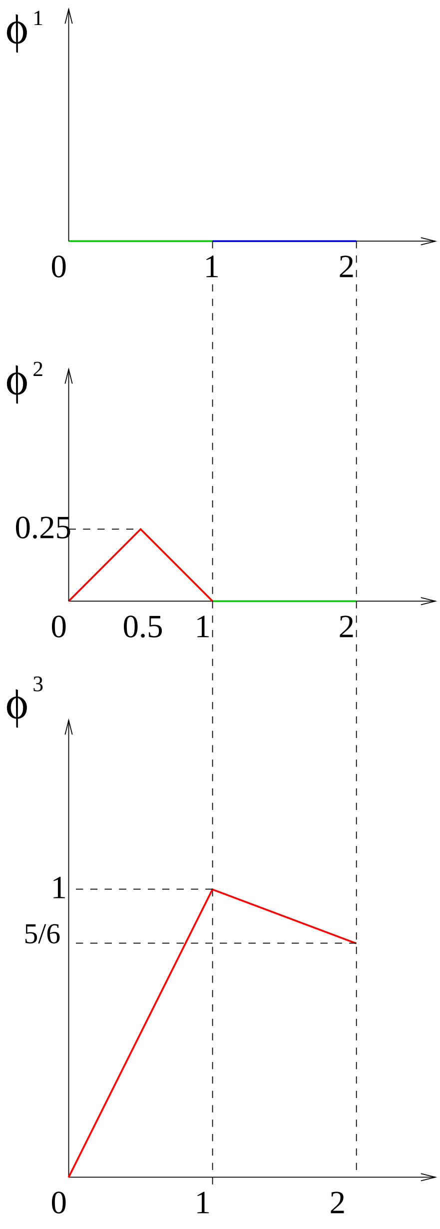

Figure 5: The fluid limit of the non-stationary merge queueing network