Chapter 14

Queueing Networks

14.1. Introduction

In this chapter we continue applying the continuous-mapping approach to establish heavy-traffic stochastic-process limits for queues, giving special attention to the possibility of having unmatched jumps in the limit process. Paralleling our application of the one-dimensional reflection map to obtain heavy-traffic limits for single queues in Chapters 5, 8, 9 and 10, we now apply the multidimensional reflection map to obtain heavy-traffic limits for queueing networks. As before, we omit some proofs. These proofs plus additional supporting material appear in Chapter 8 of the Internet Supplement.

For background on queueing (or stochastic) networks, see Kelly (1979), Whittle (1986) and Serfozo (1999). For related discussions of heavy-traffic limits for queueing networks, see Chen and Mandelbaum (1994a,b), Harrison (1988, 2000, 2001a,b), Chen and Yao (2001) and Kushner (2001). The literature is also discussed at the end of this chapter.

The (standard) multidimensional reflection map was used with the continuous mapping theorem by Harrison and Reiman (1981a,b) and Reiman (1984a) to establish heavy-traffic limits with reflected Brownian motion limit processes for vector-valued queue-length, waiting-time and workload stochastic processes in single-class open queueing networks. Since Brownian motion and reflected Brownian motion have continuous sample paths, the topology of uniform convergence over bounded intervals could be used for those results. Variants of the M_1 topologies are needed to obtain alternative stochastic-process limits with discontinuous sample paths, such as reflected Lévy processes, when the discontinuities in the sample paths of the limit process are approached gradually in the sample paths of the converging processes.

However, unlike the one-dimensional reflection map, studied in Section 13.5 and applied in Chapters 5, 8, 9 and 10, the multidimensional reflection map is not simply continuous, using either the SM_1 or WM_1 topology on D. Nevertheless, there is continuity with appropriate qualifications. In particular, the multidimensional reflection map is continuous at limits in the subset D_s of functions without simultaneous jumps of opposite sign in the coordinate functions, provided that the WM_1 topology is used on the range. As a consequence, the reflection map is continuous in the WM_1 topology at limits in the subset D_1 of functions that have discontinuities in only one coordinate at a time. That continuity property also holds for more general reflection maps and is sufficient to support limit theorems for stochastic processes in most applications.

We apply the continuity of the reflection map to obtain heavy-traffic limits for vector-valued buffer-content stochastic processes in single-class open stochastic fluid networks and for vector-valued queue-length stochastic processes in single-class open queueing networks. The M_1 topologies play a crucial role in these stochastic-process limits. In the stochastic fluid networks, just as in the single fluid queues in Chapters 5 and 8, the limit processes for the properly-scaled exogenous cumulative-input processes may have jumps even though the exogenous cumulative-input processes themselves have continuous sample paths. For example, this phenomenon occurs when the exogenous cumulative-input processes at the nodes are associated with the superpositions of independent on-off sources, where the busy (on or activity) periods and/or idle (off or inactivity) periods have heavy-tailed probability distributions (with infinite variances). Thus, in order to obtain heavy-traffic limits for properly scaled vector-valued buffer-content stochastic processes in the stochastic fluid networks, we need to invoke the continuous mapping theorem, using the continuity of the multidimensional reflection map on D with appropriate M_1 topologies.

Similarly, the M_1 topologies are needed to obtain heavy-traffic limits for open single-class queueing networks with single-server queues, where the servers are subject to rare long service interruptions. If the times between interruptions and the durations of the interruptions are allowed to increase appropriately as the traffic intensity increases in the heavy-traffic limit (with the durations of the interruptions being asymptotically negligible compared to the times between interruptions), then in the limit the interruptions occur instantaneously according to a stochastic point process, but nevertheless the interruptions have a spatial impact, causing jumps in the sample paths of the limiting queue-length process. Since these jumps in the limit process are approached gradually in the converging processes, again the M_1

topologies are needed to obtain heavy-traffic limits for the properly-scaled vector-valued queue-length stochastic processes.

Here is how this chapter is organized: We start in Section 14.2 by carefully defining the multidimensional reflection map and establishing its basic properties. Since the definition (Definition 14.2.1) is somewhat abstract, a key property is having the reflection map be well defined; i.e., we show that there exists a unique function satisfying the definition (Theorem 14.2.1). We also provide multiple characterizations of the reflection map, one alternative being as the unique fixed point of an appropriate operator (Theorem 14.2.2), while another is a basic complementarity property (Theorem 14.2.3).

A second key property of the multidimensional reflection map is Lipschitz continuity in the uniform norm on $D([0,T],\mathbb{R}^k)$ (Theorem 14.2.5). We also establish continuity of the multidimensional reflection map as a function of the reflection matrix, again in the uniform topology (Theorems 14.2.8 and 14.2.9). It is easy to see that the Lipschitz property is inherited when the metric on the domain and range is changed to d_{J_1} (Theorem 14.2.7). However, a corresonding direct extension for the SM_1 metric d_s does not hold. Much of the rest of the chapter is devoted to obtaining positive results for the M_1 topologies.

Section 14.3 provides yet another characterization of the multidimensional reflection map via an associated instantaneous reflection map on $\mathbb{R}^k$. The alternative characterization is as the limit of the multidimensional reflection map defined on the subset D_c of piecewise-constant functions in D (Theorem 14.3.4). On D_c, the reflection map can be defined in terms of the recursive application of the instantaneous reflection map on $\mathbb{R}^k$ (Theorem 14.3.1). The instantaneous reflection on $\mathbb{R}^k$ can be calculated by solving a linear program. Thus the instantaneous reflection map provides a useful way to simulate piecewise-constant approximations to reflected stochastic processes. We develop properties of the instantaneous reflection map that help us obtain the desired M_1 results (Theorem 14.3.2). In particular, we apply a monotonicity result (Corollary 14.3.2) to establish key properties of reflections of parametric representations (Lemma 14.3.4).

Sections 14.4 and 14.5 are devoted to obtaining the M_1 continuity results. In Section 14.4 we establish properties of reflection of parametric representations. We are able to extend Lipschitz and continuity results from the uniform norm to the M_1 metrics when we can show that the reflection of a parametric representation can serve as the parametric representation of the reflected function. The results are somewhat complicated, because this property holds only under certain conditions.

The basic M_1 Lipschitz and continuity results are established in Section

14.5. In addition to the positive results, we give counterexamples showing the necessity of the conditions in the theorems. Particularly interesting is Example 14.5.4, which shows that, without the regularity conditions, the fluctuations of the sequence of reflected processes can exceed the fluctuations in the reflection of the limit, thus exhibiting a kind of Gibbs phenomenon (see Remark 14.5.1). A proper limit in that example can be obtained, however, if we work in one of the more general spaces E or F introduced in Chapter 15.

In Sections 14.6 and 14.7, respectively, we apply the previous results to obtain heavy-traffic stochastic-process limits for stochastic fluid networks and conventional queueing networks. In the queueing networks we allow service interruptions. When there are heavy-tailed distributions or rare long service interruptions, the M_1 topologies play a critical role.

In Section 14.8 we consider the two-sided regulator and other reflection maps. The two-sided regulator is used to obtain heavy-traffic limits for single queues with finite waiting space, as considered in Section **??** and Chapter 8. With the scaling, the size of the waiting room is allowed to grow in the limit as the traffic intensity increases, but at a rate such that the limit process involves a two-sided regulator (reflection map) instead of the customary one-sided one. Like the one-sided reflection map, the two-sided regulator is continuous on (D^1, M_1). Moreover, the content portion of the two-sided regulator is Lipschitz, but the two regulator portions (corresponding to the two barriers) are only continuous; they are not Lipschitz.

We also give general conditions for other reflection maps to have M_1 continuity and Lipschitz properties. For these, we require that the limit function to be reflected belong to D_1, the subset of functions with discontinuities in only one coordinate at a time. We conclude the chapter with notes on the literature.

In Section 8.9 of the Internet Supplement we show that reflected stochastic processes have proper limiting stationary distributions and proper limiting stationary versions (stochastic-process limits for the entire time-shifted processes) under very general conditions. Our main result, Theorem 8.9.1 in the Internet Supplement, establishes such limits for stationary ergodic net-input stochastic processes satisfying a natural drift condition. It is noteworthy that a proper limit can exist even if there is positive drift in some (but not all) coordinates. Theorem 8.9.1 there is limited by having a special initial condition: starting out empty. Much of the rest of Section 8.9.1 in the Internet Supplement is devoted to obtaining corresponding results for other initial conditions. We establish convergence for all proper initial contents when the net input process is also a Lévy process with mutually independent coordinate processes (Theorem 8.9.6 there). Theorem 8.9.6 covers limit processes obtained in the heavy-traffic limits for the stochastic fluid networks in Section 14.6.

14.2. The Multidimensional Reflection Map

The definition of the multidimensional reflection map is somewhat indirect. So we begin by motivating the definition. Since the multidimensional reflection map arises naturally in the definition of the vector-valued buffer-content stochastic process in a stochastic fluid network, we first define the multidimensional reflection map in that special context.

14.2.1. A Special Case

We consider a single-class open stochastic fluid network with k nodes, each with a buffer of unlimited capacity. We let exogenous fluid input come to each of the k nodes. At each node the fluid is processed and released at a deterministic rate. The processed fluid from each node is then routed in a Markovian manner to other nodes or out of the network. The principal stochastic process of interest is the k-dimensional buffer-content process. The stochastic fluid network is a generalization of the fluid queue models in Chapters 5 and 8 in the case of unlimited waiting space (buffer capacity).

A single-class open stochastic fluid network with Markovian routing can be specified by a four-tuple $\{C, r, P, X(0)\}$, where $C \equiv (C^1, \ldots, C^k)$ is the vector of exogenous input stochastic processes at the k nodes, $r \equiv (r^1, \ldots, r^k)$ is the vector of deterministic output rates at the k nodes, $P \equiv (P_{i,j})$ is the $k \times k$ routing matrix and $X(0) \equiv (X^1(0), \ldots, X^k(0))$ is the nonnegative random vector of initial buffer contents at the k nodes. The stochastic process C is an element of $D_\uparrow^k \equiv D_\uparrow^1 \times \cdots \times D_\uparrow^1$, the subset of functions in $D \equiv D^k \equiv D([0,T], \mathbb{R}^k)$ that are nondecreasing and nonnegative in each coordinate. The random variable $C^i(t)$ represents the cumulative exogenous input to node i during the time interval $[0,t]$. It is natural to let the sample paths of C^i be continuous, but we do not require it.

When the buffer at node i is nonempty, there is fluid output from node i at constant rate r_i. When buffer i is empty, the output rate equals the minimum of the combined external (exogenous) plus internal input rate and the potential output rate r_i (formalized below). A proportion $P_{i,j}$ of all output from node i is immediately routed to node j, while a proportion $p_i \equiv 1 - \sum_{j=1}^k P_{ij}$ is routed out of the network. We assume that the routing matrix P is substochastic, so that $P_{i,j} \geq 0$ and $p_i \geq 0$ for all i, j. We also assume that $P^n \to 0$ as $n \to \infty$, where P^n is the n^{th} power of P, so that all input eventually leaves the network. (We can think of P as the transition matrix of a transient k-state Markov chain.)

We now proceed to mathematically define the vector-valued buffer content process. We will work with column vectors. Hence it is convenient to use the transpose of the routing matrix P. Let Q be this transpose, i.e., $Q \equiv P^t$, so that Q is a column-substochastic matrix. We start by defining a potential buffer-content (or net-input) process, which represents the potential content at each node, ignoring the emptiness condition. The *potential buffer-content process* is

$$X(t) \equiv X(0) + C(t) - (I - Q)rt, \quad t \geq 0 \ , \tag{2.1}$$

where the vectors are regarded as column vectors and the transpose Q^t is the routing matrix P. The component $X^i(t)$ in (2.1) represents what the content of buffer i would be at time t if the output occurred continuously at rate r^j from node j for all j, regardless whether or not station j had fluid to emit. That is, the content vector $X(t)$ at time t, would be the initial value $X(0)$ plus the exogenous input $C(t)$ minus the output rt plus the internal input Qrt.

We obtain the actual buffer content by disallowing the potential output (and associated internal input) that cannot occur because of emptiness. We use the componentwise partial order on D and $\mathbb{R}^k$; i.e., $c_1 \equiv (c_1^1, \ldots, c_1^k) \leq c_2 \equiv (c_2^1, \ldots, c_2^k)$ in $\mathbb{R}^k$ if $c_2^i \leq c_2^i$ in $\mathbb{R}$ for all i, $1 \leq i \leq k$, and $x_1 \leq x_2$ in D if $x_1(t) \leq x_2(t)$ in $\mathbb{R}^k$ for all t, $0 \leq t \leq T$. We then let the buffer content be

$$Z(t) \equiv X(t) + (I - Q)Y(t) \ , \tag{2.2}$$

where X is defined in (2.1) and Y is the least possible element of $D_\uparrow^k$ (the subset of functions in D^k that are nonnegative and nondecreasing in each coordinate) such that $Z \geq 0$. We call the map from X to (Y, Z) the *reflection map.*

14.2.2. Definition and Characterization

We will prove that the vector-valued buffer-content process Z is properly defined by (2.1) and (2.2) by showing that such a process Y is well defined (exists and is unique). It turns out that the reflection map from X to Y and Z is well defined, even if X does not have the special structure in (2.1).

More generally, we call the transposed routing matrix Q the *reflection matrix.* Let $\mathcal{Q}$ be the set of all reflection matrices, i.e., the set of all column-stochastic matrices Q (with $Q^t_{i,j} \geq 0$ and $\sum_{j=1}^k Q^t_{i,j} \leq 1$) such that $Q^n \to 0$ as $n \to \infty$, where Q^n is the n^{th} power of Q.

Definition 14.2.1. (reflection map) *For any* $x \in D^k \equiv D([0,T], \mathbb{R}^k)$ *and any reflection matrix* $Q \in \mathcal{Q}$, *let the* feasible regulator *set be*

$$\Psi(x) \equiv \{w \in D^k_\uparrow : x + (I - Q)w \geq 0\} \tag{2.3}$$

and let the reflection map *be* $R \equiv (\psi, \phi) : D^k \to D^{2k}$ *with* regulator component

$$y \equiv \psi(x) \equiv \inf \Psi(x) \equiv \inf\{w : w \in \Psi(x)\} , \tag{2.4}$$

i.e.,

$$y^i(t) \equiv \inf\{w^i(t) \in \mathbb{R} : w \in \Psi(x)\} \quad \textit{for all} \quad i \quad \textit{and} \quad t , \tag{2.5}$$

and content component

$$z \equiv \phi(x) \equiv x + (I - Q)y . \tag{2.6}$$

It remains to show that the reflection map is well defined by Definition 14.2.1; i.e., we need to know that the feasible regulator set $\Psi(x)$ is nonempty and that its infimum y (which necessarily is well defined and unique for nonempty $\Psi(x)$) is itself an element of $\Psi(x)$, so that $z \in D^k$ and $z \geq 0$.

To show that $\Psi(x)$ in (2.3) is nonempty, we exploit the well known fact that the matrix $I - Q$ has nonnegative inverse.

Lemma 14.2.1. (nonnegative inverse of reflection matrix) *For all* $Q \in \mathcal{Q}$, $I - Q$ *is nonsingular with nonnegative inverse*

$$(I - Q)^{-1} = \sum_{n=0}^{\infty} Q^n ,$$

where $Q^0 = I$.

The key to showing that the infimum belongs to the feasibility set is a basic result about semicontinuous functions. Recall that a real-valued function x on $[0,T]$ is *upper semicontinuous* at a point t in its domain if

$$\limsup_{t_n \to t} x(t_n) \leq x(t)$$

for any sequence $\{t_n\}$ with $t_n \in [0,T]$ and $t_n \to t$ as $n \to \infty$. The function x is upper semicontinuous if it is upper semicontinuous at all arguments t in its domain.

Lemma 14.2.2. (preservation of upper semicontinuity) *Suppose that* $\{x_s : s \in S\}$ *is a set of upper semicontinuous real-valued function on a subinterval of* $\mathbb{R}$. *Then the infimum* $\underline{x} \equiv \inf\{x_s : s \in S\}$ *is also upper semicontinuous.*

Recall that $x^\uparrow \equiv \sup_{0 \le s \le t} x(s)$, $t \ge 0$, for $x \in D^1$. For $x \equiv (x^1, \ldots, x^k) \in D^k$, let $x^\uparrow \equiv ((x^1)^\uparrow, \ldots, (x^k)^\uparrow)$.

Theorem 14.2.1. (existence of the reflection map) *For any* $x \in D^k$ *and* $Q \in \mathcal{Q}$,

$$(I - Q)^{-1}[(-x)^\uparrow \vee 0] \in \Psi(x) \ , \tag{2.7}$$

so that $\Psi(x) \ne \phi$,

$$y \equiv \psi(x) \in \Psi(x) \subseteq D^k_\uparrow \tag{2.8}$$

for y *in (2.4) and*

$$z \equiv \phi(x) = x + (I - Q)y \ge 0 \ . \tag{2.9}$$

Proof. Using Lemma 14.2.1, it is easy to see that (2.7) holds, because

$$x + (I - Q)(I - Q)^{-1}[(-x)^\uparrow \vee 0] \ge x + [(-x)^\uparrow \vee 0] \ge x + (-x)^\uparrow \ge x - x \ge 0$$

and $(-x)^\uparrow \vee 0 \in D^k_\uparrow$. Let y be the infimum in (2.4). Since $\Psi(x) \subseteq D^k_\uparrow$, necessarily y is nondecreasing and nonnegative. Since the elements of $D^k_\uparrow$ are nondecreasing, right-continuity coincides with upper semicontinuity, so we can apply Lemma 14.2.2 to conclude that $y \in D^k_\uparrow$. It now remains to show that $x + (I - Q)y \ge 0$. Fix ϵ, i and t. By the definition of the infimum, there exists $w \in \Psi(x)$ such that $w^i(t) \le y^i(t) + \epsilon$ and $w^j(t) \ge y^j(t)$ for all j. Thus

$$x^i(t) + y^i(t) - \sum_{j=1}^{k} Q_{j,i} y^j(t) \ge x^i(t) + w^i(t) - \sum_{j=1}^{k} Q_{j,i} w^j(t) - \epsilon \, .$$

Since ϵ, i and t were arbitrary,

$$x + (I - Q)y \ge x + (I - Q)w \ge 0$$

so (2.8) holds and the proof is complete. ∎

We now characterize the regulator function $y \equiv \psi(x)$ as the unique fixed point of a mapping $\pi \equiv \pi_{x,Q} : D^k_\uparrow \to D^k_\uparrow$, defined by

$$\pi(w) = (Qw - x)^\uparrow \vee 0 \tag{2.10}$$

for $w \in D^k_\uparrow$. For this purpose, we use two elementary lemmas.

Lemma 14.2.3. (feasible regulator set characterization) *The feasible regulator set* $\Psi(x)$ *in* (2.3) *can be characterized by*

$$\Psi(x) = \{w \in D_{\uparrow}^{k} : w \geq \pi(w)\}$$

for π *in* (2.10).

Proof. For each $w \in \Psi(x)$, $x + (I - Q)w \geq 0$ or, equivalently, $w \geq Qw - x$. Since $\Psi(x) \subseteq D_{\uparrow}^{k}$, we must also have $w \geq \pi(y)$. On the other hand, if $w \in D_{\uparrow}^{k}$ and $w \geq \pi(y)$, then we must have $w \geq Qw - x$, so that $w \in \Psi(x)$. ∎

Remark 14.2.1. *Semilattice structure.* It is also easy to see that if $y_1, y_2 \in \Psi(x)$, then $y_1 \wedge y_2 \in \Psi(x)$; that makes $\Psi(x)$ a meet semilattice.

Lemma 14.2.4. (closed subset of D) *With the uniform topology on* D*, The feasible regulator set* $\Psi(x)$ *is a closed subset of* $D_{\uparrow}^{k}$*, while* $D_{\uparrow}^{k}$ *is a closed subset of* D.

Theorem 14.2.2. (fixed-point characterization) *For each* $Q \in \mathcal{Q}$*, the regulator map* $y \equiv \psi(x) \equiv \psi_Q(x) : D^k \to D_{\uparrow}^{k}$ *can be characterized as the unique fixed point of the map* $\pi \equiv \pi_{x,Q} : D_{\uparrow}^{k} \to D_{\uparrow}^{k}$ *defined in* (2.10).

Proof. We use a standard argument to establish fixed points of monotone maps on ordered sets; e.g., see Section 3.8 of Edwards (1965). As before, we use the componentwise partial order on $\mathbb{R}^k$ and D^k. It is immediate that the map π in (2.10) is monotone. Note that

$$0 \leq \pi(0) = (-x)^{\uparrow} \vee 0 \ ,$$

where 0 is used as the vector and vector-valued function with zero values. Note that the functions 0 and $\pi(0)$ are both elements of $D_{\uparrow}^{k}$. Hence, the iterates $\pi^n(0) \equiv \pi(\pi^{n-1}(0))$ are elements of $D_{\uparrow}^{k}$ that are nondecreasing in n. On the other hand, $w \geq \pi(w)$ for any $w \in \Psi(x)$ by Lemma 14.2.3. Consequently, $\pi^n(w)$ is decreasing in n for any $w \in \Psi(x)$. Since $\Psi(x)$ is nonempty by Theorem 14.2.1, there exists $w \in \Psi(x)$. Since $0 \leq w$, $\pi^n(0) \leq \pi^n(w) \leq w$ for all n, so that $\pi^n(0)$ is bounded above. It is easy to see, using the addition and supremum maps, that $\pi : (D, U) \to (D, U)$ is also continuous. Hence $\pi^n(0) \uparrow w^*$ in $D_{\uparrow}^{k}$ as $n \to \infty$, where $\pi(w^*) = w^*$. Since $D_{\uparrow}^{k}$ is a closed subset of D by Lemma 14.2.4, $w^* \in D_{\uparrow}^{k}$. Since $w^* \geq \pi(w^*)$,

$w^* \in \Psi(x)$ too. Since the regulator $y \equiv \psi(x)$ is the infimum by Definition 14.2.1, necessarily $y \leq w^*$. Since $0 \leq y \leq w^*$,

$$\pi^n(0) \leq \pi^n(y) \leq y \quad \text{for all} \quad n \ .$$

Letting $n \to \infty$, we see that $w^* \leq y$. Hence we must have $y = w^*$. ∎

Theorem 14.2.3. (complementarity characterization) *A function y in the feasible regulator set $\Psi(x)$ in* (2.3) *is the infimum $\psi(x)$ in* (2.4) *if and only if the pair (y, z) for $z \equiv x + (I - Q)y$ satisfies the complementarity property*

$$\int_0^\infty z^i dy^i = 0, \quad 1 \leq i \leq k \ . \tag{2.11}$$

Proof. We first prove that the infimum satisfies the complementarity property. We will show that failing to satisfy the complementarity implies failing the infimum property. Hence, suppose that (y, z) fails to satisfy the complementarity property (2.11). Thus there is t and j such that $z^j(t) > 0$ and y^j increases at t. We consider two cases:

Case 1. Suppose that $y^j(t) > y^j(t-)$. There must exist $\epsilon, \delta > 0$ such that $y^j(t) - y^j(t-) > \epsilon$ and $z^j(s) \geq \epsilon$, $t \leq s \leq t + \delta$. Let $\tilde{y}$ be defined by

$$\tilde{y}^j(s) = \begin{cases} y^j(s), & 0 \leq s < t \ , \\ y^j(s) - \epsilon, & t \leq s < t + \delta \ , \\ y^j(s), & t + \delta \leq s \ , \end{cases}$$

with $\tilde{y}^i = y^i$ for $i \neq j$. Then $\tilde{y} \in D_\uparrow^k$ and $x + (I - Q)\tilde{y} \geq 0$, so that $\tilde{y} \in \Psi(x)$, $\tilde{y} \leq y$ and $\tilde{y} \neq y$, which implies that y is not the infimum.

Case 2. Suppose that $y^j(t) = y^j(t-)$. Now there must exist $\epsilon, \delta > 0$ such that $z^j(s) \geq \epsilon$ and $0 \leq y^j(s) - y^j(t) \leq \epsilon$ for $t \leq s < t + \delta$ and $y^j(s) - y^j(t) > 0$ for $s > t$. Now let $\tilde{y}$ be defined by

$$\tilde{y}^j(s) = \begin{cases} y^j(s), & 0 \leq s < t \ , \\ y^j(t), & t \leq s < t + \delta \ , \\ y^j(s), & t + \delta \leq s \ , \end{cases}$$

with $\tilde{y}^i = y^i$ for $i \neq j$. Again $\tilde{y} \in D_\uparrow^k$ and $x + (I - Q)\tilde{y} \geq 0$. Since $\tilde{y} \leq y$ and $\tilde{y} \neq y$, y must not be the infimum. We now prove that the complementarity property implies the infimum property. Invoking Theorem 14.2.2, it suffices to show that if (y, z) satisfies the complementarity property, then necessarily y is the unique solution to the fixed point equation $y = \pi(y)$. Thus let $v = \pi(y)$. Since $y \in \Psi(x)$, $y \geq \pi(y)$ by Lemma 14.2.3, so that it suffices to show that $v = \pi(y) \geq y$. Suppose that $y(t) > v(t)$ for some t. Then necessarily $y(t_0) > v(t_0)$ for some t_0 that is a point of increase of y, but $y(t_0) > \pi(y)(t_0)$ implies that

$$z(t_0) = y(t_0) - (Qy - x)(t_0) \geq y(t_0) - \pi(y)(t_0) > 0 ,$$

which contradicts the complementarity condition. ∎

14.2.3. Continuity and Lipschitz Properties

We now establish continuity and Lipschitz properties of the reflection map as a function of the function x and the reflection matrix Q. We use the *matrix norm*, defined for any $k \times k$ real matrix A by

$$\|A\| \equiv \max_j \sum_{i=1}^{k} |A_{i,j}| . \tag{2.12}$$

We use the maximum column sum in (2.12) because we intend to work with the column-substochastic matrices in $\mathcal{Q}$. Note that

$$\|A_1 A_2\| \leq \|A_1\| \cdot \|A_2\|$$

for any two $k \times k$ real matrices A_1 and A_2. Also, using the sum (or l_1) norm

$$\|u\| \equiv \sum_{i=1}^{k} |u^i| \tag{2.13}$$

on $\mathbb{R}^k$, we have

$$\|Au\| \leq \|A\| \cdot \|u\| \tag{2.14}$$

for each $k \times k$ real matrix A and $u \in \mathbb{R}^k$. Indeed, we can also define the matrix norm by

$$\|A\| \equiv \max\{\|Au\| : u \in \mathbb{R}^n, \|u\| = 1\} , \tag{2.15}$$

using the sum norm in (2.13) in both places on the right. Then (2.12) becomes a consequence. Consistent with (2.13), we let

$$\|x\| \equiv \sup_{0 \leq t \leq T} \|x(t)\| \equiv \sup_{0 \leq t \leq T} \sum_{i=1}^{k} \|x^i(t)\| \tag{2.16}$$

for $x \in D([0,T],\mathbb{R}^k)$. Combining (2.14) and (2.16), we have

$$\|Ax\| \le \|A\| \cdot \|x\| \tag{2.17}$$

for each $k \times k$ real matrix A and $x \in D([0,T],\mathbb{R}^k)$.

We use the following basic lemma.

Lemma 14.2.5. (reflection matrix norms) *For any $k \times k$ matrix $Q \in \mathcal{Q}$,*

$$\|Q\| \le 1, \quad \|Q^k\| = \gamma < 1 \tag{2.18}$$

and

$$\|(I-Q)^{-1}\| \le \frac{k}{1-\gamma} \ . \tag{2.19}$$

Example 14.2.1. *Need for k-stage contraction.* The standard example in which $\|Q^j\| = 1$ for all j, $1 \le j \le k-1$, has $Q^t_{i,i+1} = 1$ for $1 \le i \le k-1$, $Q^t_{k,1} = \gamma$, $0 < \gamma < 1$, and $Q^t_{k,i} = 0$, $2 \le i \le k$. Then $Q^k = \gamma I$ and $\|Q^k\| = \gamma$. ■

We now show that $\pi \equiv \pi_{x,Q}$ in (2.10) is a k-stage contraction map on $D^k_\uparrow$. Recall that for $x \in D$, $|x|$ denotes the function $\{|x(t)| : t \ge 0\}$ in D, where $|x(t)| = (|x^1(t)|, \dots, |x^k(t)|) \in \mathbb{R}^k$. Thus, for $x \in D$, $|x|^\uparrow = (|x^1|^\uparrow, \dots, |x^k|^\uparrow)$, where $|x^i|^\uparrow(t) = \sup_{0 \le s \le t} |x^i(s)|$, $0 \le t \le T$.

Lemma 14.2.6. (π is a k-stage contraction) *For any $Q \in \mathcal{Q}$ and $w_1, w_2 \in D^k_\uparrow$,*

$$|\pi^n(w_1) - \pi^n(w_2)|^\uparrow \le |Q^n(|w_1 - w_2|^\uparrow)| \quad for \quad n \ge 1 \ , \tag{2.20}$$

so that

$$\|\pi^n(w_1) - \pi^n(w_2)\| \le \|Q^n\| \cdot \|w_1 - w_2\| \le \|w_1 - w_2\| \tag{2.21}$$

for $n \ge 1$ and

$$\|\pi^n(w_1) - \pi^n(w_2)\| \le \gamma \|w_1 - w_2\| \quad for \quad n \ge k \ ,$$

where

$$\|Q^k\| \equiv \gamma < 1 \ .$$

Hence

$$\|\pi^n(w) - \psi(x)\| \to 0 \quad as \quad n \to \infty \ .$$

Proof. First,

$$\begin{aligned} |\pi(w_1) - \pi(w_2)|^{\uparrow} &= |(Qw_1 - x)^{\uparrow} \vee 0 - (Qw_2 - x)^{\uparrow} \vee 0|^{\uparrow} \\ &\leq |(Qw_1 - x)^{\uparrow} - (Qw_2 - x)^{\uparrow}|^{\uparrow} \\ &\leq |Qw_1 - x - Qw_2 + x|^{\uparrow} \\ &\leq Q|w_1 - w_2|^{\uparrow} , \end{aligned}$$

which implies (2.20) for $n = 1$. We prove (2.20) for arbitrary n by induction. Suppose that it has been established up to n. Then

$$\begin{aligned} |\pi^{n+1}(w_1) - \pi^{n+1}(w_2)|^{\uparrow} &= |(Q\pi^n(w_1) - x)^{\uparrow} \vee 0 - (Q\pi(w_2) - x)^{\uparrow} \vee 0|^{\uparrow} \\ &\leq Q|\pi^n(w_1) - \pi^n(w_2)|^{\uparrow} \leq Q^{n+1}|w_1 - w_2|^{\uparrow} , \end{aligned}$$

using the induction hypothesis. Finally,

$$\|x\| = \sum_{i=1}^{k} |x^i|^{\uparrow}(T) ,$$

so that (2.21) follows directly from (2.20). By the Banach-Picard contraction fixed-point theorem, for any $w \in D_{\uparrow}^k$,

$$\|\pi^n(w) - w^*\| \to 0 \quad \text{as} \quad n \to \infty$$

geometrically fast, where w^* is the unique fixed point of π, but by Theorem 14.2.2 that fixed point is $\psi(x)$. ■

We now establish inequalities that imply that the reflection map is a Lipschitz continuous map on $(D, \|\cdot\|)$. We will use the stronger inequalities themselves in Section **??**.

Theorem 14.2.4. (one-sided bounds) *For any $Q \in \mathcal{Q}$ and $x_1, x_2 \in D$,*

$$-(I-Q)^{-1}\eta_1(x_1 - x_2) \leq \psi(x_1) - \psi(x_2) \leq (I-Q)^{-1}\eta_1(x_2 - x_1) \quad (2.22)$$

where $\eta_1(x) \equiv (\hat{\eta}_1(x^1), \ldots, \hat{\eta}_1(x^k))$ with $\hat{\eta}_1 : D^1 \to D^1$ defined by

$$\hat{\eta}_1(x^i) \equiv (x^i)^{\uparrow} \vee 0 .$$

Proof. Use the map π_x in (2.10). Assuming that

$$\psi(x_1) \leq \pi_{x_2}^n(\psi(x_1)) + \sum_{i=0}^{n-1} Q^i \eta_1(x_2 - x_1) ,$$

with $\pi^0_{x_2}$ being the identity (which is true for $n = 0$), we have

$$\begin{aligned}\psi(x_1) &= \pi_{x_1}(\psi(x_1)) = \eta_1(Q\psi(x_1) - x_1) \\ &\leq \eta_1(Q\pi^n_{x_2}(\psi(x_1)) - x_2 + Q\sum_{i=0}^{n-1} Q^i \eta_1(x_2 - x_1) + (x_2 - x_1)) \\ &\leq \eta_1(Q\pi^n_{x_2}(\psi(x_1)) - x_2 + \sum_{i=0}^{n} Q^i \eta_1(x_2 - x_1)) \\ &\leq \eta_1(Q\pi^n_{x_2}(\psi(x_1) - x_2) + \sum_{i=0}^{n} Q^i \eta_1(x_2 - x_1) \\ &\leq \pi^{n+1}_{x_2}(\psi(x_1)) + \sum_{i=0}^{n} Q^i \eta_1(x_2 - x_1) \ .\end{aligned}$$

Letting $n \to \infty$, we obtain the second inequality in (2.22). The first inequality in (2.22) follows by symmetry. ∎

We also have the following consequence.

Corollary 14.2.1. (more bounds) *For any $Q \in \mathcal{Q}$, $x \in D^k$ and $w \in \mathbb{R}^k_+$,*

$$0 \leq \psi(x + w) \leq \psi(x) \leq \psi(x + w) + (I - Q)^{-1} w \ .$$

Proof. Apply Theorem 14.2.4, letting $x_1 \equiv x$ and $x_2 = x + w$. ∎

As a direct consequence of Theorem 14.2.4, we obtain the desired Lipschitz property.

Theorem 14.2.5. (Lipschitz property with uniform norm) *For any $Q \in \mathcal{Q}$ and $x_1, x_2 \in D$,*

$$\begin{aligned}\|\psi(x_1) - \psi(x_2)\| &\leq \|(I - Q)^{-1}\| \cdot \|x_1 - x_2\| \\ &\leq \sum_{n=0}^{\infty} \|Q^n\| \cdot \|x_1 - x_2\| \\ &\leq \frac{k}{1-\gamma} \|x_1 - x_2\| \ , \end{aligned} \tag{2.23}$$

where $\gamma \equiv \|Q^k\| < 1$, and

$$\begin{aligned}\|\phi(x_1) - \phi(x_2)\| &\leq (1 + \|I - Q\| \cdot \|(I - Q)^{-1}\|)\|x_1 - x_2\| \\ &\leq \left(1 + \frac{2k}{1-\gamma}\right) \|x_1 - x_2\| \ . \end{aligned} \tag{2.24}$$

Proof. Since

$$\left\|\sum_{n=0}^{\infty} Q^n\right\| \le \sum_{n=0}^{\infty} \|Q^n\| \le \sum_{n=0}^{k-1} \|Q^n\| + \gamma \sum_{n=0}^{\infty} \|Q^n\| ,$$

we have

$$\sum_{n=0}^{\infty} \|Q^n\| \le \frac{\sum_{n=0}^{k-1} \|Q^n\|}{1-\gamma} \le \frac{k}{1-\gamma} .$$

By (2.6) and (2.23),

$$\begin{aligned} \|\phi(x_1) - \phi(x_2)\| &\le \|x_1 - x_2\| + \|I - Q\| \cdot \|\psi(x_1) - \psi(x_2)\| \\ &\le (1 + \|I - Q\| \cdot \|(I - Q)^{-1}\|)\|x_1 - x_2\| . \quad \blacksquare \end{aligned}$$

Remark 14.2.2. *Alternate proof.* Instead of applying Theorem 14.2.4, we could prove Theorem 14.2.5 by reasoning as in Lemma 14.2.6 to get

$$|\pi_{x_1}^n(0) - \pi_{x_2}^n(0)|^{\uparrow} \le (I + Q + \cdots + Q^{n-1})|x_1 - x_2|^{\uparrow} ,$$

which implies that

$$|\psi(x_1) - \psi(x_2)|^{\uparrow} \le (I - Q)^{-1}|x_1 - x_2|^{\uparrow}$$

and then (2.23). ■

Remark 14.2.3. *Lipschitz constant.* The upper bounds in Theorem 14.2.5 are minimized by making $Q_{i,j} = 0$ for all i, j. Let K^* be the infimum of K such that

$$\|R(x_1) - R(x_2)\| \le K\|x_1 - x_2\| \quad \text{for all} \quad x_1, x_2 \in D . \tag{2.25}$$

We call K^* the *Lipschitz constant.* The bounds yield $K^* \le 2$ when $Q_{i,j} = 0$ for all i, j, but the following example in shows that $K^* = 2$ in that case. Hence $K^* \ge 2$ in general.

Example 14.2.2. *Lower bound on the Lipschitz constant.* Let $k = 1$, $Q = 0$, $x_1(t) = 0$, $0 \le t \le 1$, and $x_2 = -I_{[1/3,1/2)} + I_{[1/2,1]}$ in $D([0,1], \mathbb{R})$. Then $y_1 = z_1 = x_1$, but $y_2 = I_{[1/3,1]}$ and $z_2 = 2I_{[1/2,1]}$, so that $\|z_1 - z_2\| = 2$. Hence $K^* \ge 2$ for all Q.

Example 14.2.3. *No upper bound on the Lipschitz constant.* To see that there is no upper bound on the Lipschitz constant K^* independent of Q, let $x_1(t) = 0$, $0 \le t \le 2$, and $x_2 = -I_{[1,2]}$ in $D([0,2],\mathbb{R})$, so that $\|x_1 - x_2\| = 1$. Let $Q = 1 - \epsilon$, so that (2.6) becomes $z = x + \epsilon y$. Then $z_2 = z_1 = y_1 = x_1$, but $y_2 = \epsilon^{-1} I_{[1,2]}$, so that $\|y_1 - y_2\| = \epsilon^{-1}$.

Example 14.2.4. *No upper bound on the Lipschitz constant for ϕ.* To see that the Lipschitz constant for the component map ϕ can be arbitrarily large as well, consider the two-dimensional example with $Q_{1,1} = 1 - \epsilon$, $Q_{2,1} = 1$ and $Q_{2,2} = Q_{1,2} = 1/2$, so that

$$(I - Q^t) = \begin{pmatrix} \epsilon & -1 \\ -1/2 & 1/2 \end{pmatrix} .$$

Let $x_1^1 = -I_{[1,2]}$, $x_2^1(t) = 0$, $0 \le t \le 2$, and $x_1^2 = x_2^2 = \epsilon^{-1} I_{[0,2]}$ in $D([0,2],\mathbb{R}^2)$. Then $\|x_1 - x_2\| = 1$, but $z_1^1(t) = z_2^1(t) = 0$, $0 \le t \le 2$, $z_1^2 = \epsilon^{-1} I_{[0,1)}$ and $z_2^2 = \epsilon^{-1} I_{[0,2]}$, so that $\|z_1 - z_2\| = \epsilon^{-1}$.

We now summarize some elementary but important properties of the reflection map.

Theorem 14.2.6. (reflection map properties) *The reflection map satisfies the following properties:*

(i) adaptedness*: For any $x \in D$ and $t \in [0,T]$, $R(x)(t)$ depends upon x only via $\{x(s) : 0 \le s \le t\}$.*

(ii) monotonicity*: If $x_1 \le x_2$ in D, then $\psi(x_1) \ge \psi(x_2)$.*

(iii) rescaling*: For each $x \in D([0,T],\mathbb{R}^k)$, $\eta \in \mathbb{R}^k$, $\beta > 0$ and γ nondecreasing right-continuous function mapping $[0,T_1]$ into $[0,T]$, $\eta + \beta(x \circ \gamma) \in D([0,T_1],\mathbb{R}^k)$ and*

$$R(\eta + \beta(x \circ \gamma)) = \beta R(\beta^{-1}\eta + x) \circ \gamma .$$

(iv) shift*: For all $x \in D$ and $0 < t_1 < t_2 < T$,*

$$\psi(x)(t_2) = \psi(x)(t_1) + \psi(\phi(x)(t_1) + x(t_1 + \cdot) - x(t_1))(t_2 - t_1)$$

and

$$\phi(x)(t_2) = \phi(\phi(x)(t_1) + x(t_1 + \cdot) - x(t_1))(t_2 - t_1)$$

(v) continuity preservation*: If $x \in C$, then $R(x) \in C$.*

We can apply Theorems 14.2.5 and 14.2.6 (iii) to deduce that the reflection map inherits the Lipschitz property on (D, J_1) from (D, U). Unfortunately, we will have to work harder to obtain related results for the M_1 topologies.

Theorem 14.2.7. (Lipschitz property with d_{J_1}) *For any $Q \in \mathcal{Q}$, there exist constants K_1 and K_2 (the same as in Theorem 14.2.5) such that*

$$d_{J_1}(\psi(x_1), \psi(x_2)) \leq K_1 d_{J_1}(x_1, x_2) \tag{2.26}$$

and

$$d_{J_1}(\phi(x_1), \phi(x_2)) \leq K_2 d_{J_1}(x_1, x_2) \tag{2.27}$$

for all $x_1, x_2 \in D$.

Proof. The argument is the same for (2.26) and (2.27), so we prove only (2.26). By the definition of d_{J_1} and Theorems 14.2.6(iii) and 14.2.5,

$$\begin{aligned} d_{J_1}(\psi(x_1), \psi(x_2)) &\equiv \inf_{\lambda \in \Lambda} \{\|\psi(x_1) \circ \lambda - \psi(x_2)\| \vee \|\lambda - e\|\} \\ &= \inf_{\lambda \in \Lambda} \{\|\psi(x_1 \circ \lambda) - \psi(x_2)\| \vee \|\lambda - e\|\} \\ &\leq \inf_{\lambda \in \Lambda} \{K_1 \|x_1 \circ \lambda - x_2\| \vee \|\lambda - e\|\} \\ &\leq (K_1 \vee 1) d_{J_1}(x_1, x_2) \ , \end{aligned}$$

which implies (2.26) because $K_1 \geq 1$. ∎

Even when we establish convergence to stochastic-process limits with continuous sample paths, we need maps on D to be measurable as well as continuous at $x \in C$ in order to apply the continuous mapping theorem on D. From Theorem 14.2.5, it follows that the reflection map is measurable with respect to the Borel σ-field associated with the uniform topology (on both the domain and range), but that does not imply measurability with respect to the usual Kolmogorov σ-field generated by the coordinate projections, because the Borel σ-field on (D, U) is much larger than the Kolmogorov σ-field. Indeed, the Borel σ-field on (D, U) tends to be too large, causing severe measurability problems; see Section 18 of Billingsley (1968). Fortunately, Theorem 14.2.7 directly implies the desired measurability.

Corollary 14.2.2. (measurability) *The reflection map $R : D^k \to D^{2k}$ is measurable, using the Kolmogorov σ-field on the domain and range.*

Proof. The continuity established in Theorem 14.2.7 implies measurability with respect to the Borel σ-fields, but the Borel σ-field on (D, J_1) coincides with the Kolmogorov σ-field generated by the projection map; see Theorem 11.5.2. ∎

We now want to consider the reflection map R as a function of the reflection matrix Q as well as the net input function x. We first consider the maps $\pi \equiv \pi^n_{x,Q}(0)$ in (2.10) and $\psi \equiv \psi_Q$ in (2.4) as functions of Q when Q is a strict contraction in the matrix norm (2.12), i.e., when $\|Q\| < 1$.

Theorem 14.2.8. (stability bounds for different reflection matrices) *Let $Q_1, Q_2 \in \mathcal{Q}$ with $\|Q_1\| = \gamma_1 < 1$ and $\|Q_2\| = \gamma_2 < 1$. For all $n \geq 1$,*

$$\|\pi^n_{x,Q_j}(0)\| \leq (1 + \gamma_j + \cdots + \gamma_j^{n-1})\|x\| \tag{2.28}$$

and

$$\|\pi^n_{x,Q_1}(0) - \pi^n_{x,Q_2}(0)\| \leq (1 + \gamma_2 + \cdots + \gamma_2^{n-1})\frac{\|x\| \cdot \|Q_1 - Q_2\|}{1 - \gamma_1}\ , \tag{2.29}$$

so that

$$\|\psi_{Q_j}(x)\| \leq \frac{\|x\|}{1 - \gamma_j} \tag{2.30}$$

and

$$\|\psi_{Q_1}(x) - \psi_{Q_2}(x)\| \leq \frac{\|x\| \cdot \|Q_1 - Q_2\|}{(1 - \gamma_1)(1 - \gamma_2)}\ . \tag{2.31}$$

Corollary 14.2.3. (continuity as a function of Q) *If $Q_n \to Q$ in $\mathcal{Q}$,*

$$\|R_{Q_n}(x) - R_Q(x)\| \to 0 \quad as \quad n \to \infty$$

for each $x \in D$.

Proof. Apply Theorem 14.2.8, noting that $\|Q_n\| \to \|Q\|$ if $Q_n \to Q$. ∎

We now exploit the fact that any $Q \in \mathcal{Q}$ can be transformed into another matrix Q_* in $\mathcal{Q}$ with $\|Q_*\| < 1$ by letting

$$Q_* \equiv \Lambda^{-1} Q \Lambda \tag{2.32}$$

for a suitable strictly positive diagonal matrix Λ. We note that a non-negative matrix Q is in $\mathcal{Q}$ if and only if $|sp(Q)| < 1$, where $sp(Q)$ is its spectrum (set of eigenvalues) and $|sp(Q)|$ is its spectral radius (supremum of the eigenvalue norms): There is one real eigenvalue equal to $|sp(Q)|$ and all other eigenvalues λ (in general complex valued) satisfy $|\lambda| \leq |sp(Q)|$;

e.g., see Seneta (1981). Next note that the transformation from Q to Q_* in (2.32) leaves the eigenvalues unchanged: If λ is an eigenvalue of Q with an associated left eigenvector u, i.e., if $uQ = \lambda u$, then

$$(u\Lambda)Q_* = u\Lambda(\Lambda^{-1}Q\Lambda) = (uQ\Lambda) = \lambda(u\Lambda) ,$$

so that λ is also an eigenvalue of Q_* with left eigenvector $u\Lambda$.

Lemma 14.2.7. (equivalence to a contractive reflection matrix) *For any* $Q \in \mathcal{Q}$, *there exists a positive diagonal matrix* Λ *such that* $\|Q_*\| < 1$ *for* Q_* *in* (2.32).

Proof. By Corollary 14.3.3 in Section 14.3 below, for any $u \in \mathbb{R}^k$, the map $\pi_0(v) \equiv (Qv - u)^+$ has a unique fixed point in $\mathbb{R}^k$. Let v be a vector in $\mathbb{R}^k$ such that $1 + Qv = v$. Necessarily $v \geq 1$. Let $\Lambda \equiv diag(1/v^j)$. Then

$$\Lambda v - \Lambda 1 = \Lambda(1 + Qv) - \Lambda 1 ,$$

so that

$$1 - \Lambda 1 = \Lambda Q v = \Lambda Q \Lambda^{-1} \equiv Q_* .$$

Since $0 \leq 1 - \Lambda 1 < 1$, $\|Q_*\| < 1$. ∎

By Theorem 14.2.6(iii), we can relate the reflection maps R_Q and R_{Q_*}.

Lemma 14.2.8. (reflections associated with equivalent reflection matrices) *Let* Q *and* Q_* *be reflection matrices in* $\mathcal{Q}$ *related by* (2.32) *for some strictly positive diagonal matrix* Λ. *Then*

$$\Psi_{Q_*}(\Lambda^{-1}x) = \Lambda^{-1}\Psi_Q(x) \equiv \{\Lambda^{-1}x : x \in \Psi_Q(x)\}$$

and

$$R_{Q_*}(\Lambda^{-1}x) = \Lambda^{-1}R_Q(x)$$

for all $x \in D$.

Proof. For any $w \in \Psi_Q(x)$,

$$\begin{aligned} 0 \leq x + (I - Q)w &= x + (I - \Lambda Q_* \Lambda^{-1})w \\ &= \Lambda\Lambda^{-1}x + (\Lambda\Lambda^{-1} - \Lambda Q_* \Lambda^{-1})w \\ &= \Lambda(\Lambda^{-1}x + (I - Q_*)(\Lambda^{-1}w)) , \end{aligned}$$

so that $\Lambda^{-1}w \in \Psi_{Q_*}(\Lambda^{-1}x)$, with $w_1 \leq w_2$ if and only if $\Lambda^{-1}w_1 \leq \Lambda^{-1}w_2$, so that $\Psi_{Q_*}(\Lambda^{-1}x) = \Lambda^{-1}\psi_Q(x)$ and

$$\Lambda^{-1}z = \Lambda^{-1}x + (I - Q_*)(\Lambda^{-1}y) ,$$

so that $\phi_{Q_*}(\Lambda^{-1}x) = \Lambda^{-1}\psi_Q(x)$. ∎

Theorem 14.2.9. (continuity as a function of x and Q) *If* $\|x_n - x\| \to 0$ *in* D^k *and* $Q_n \to Q$ *in* $\mathcal{Q}$, *then*

$$\|R_{Q_n}(x_n) - R_Q(x)\| \to 0 \quad in \quad D^{2k} .$$

Proof. By Lemma 14.2.7, we can find a positive diagonal matrix Λ so that $Q_* = \Lambda^{-1}Q\Lambda$ and $\|Q_*\| = \gamma < 1$. Since $Q_n \to Q$ as $n \to \infty$, $\|Q_{n*}\| \equiv \gamma_n \to \gamma$, where $Q_{n*} \equiv \Lambda^{-1}Q_n\Lambda$ with the same diagonal matrix used above. Consider n sufficiently large that $\gamma_n < 1$. Since $\psi_{Q^*}(\Lambda^{-1}x) = \Lambda^{-1}\psi_Q(x)$, for such n we have

$$\begin{aligned}\|\psi_{Q_n}(x_n) - \psi_Q(x)\| &= \|\Lambda\Lambda^{-1}\psi_{Q_n}(x_n) - \Lambda\Lambda^{-1}\psi_Q(x)\| \\ &\le \|\Lambda\| \cdot \|\psi_{Q_{n*}}(\Lambda^{-1}x_n) - \psi_{Q_*}(\Lambda^{-1}x)\| \\ &\le \|\Lambda\|(\|\psi_{Q_{n*}}(\Lambda^{-1}x_n) - \psi_{Q_{n*}}(\Lambda^{-1}x)\| \\ &\qquad + \|\psi_{Q_{n*}}(\Lambda^{-1}x) - \psi_{Q_*}(\Lambda^{-1}x)\|) .\end{aligned}$$

Thus, by (2.30) and (2.31),

$$\begin{aligned}\|\psi_{Q_n}(x_n) - \psi_Q(x)\| &\le \|\Lambda^{-1}\| \left(\frac{\|\Lambda x_n - \Lambda x\|}{1-\gamma_n} + \frac{\|\Lambda x\| \cdot \|Q_{n*} - Q_*\|}{(1-\gamma_n)(1-\gamma)} \right) \\ &\le M_n \left(\|x_n - x\| + \frac{\|x\| \cdot M_n \cdot (1-\gamma_n) \cdot \|Q_n - Q\|}{1-\gamma} \right)\end{aligned} \tag{2.33}$$

for

$$M_n \equiv \frac{\|\Lambda^{-1}\| \cdot \|\Lambda\|}{1-\gamma_n} .$$

Hence,

$$\|\psi_{Q_n}(x_n) - \psi_Q(x)\| \to 0 \quad \text{as} \quad n \to \infty . \quad \blacksquare$$

14.3. The Instantaneous Reflection Map

In this section we introduce yet another characterization of the reflection map. We represent the reflection map on D as the limit of reflections of functions in D_c, the subset of piecewise-constant functions in D. (Recall from Section 12.2 that any function in D can be approximated uniformly by functions in D_c.) On the subset D_c, the reflection map reduces to an iterative application of an instantaneous reflection map on $\mathbb{R}^k$, which is of interest itself.

The instantaneous reflection map describes how

$$(y(0), z(0)) \equiv (\psi(x)(0), \phi(x)(0))$$

depends upon $x(0)$ and characterizes the behavior of the full reflection map at each discontinuity point. We will apply the instantaneous reflection map to establish monotonicity results (in particular, Corollary 14.3.2 below), which we will in turn apply to establish the M_1 continuity results. We use the final Lemma 14.3.4 in the next section to study reflections of parametric representations.

14.3.1. Definition and Characterization

Let the instantaneous reflection map be $R_0 \equiv (\phi_0, \psi_0) : \mathbb{R}^k \rightarrow \mathbb{R}^{2k}$, where $\psi_0 : \mathbb{R}^k \rightarrow \mathbb{R}^k$ is defined by

$$\psi_0(u) \equiv \inf\{v \in \mathbb{R}_+^k : u + (I - Q)v \geq 0\} , \tag{3.1}$$

where $u_1 \leq u_2$ in $\mathbb{R}^k$ if $u_1^i \leq u_2^i$ in $\mathbb{R}$, $1 \leq i \leq k$. It turns out that the infimum in (3.1) is attained (so that we can refer to it as the minimum) and there are useful expressions for it. Given the solution to (3.1), we can define the other component of the instantaneous reflection map $\phi_0 : \mathbb{R}^k \rightarrow \mathbb{R}^k$ by

$$\phi_0(u) = u + (I - Q)\psi_0(u) . \tag{3.2}$$

The instantaneous reflection map is a version of the *linear complementarity problem* (LCP), which has a long history; see Cottle, Pang and Stone (1992). Given a vector $u \in R^k$ and a $k \times k$ matrix M, the LCP is to find a vector $v \in \mathbb{R}_+^k$ such that

$$u + Mv \geq 0 \tag{3.3}$$

and

$$v_i(u + Mv)_i = 0, \quad 1 \leq i \leq k . \tag{3.4}$$

It turns out that the vector v satisfying (3.1) also satisfies (3.4) for $M = I - Q$ with $Q \in \mathcal{Q}$; see Corollary 14.3.1 below. Moreover, it turns out that the LCP based on (u, M) has solutions for all vectors u for more general matrices M than $I - Q$ for $Q \in \mathcal{Q}$. Such generalizations provide a basis for defining reflection maps in terms of other matrices, but we do not pursue that generalization.

Since the instantaneous reflection map is defined on $\mathbb{R}^k$ instead of D^k, it is much easier to calculate. For example, we can calculate v satisfying (3.1) by solving the *linear program*

$$\begin{aligned} &\min c^t v \\ \text{subject to:} \quad &u + (I - Q)v \geq 0 \\ &v \geq 0 \end{aligned} \tag{3.5}$$

for any strictly positive vector c in $\mathbb{R}^k$. We can thus simulate piecewise-constant approximations to reflected stochastic processes by generating finitely many linear programs. (We briefly discuss simulation some more at the end of the section.)

Paralleling the stochastic network example used to motivate the definition of the reflection map in the beginning of Section 14.2, we can introduce a discrete-time fluid-network model to motivate the instantaneous reflection map. At each transition epoch there is an input and a potential output. We might have an instantaneous nonnegative input vector u_1 and potential instantaneous nonnegative output vector u_2 which is routed to other queues by the stochastic matrix Q, so that the overall potential instantaneous net input is

$$u = u_1 - u_2 + u_2 Q \ . \tag{3.6}$$

However, if the potential output u_2 in (3.6) exceeds the available supply, then we may have to disallow some of the output u_2. That can be accomplished by adding a minimal $(I - Q^t)v$ to u in (3.6), which gives (3.1). In fact, as we will show, the instantaneous reflection map in (3.1) and (3.2) is well defined for any $u \in \mathbb{R}^k$, not just for u of the form (3.6).

Given the instantaneous reflection map R_0 in (3.1) and (3.2) (which we have yet to show is well defined), it is straightforward to define the associated reflection map R on D_c. For any $x \in D_c$, the set of discontinuities is $Disc(x) = \{t_1, \ldots, t_m\}$ for some integer m and some time points t_i satisfying and $t_0 \equiv 0 < t_1 < \cdots < t_m < T$. Clearly we should have

$$\psi(x)(t_i) \equiv y(t_i) = \psi_0(z(t_{i-1}) + x(t_i) - x(t_{i-1})) + y(t_{i-1}) \tag{3.7}$$

and

$$\phi(x)(t_i) \equiv z(t_i) = \phi_0(z(t_{i-1}) + x(t_i) - x(t_{i-1})) \tag{3.8}$$

for $0 \le i \le m$, where $z(t_{-1}) \equiv y(t_{-1}) \equiv x(t_{-1}) \equiv 0$, and we let (y, z) be piecewise constant with $Disc(y, z) = Disc(x)$.

Theorem 14.3.1. (reflection map for piecewise-constant functions) *For all $x \in D_c$, the reflection map defined by (3.1), (3.7) and (3.8) is equivalent to the reflection map in Definition 14.2.1.*

Proof. By induction, we can reexpress (3.7) as

$$\begin{aligned}
y(t_i) &= \psi_0(z(t_{i-1}) + x(t_i) - x(t_{i-1})) + y(t_{i-1}) \\
&= \min\{v \in \mathbb{R}^k_+ : z(t_{i-1}) + x(t_i) - x(t_{i-1}) + (I - Q)v \ge 0\} + y(t_{i-1}) \\
&= \min\{v \in \mathbb{R}^k_+ : x(t_i) + (I - Q)y(t_{i-1}) + (I - Q)v \ge 0\} + y(t_{i-1}) \\
&= \min\{v \ge y(t_{i-1}) : x(t_i) + (I - Q)v \ge 0\} \ ,
\end{aligned}$$

which corresponds to Definition 14.2.1. ∎

Since D_c is a subset of D, all the theorems in Section 14.2 apply to the reflection map on D_c in (3.7) and (3.8) by virtue of Theorem 14.3.1. The instantaneous reflection map itself corresponds to the reflection map in (3.7) and (3.8) applied to constant functions. Thus, from Section 14.2, we already know that the instantaneous reflection map is well defined. However, we can deduce additional structure of the reflection map by focusing directly on the instantaneous reflection map.

We first establish upper and lower bounds on $\psi_0(u)$. For $u \in \mathbb{R}^k$, let

$$u^+ \equiv u \vee 0 \equiv (u^1 \vee 0, \ldots, u^k \vee 0) \quad \text{and} \quad u^- \equiv u \wedge 0 \equiv (u^1 \wedge 0, \ldots, u^k \wedge 0) \ .$$

Lemma 14.3.1. *(bounds on $\psi_0(u)$) For any $u \in \mathbb{R}^k$,*

$$0 \le -(u^-) \le \psi_0(u) \le -(I-Q)^{-1}u^- \ .$$

Proof. Let $v = -(I-Q^t)^{-1}u^-$ and note that

$$u + (I-Q^t)v = u - (I-Q^t)(I-Q^t)^{-1}u^- = u - u^- = u^+ \ge 0 \ .$$

Then, by the definition of ψ_0 in (3.1), $\psi_0(u) \le v$, which establishes the upper bound. By (3.2),

$$\phi_0(u) = u + (I-Q^t)\psi_0(u) \ge 0.$$

Since $\psi_0(u) \ge 0$ and $Q \ge 0$,

$$\psi_0(u) \ge -u + Q^t\psi_0(u) \ge -u \ ,$$

which implies the lower bound. ∎

We now establish an additivity property of ψ_0.

Lemma 14.3.2. *(additivity of ψ_0) If $0 \le v_0 \le \psi_0(u)$ in $\mathbb{R}^k$, then*

$$\psi_0(u) = \psi_0(u + (I-Q)v_0) + v_0 \ .$$

Proof. By (3.1),

$$\begin{aligned} \psi_0(u) &= \min\{v \in \mathbb{R}^k_+ : u + (I-Q)v \ge 0\} \\ &= \min\{v \in \mathbb{R}^k_+ : u + (I-Q)v_0 + (I-Q)(v - v_0) \ge 0\} \\ &= v_0 + \min\{v' \in \mathbb{R}^k_+ : u + (I-Q)v_0 + (I-Q)v \ge 0\} \\ &= v_0 + \psi_0(u + (I-Q)v_0) \ , \end{aligned}$$

using the condition in the penultimate step. ∎

We now characterize the instantaneous reflection map in terms of a linear map applied to the vector (u^+, u^-) in $\mathbb{R}^{2k}$ involving the positive and negative parts of the vector u. In particular, for any $u \in \mathbb{R}^k$, let

$$T(u) \equiv u^+ + Qu^- \tag{3.9}$$

and let T^k be the k-fold iterate of the map T, i.e., $T^k(u) = T(T^{k-1}(u))$ for $k \geq 1$ with $T^0(u) \equiv u$. Note that T is a nonlinear function from $\mathbb{R}^k$ to $\mathbb{R}^k$.

Theorem 14.3.2. (characterization of the instantaneous reflection map) *Let $u_n \equiv T(u_{n-1})$ for T in (3.9) and $u_0 \equiv u$. Then, for any $u \in \mathbb{R}^k$,*

$$u_{n-1}^+ \geq u_n^+ \geq 0 \tag{3.10}$$

and

$$0 \geq u_n^- \geq Q^n u_0^- \quad \textit{for all} \quad n\ , \tag{3.11}$$

so that

$$u_n^- \to 0, \quad u_n \to u_\infty \geq 0 \quad \textit{and} \quad \psi_0(u_n) \to 0 \quad \textit{as} \quad n \to \infty\,. \tag{3.12}$$

For each $n \geq 1$,

$$\psi_0(u) = -\sum_{k=0}^{n-1} u_k^- + \psi_0(u_n) \tag{3.13}$$

and

$$u_n = u - (I - Q)\sum_{k=0}^{n-1} u_k^-\ , \tag{3.14}$$

so that ψ_0 in (3.1) is well defined with

$$\psi_0(u) = -\sum_{k=0}^{\infty} u_k^- \equiv -\sum_{k=0}^{\infty} T^k(u)^- \tag{3.15}$$

and

$$\phi_0(u) = u_\infty \equiv \lim_{n\to\infty} T^n(u)\ . \tag{3.16}$$

Proof. Since $u_n = u_{n-1}^+ + Qu_{n-1}^-$ by (3.9), $Qu_{n-1}^- \leq u_n \leq u_{n-1}^+$, which implies (3.10) and $Qu_{n-1}^- \leq u_n^- \leq 0$. By induction, these inequalities imply (3.11). Since $Q^n \to 0$ as $n \to \infty$, (3.10) and (3.11) imply the first two limits in (3.12). By Lemma 14.3.1,

$$\psi_0(u_n) \leq -(I-Q)^{-1}u_n^- \ . \tag{3.17}$$

Since $u_n^- \to 0$, (3.16) implies the last limit in (3.12). Formula (3.13) follows from Lemmas 14.3.1 and 14.3.2 by induction. From (3.9), $u_n - u_{n-1} = -(I-Q)u_{n-1}^-$, from which (3.14) follows by induction. Since $\psi_0(u_n) \to 0$, (3.13) implies (3.15), where the sum is finite. Moreover, (3.12)–(3.15) imply that $u_\infty = u + (I-Q)\psi_0(u)$, which in turn implies (3.16). ∎

We can apply Theorem 14.3.2 to deduce the complementarity property in (3.4). This corollary provides an alternative proof to half of Theorem 14.2.3.

Corollary 14.3.1. (complementarity) *For any* $u \in \mathbb{R}^k$,

$$\phi_0^i(u)\psi_0^i(u) = 0 \quad \textit{for all} \quad i \ .$$

Proof. If $\phi_0^i(u) > 0$, then $u_n^i > 0$ for all n by (3.10), which implies that $(u_n^-)^i = 0$ for all n and $\psi_0^i(u) = 0$ by (3.15). On the other hand, if $\psi_0^i(u) > 0$, then $u_k^i < 0$ for some k by (3.15), which implies that $(u_k^+)^i = 0$ for some k, so that $u_\infty^i = 0$ by (3.10). ∎

Theorem 14.3.2 implies the following important monotonicity property.

Corollary 14.3.2. (monotonicity) *If* $u_1 \leq u_2$ *in* $\mathbb{R}^k$, *then*

$$\phi_0(u_1) \leq \phi_0(u_2) \quad \textit{and} \quad \psi_0(u_1) \geq \psi_0(u_2) \ .$$

We now apply Corollary 14.3.2 and Lemma 14.3.2 to show that there is regularity in the reflection map when increments have a common sign in all coordinates. First, it is elementary that, if $u \geq 0$ and $\Delta \geq 0$ in $\mathbb{R}^k$, then

$$\phi_0(u+\Delta) = u + \Delta \quad \text{and} \quad \psi(u+\Delta) = 0 \ .$$

The situation is more delicate when $u \geq 0$ but $\Delta \not\geq 0$, but we obtain regularity when $\Delta \leq 0$.

In particular, we now consider the instantaneous reflection map R_0 applied to $u - \sum_{i=1}^m \Delta_i$, where $u \geq 0$ and $\Delta_i \geq 0$ for all i. We show that the instantaneous reflection of $u - \sum_{i=1}^m \Delta_i$ is the same as the m^{th} iteration of the iterative reflection introducing the increments Δ_i one at a time (in any order). Specifically, let

$$u_j \equiv \phi_0(u_{j-1} - \Delta_j), \quad 1 \le j \le m \ , \tag{3.18}$$

with $u_0 \equiv u$.

Theorem 14.3.3. (iterations of the instantaneous reflection mapping) *If $u \ge 0$ and $\Delta_i \ge 0$ for $i \ge 1$, then*

$$\phi_0(u - \sum_{i=1}^{m} \Delta_i) = u_m \tag{3.19}$$

for u_m in (3.18) and

$$\psi_0(u - \sum_{i=1}^{m} \Delta_i) = v_m \equiv \sum_{i=1}^{m} \psi_0(u_{i-1} - \Delta_i) \tag{3.20}$$

for all $m \ge 1$.

Proof. We use induction on m. For $m = 1$, relations (3.19) and (3.20) hold by definition. Suppose that (3.19) and (3.20) hold for all positive integers up to m; we will show that they must also hold for $m + 1$. By Corollary 14.3.2,

$$0 \le \psi_0(u - \sum_{i=1}^{m} \Delta_i) \le \psi_0(u - \sum_{i=1}^{m+1} \Delta_i) \ .$$

By Lemma 14.3.2,

$$\psi_0(u - \sum_{i=1}^{m+1} \Delta_i) = \psi_0(u - \sum_{i=1}^{m+1} \Delta_i + (I - Q^t)v_m) + v_m$$

for $v_j \equiv \psi_0(u - \sum_{i=1}^{j} \Delta_i)$ for $j \ge 1$. By the induction hypothesis,

$$\begin{aligned} \psi_0(u - \sum_{i=1}^{m+1} \Delta_i) &= \psi_0(\phi_0(u - \sum_{i=1}^{m} \Delta_i) - \Delta_{m+1}) + v_m \\ &= \psi_0(u_m - \Delta_{m+1}) + v_m = v_{m+1} \ . \end{aligned}$$

Moreover, by (3.2),

$$\begin{aligned}
\phi_0(u - \sum_{i=1}^{m+1} \Delta_i) &= u - \sum_{i=1}^{m+1} \Delta_i + (I - Q^t)\psi_0(u - \sum_{i=1}^{m+1} \Delta_i) \\
&= u - \sum_{i=1}^{m+1} \Delta_i + (I - Q^t)(\psi_0(u_m - \Delta_{m+1}) + v_m) \\
&= u_m - \Delta_{m+1} + (I - Q^t)(\psi_0(u_m - \Delta_{m+1})) \\
&= u_{m+1} \; . \quad \blacksquare
\end{aligned}$$

It is easy to see that the conclusion of Theorem 14.3.3 does not hold even for $m = 2$ and $k = 1$ when $\Delta_1 > 0$ and $\Delta_2 < 0$.

Example 14.3.1. *Need for common signs.* To see the need for having $\Delta_i \geq 0$ in $\mathbb{R}^k$ for all i in Theorem 14.3.3, suppose that $k = 2$, Q is as in Example 14.5.4, $u = (1, 0)$, $\Delta_1 = (3, -1)$ and $\Delta_2 = (3, -10)$. Then $\phi_0(u - \Delta_1 - \Delta_2) = \phi_0(-5, 11) = (0, 6)$, but $\phi_0(\phi_0(u - \Delta_1) - \Delta_2)) = (0, 7)$. ■

We can also deduce the following alternative characterization of the instantaneous reflection map.

Corollary 14.3.3. (fixed-point characterization) *For any $u \in \mathbb{R}^k$, $\psi_0(u)$ can be characterized as the unique solution v in $\mathbb{R}^k_+$ to the equation*

$$v = \pi_0(v) \equiv (Qv - u)^+ \; . \tag{3.21}$$

Proof. We use the fact that

$$-Qu_k^- = u_k^+ - u_{k+1} \; ,$$

drawing on (3.9). Then

$$\begin{aligned}
(Q\psi(u) - u)^+ &= \lim_{n\to\infty} \left(Q \sum_{k=0}^{n-1} (-u_k^-) - u_0 \right)^+ \\
&= \lim_{n\to\infty} \left(\sum_{k=0}^{n-1} u_k^+ - \sum_{k=1}^{n} u_k - u_0 \right)^+ \\
&= \lim_{n\to\infty} \left(-u_n^+ - \sum_{k=0}^{n} u_k^- \right)^+ \\
&= (-\phi(u) + \psi(u))^+ \; .
\end{aligned}$$

By Corollary 14.3.1, the last expression equals $\psi(u)$: If $\psi^i(u) > 0$, then $\phi^i(u) = 0$ and $(-\phi^i(u) + \psi^i(u)) = \psi^i(u) > 0$; if $\phi^i(u) > 0$, then $\psi^i(u) = 0$ and $-\phi^i(u) + \psi^i(u) < 0$, so that $(-\phi^i(u) + \psi^i(u))^+ = 0 = \psi^i(u)$. It remains to establish uniqueness. Suppose that v_1 and v_2 are two solutions to equation (3.21). Then

$$\begin{aligned} \|v_1 - v_2\| &= \|(Qv_1 - u)^+ - (Qv_2 - u)^+\| \\ &\leq \|(Qv_1 - u) - (Qv_2 - u)\| = \|Q(v_1 - v_2)\| \end{aligned}$$

Suppose that $\|z\| = \|Qz\|$. Then, by induction, $\|z\| = \|Q^n z\|$ for all $n \geq 1$, but $Q^n \to 0$ as $n \to \infty$, so that we must have $z = 0$. Hence $v_1 = v_2$ and the solution to (3.21) is unique. ∎

14.3.2. Implications for the Reflection Map

We can now extend the reflection map defined on D_c to D. We can *define* $R(x)$ for $x \in D$ by

$$R(x) \equiv \lim_{n\to\infty} R(x_n) \tag{3.22}$$

for $x_n \in D_c$ with $\|x_n - x\| \to 0$.

Theorem 14.3.4. (extension of the reflection map from D_c to D) *For all $x \in D$, the limit in* (3.22) *exists and is unique. Moreover, R is Lipschitz as a map from $(D, \|\cdot\|)$ to $(D, \|\cdot\|)$ and satisfies properties* (2.4)–(2.6).

Proof. For $x \in D$ given, choose $x_n \in D_c$ with $\|x_n - x\| \to 0$. Since $\|x_n - x\| \to 0$ for $x_n \in D_c$, $\|x_n - x_m\| \to 0$ as $m, n \to \infty$. As noted above, we can deduce the Lipschitz property of R on D_c by applying Theorem 14.2.5. By that Lipschitz property on D_c, $\|R(x_n) - R(x_m)\| \leq K\|x_n - x_m\| \to 0$. Since $(D, \|\cdot\|)$ is a complete metric space, there exists $(y, z) \in D$ such that $\|R(x_n) - (y, z)\| \to 0$. To show uniqueness, suppose that $\|x_{jn} - x\| \to 0$ for $j = 1, 2$. Then $\|x_{1n} - x_{2n}\| \to 0$ and $\|R(x_{1n}) - R(x_{2n})\| \leq K\|x_{1n} - x_{2n}\| \to 0$, so that the limits necessarily coincide. Given that $x_n \in D_c$, so that (x_n, y_n, z_n) satisfy (2.4), (2.6) and (2.11) with $\|(x_n, y_n, z_n) - (x, y, z)\| \to 0$, it follows that (x, y, z) satisfies (2.4), (2.6) and (2.11) too. (If (2.11) were to be violated for (z^i, y^i) for some i, then it follows that (2.11) would necessarily be violated by (z_n^i, y_n^i) for some n, because there would exist an interval $[a, b]$ in $[0, T]$ such that $z^i(t) \geq \epsilon > 0$ for $a \leq t \leq b$ and $y^i(b) > y^i(a)$.) Alternatively, since there exists a unique solution to (2.4) and (2.6), it must coincide with the one obtained via the limit (3.22). To directly verify the Lipschitz property on D given the Lipschitz property on D_c, for any $x_1, x_2 \in D$, let x_{1n}, $x_{2n} \in D_c$ with $\|x_{1n} - x_1\| \to 0$ and $\|x_{2n} - x_2\| \to 0$. Then, for any $\epsilon > 0$, there is an n_0 such that

$$\begin{aligned}\|R(x_1) - R(x_2)\| &\leq \|R(x_1) - R(x_{1n})\| + \|R(x_{1n}) - R(x_{2n})\| + \|R(x_{2n}) - R(x_2)\| \\ &\leq K(\|x_1 - x_{1n}\| + \|x_{1n} - x_{2n}\| + \|x_{2n} - x_2\|) \\ &\leq K\|x_1 - x_2\| + 2K(\|x_1 - x_{1n}\| + \|x_{2n} - x_2\|) \\ &\leq K\|x_1 - x_2\| + \epsilon\end{aligned}$$

for all $n \geq n_0$. Since ϵ was arbitrary, the Lipschitz property is established. ∎

From Theorems 14.2.6(iv), 14.3.1 and 14.3.2, we have the following result.

Lemma 14.3.3. (the reflection map at discontinuity points) *For any $x \in D$ and t, $0 < t < T$,*

$$\psi(x)(t) \equiv y(t) = \psi_0(z(t-) + x(t) - x(t-)) + y(t-)$$

and

$$\phi(x)(t) \equiv z(t) = \phi_0(z(t-) + x(t) - x(t-)) \ .$$

We can apply Lemma 14.3.3 to relate the set of discontinuity points of $R(x)$ to the set of discontinuity points of x, which we denote by $Disc(x)$.

Corollary 14.3.4. (the set of discontinuity points) *For any $x \in D$,*

$$Disc(R(x)) = Disc(x) \ .$$

Proof. By Lemma 14.3.3, we can write

$$z(t) - z(t-) = x(t) - x(t-) + (I - Q)(y(t) - y(t-)) \ ,$$

where $y^i(t) - y^i(t-)$ is minimal, $1 \leq i \leq k$. If $x(t) - x(t-) = 0$ (where here 0 is the zero vector), then necessarily $y(t) - y(t-) = 0$, which then forces $z(t) - z(t-) = 0$. On the other hand, if $x(t) - x(t-) \neq 0$, then we cannot have both $z(t) - z(t-) = 0$ and $y(t) - y(t-) = 0$, so we must have $t \in Disc(R(x))$. ∎

We obtain our strongest results for the case in which no coordinate of x has a negative jump. Let D_+ be the subset of functions x for which $x(t) - x(t-) \geq 0$ for all t.

Corollary 14.3.5. (stronger result in D_+) *For any* $x \in D_+$, *we have* $\psi(x) \in C$, $\phi(x) \in D_+$ *and*

$$\phi(x)(t) - \phi(x)(t-) = x(t) - x(t-) \ .$$

Finally, we can apply Lemma 14.3.3 and Corollary 14.3.2 to show how reflections of parametric representations perform. We consider the reflection map applied to α times the increment as a function of α for $0 \le \alpha \le 1$. We will apply the following lemma in the next section.

Lemma 14.3.4. (instantaneous reflection at discontinuity points) *Suppose that* $x \in D$, $t \in Disc(x)$ *and* $0 \le \alpha \le 1$.

(a) If $x(t) \ge x(t-)$, *then*

$$\hat{\psi}(x,t,\alpha) \equiv \psi_0(z(t-)+\alpha[x(t)-x(t-)])+y(t-) = \hat{\psi}(x,t,0) = y(t-) \quad (3.23)$$

and

$$\hat{\phi}(x,t,\alpha) \equiv \phi_0(z(t-)+\alpha[x(t)-x(t-)]) = \hat{\phi}(x,t,0)+\alpha[x(t)-x(t-)] \quad (3.24)$$

for $0 \le \alpha \le 1$.

(b) If $x(t) \le x(t-)$ *and* $0 \le \alpha_1 < \alpha_2 \le 1$, *then*

$$\hat{\psi}(x,t,\alpha_1) \le \hat{\psi}(x,t,\alpha_2)$$

and

$$\hat{\phi}(x,t,\alpha_1) \ge \hat{\phi}(x,t,\alpha_2)$$

for $\hat{\psi}$ *in* (3.23) *and* $\hat{\phi}$ *in* (3.24).

We conclude this section by further discussing the possibility of using the instantaneous reflection map to simulate reflected stochastic processes in D^k. Given a stochastic process X, we can approximate it by the associated discrete-time stochastic process

$$X_n(t) \equiv X(\lfloor nt \rfloor/n), \quad t \ge 0 \ ,$$

for some suitably large n. For each n, X_n has sample paths in D_c. Given that X has sample paths in D, it is easy to see that $X_n \to X$ in (D, J_1) w.p.1 as $n \to \infty$. We simulate X_n for suitably large n by generating the random vectors $X(k/n) - X((k-1)/n)$ for $k \ge 1$. When X is a Lévy process, $X(k/n) - X((k-1)/n)$ for $k \ge 1$ will be IID random vectors with an infinitely divisible distribution. When X is a stable Lévy motion, the random vectors will have a stable law. For discussion about simulation of stable random vectors and processes, see Janicki and Weron (1993).

Given a sample path of $\{X(k/n) : k \geq 1\}$, we can calculate the sample path of the associated reflected process $\phi(X_n)$ by solving the linear complementarity problem (LCP) at each transition epoch k/n. Thus the substantial literature on LCP can be applied; see Cottle, Pang and Stone (1992). For example, as noted before, we can use linear programming to solve the LCP, recognizing that only the trivial calculation associated with $v = 0$ in (3.1) occurs whenever $u \geq 0$. (See (3.5) above.)

Instead of linear programming, we can also use Theorem 14.3.2 to do the calculation. We can approximate $\phi(u)$ by $T^n(u)$ for the map in (3.9). Even though the operator T on $\mathbb{R}^k$ in (3.9) must be applied many times to calculate each instantaneous reflection, the algorithm can be effective, because T itself is remarkably simple. By (3.10) and (3.16), u_n^+ is an upper bound for $\phi(u)$, where $u_n = T^n(u)$. Moreover, we can apply (3.17) to bound the error $\|u_n^+ - \phi(u)\|$. Since

$$\phi(u) = \phi(u_n) = u_n + (I - Q)\psi(u_n) \ ,$$

$$\begin{aligned} \|u_n^+ - \phi(u)\| &\leq \|u_n - \phi(u_n)\| \leq \|I - Q\| \cdot \|\psi(u_n)\| \\ &\leq \|I - Q\| \cdot \|(I - Q)^{-1} u_n^-\| \ . \end{aligned}$$

So $T^n(u) \equiv u_n$ provides an upper bound on $\phi(u)$ via u_n^+ and an upper bound on the error $\|u_n^+ - \phi(u)\|$ via u_n^-.

14.4. Reflections of Parametric Representations

In order to establish continuity and stronger Lipschitz properties of the reflection map R on D with the M_1 topologies, we would like to have $(R(u), r)$ be a parametric representation of $R(x)$ when (u, r) is a parametric representation of x. We now obtain positive results in this direction. Proofs appear in the Internet Supplement.

Theorem 14.4.1. (reflections of parametric representations) *Suppose that* $x \in D$, $(u, r) \in \Pi_s(x)$ *and* $r^{-1}(t) = [s_-(t), s_+(t)]$.

(a) If $t \in Disc(x)^c$, *then*

$$R(u)(s) = R(x)(t) \quad for \quad s_-(t) \leq s \leq s_+(t) \ .$$

(b) If $t \in Disc(x)$, *then*

$$R(u)(s_-(t)) = R(x)(t-) \quad and \quad R(u)(s_+(t)) = R(x)(t) \ .$$

(c) If $t \in Disc(x)$ and $x(t) \geq x(t-)$, then

$$\phi(u)(s) = \phi(x)(t-) + \left(\frac{u^j(s) - u^j(s_-(t))}{u^j(s_+(t)) - u^j(s_-(t))} \right) [x(t) - x(t-)]$$

for any j, $1 \leq j \leq k$, and

$$\psi(u)(s) = \psi(x)(t-) = \psi(x)(t) \quad \textit{for} \quad s_-(t) \leq s \leq s_+(t) \ ,$$

so that

$$R(u)(s) \in [R(x)(t-), R(x)(t)] \quad \textit{for} \quad s_-(t) \leq s \leq s_+(t) \ .$$

(d) If $t \in Disc(x)$ and $x(t) \leq x(t-)$, then $\phi^i(u)$ and $\psi^i(u)$ are monotone in $[s_-(t), s_+(t)]$ for each i, so that

$$R(u)(s) \in [[R(x)(t-), R(x)(t)]] \quad \textit{for} \quad s_-(t) \leq s \leq s_+(t) \ .$$

We can draw the desired conclusion that $(R(u), r)$ is a parametric representation of $R(x)$ if we can apply parts (c) and (d) of Theorem 14.4.1 to all jumps. Recall that D_+ (D_s) is the subset of D for which condition (c) (condition (c) or (d)) holds at all discontinuity points of x. For $x \in D_s$, the direction of the inequality is allowed to depend upon t.

Theorem 14.4.2. (preservation of parametric representations under reflection) *Suppose that $x \in D$ and $(u, r) \in \Pi_s(x)$.*
(a) If $x \in D_+$, then $(R(u), r) \in \Pi_s(R(x))$.
(b) If $x \in D_s$, then $(R(u), r) \in \Pi_w(R(x))$.

We also have an analog of Theorems 14.4.1 and 14.4.2 for the case $x \in D_s$ and $(u, r) \in \Pi_w(x)$.

Theorem 14.4.3. (preservation of weak parametric representations) *If $x \in D_s$ and $(u, r) \in \Pi_w(x)$, then $(R(u), r) \in \Pi_w(R(x))$.*

As a basis for proving Theorem 14.4.1, we exploit piecewise-constant approximations.

Lemma 14.4.1. (left and right limits) *For any $x \in D_c$, $(u, r) \in \Pi_s(x)$ and $r^{-1}(t) = [s_-(t), s_+(t)]$,*

$$R(u)(s_-(t)) = R(x)(t-) \quad \textit{and} \quad R(u)(s_+(t)) = R(x)(t) \ . \tag{4.1}$$

We now show that it is essential in Lemma 14.4.1 to have $(u,r) \in \Pi_s(x)$ instead of just $(u,r) \in \Pi_w(x)$. We also show that we cannot improve upon Lemma 14.4.1 to conclude that $(R(u),r) \in \Pi_w(R(x))$ when $(u,r) \in \Pi_s(x)$.

Example 14.4.1. *Impossibility of improvements.* To demonstrate the points above, let $x \in D_c$ and R be defined by

$$x^1 = I_{[0,1)} - 3I_{[1,2]}, \quad x^2 = I_{[0,1)} + 2I_{[1,2]} \ ,$$

$$Q^t = \begin{pmatrix} 0 & 1 \\ 0 & .9 \end{pmatrix} , \quad \text{so that} \quad I - Q = \begin{pmatrix} 1 & 0 \\ -1 & .1 \end{pmatrix} .$$

Then $z^1 = z^2 = I_{[0,1)}$, $y^1 = 3I_{[1,2]}$ and $y^2 = 10I_{[1,2]}$. To see that the conclusion of Lemma 14.4.1 fails when we only have $(u,r) \in \Pi_w(x)$, let a parametric representation (u,r) in $\Pi_w(x)$ be defined by

$$\begin{aligned} & r(0) = 0, \quad r(1/3) = r(2/3) = 1, \quad r(1) = 2 \\ & u^1(0) = u^1(1/3) = 1, \quad u^1(1/2) = u^1(1) = -3 \qquad (4.2) \\ & u^2(0) = u^2(1/2) = 1, \quad u^2(2/3) = u^2(1) = 2 \end{aligned}$$

with r, u^1 and u^2 defined by linear interpolation elsewhere. Notice that $[s_-(1), s_+(1)] = [1/3, 2/3]$, $\phi^2(u)(1/2) = 0$ and $\phi^2(u)(2/3) = 1 > 0 = z^2(1)$. Moreover, $\phi^2(u)(s) = 1$ on $[2/3, 1]$.

Next, to see that we need not have $(R(u),r) \in \Pi_w(R(x))$ when $(u,r) \in \Pi_s(x)$, let r be defined in (4.2) and let the parametric representation (u,r) in $\Pi_s(x)$ be defined by

$$\begin{aligned} & u^1(0) = u^1(1/3) = 1, \ u^1(2/3) = u^1(1) = -3 \\ & u^2(0) = u^2(1/3) = 1, \ u^2(2/3) = u^2(1) = 2 \end{aligned}$$

with r, u^1, u^2 defined at other points by linear interpolation. Clearly $(u,r) \in \Pi_s(x)$. Note that $u^i(s) \ge 0$ for all $s \le 5/12$. Then $r(5/12) = 1$, $u^1(5/12) = 0$ and $u^2(5/12) = 5/4$. Clearly $\phi(u)(5/12) = u(5/12) = (0, 5/4)$, which is not in $[[(0,0),(1,1)]]$, the weak range of $z = \phi(x)$. Further analysis shows that $\phi^1(u)(s) = 0$ for $s \ge 5/12$, while $\phi^2(u) = u^2$ on $[0,1/3]$, $\phi^2(u)(5/12) = 5/4$, $\phi^2(u)(5/9) = \phi^2(u)(1) = 0$, with $\phi^2(u)$ defined elsewhere by linear interpolation. Similarly, ψ has slope $(12,0)$ over $(5/12, 5/9)$ and slope $(12,90)$ over $(5/9, 2/3)$, so that $\psi(u)(5/12) = (0,0)$, $\psi(u)(5/9) = (5/3, 0)$, $\psi(u)(2/3) = \psi(u)(1) = (3,10)$ and ψ is defined by linear interpolation elsewhere. ■

14.5. M_1 Continuity Results and Counterexamples

In this section we establish positive results and give counterexamples showing that candidate stronger results do not hold.

14.5.1. M_1 Continuity Results

We first state continuity and Lipschitz properties of the reflection map on $D \equiv D^k \equiv D([0,T],\mathbb{R}^k)$ with the M_1 topologies. Our first result establishes continuity of the reflection map R (for an arbitrary reflection matrix Q) as a map from (D, SM_1) to (D, L_1), where L_1 is the topology on D induced by the L_1 norm

$$\|x\|_{L_1} \equiv \int_0^T \|x(t)\| dt \ . \tag{5.1}$$

Under a further restriction, the map from (D, WM_1) to (D, WM_1) will be continuous.

Recall that D_s is the subset of functions in D without simultaneous jumps of opposite sign in the coordinate functions; i.e., $x \in D_s$ if, for all $t \in (0,T)$, either $x(t) - x(t-) \le 0$ or $x(t) - x(t-) \ge 0$, with the sign allowed to depend upon t. The subset D_s is a closed subset of D in the J_1 topology and thus a measurable subset of D with the SM_1 and WM_1 topologies (since the Borel σ-fields coincide). The proofs of the main theorems here appear in Section 8.5 of the Internet Supplement.

Theorem 14.5.1. (continuity with the SM_1 topology on the domain) *Suppose that $x_n \to x$ in (D, SM_1).*

(a) Then

$$R(x_n)(t_n) \to R(x)(t) \quad in \quad \mathbb{R}^{2k} \tag{5.2}$$

for each $t \in Disc(x)^c$ and sequence $\{t_n : n \ge 1\}$ with $t_n \to t$,

$$\sup_{n \ge 1} \|R(x_n)\| < \infty \ , \tag{5.3}$$

$$R(x_n) \to R(x) \quad in \quad (D, L_1) \tag{5.4}$$

and

$$\psi(x_n) \to \psi(x) \quad in \quad (D, WM_1) \ . \tag{5.5}$$

(b) If in addition $x \in D_s$, then

$$\phi(x_n) \to \phi(x) \quad in \quad (D, WM_1) \ , \tag{5.6}$$

so that

$$R(x_n) \to R(x) \quad in \quad (D, WM_1) \ . \tag{5.7}$$

Under the extra condition in part (b), the mode of convergence on the domain actually can be weakened. However, little positive can be said if only $x_n \to x$ in (D, WM_1) without $x \in D_s$; see Example 14.5.3 below.

Theorem 14.5.2. (continuity with the WM_1 topology on the domain) *If $x_n \to x$ in (D, WM_1) and $x \in D_s$, then* (5.7) *holds.*

Remark 14.5.1. *Gibbs phenomenon.* Interestingly, the limit of $\phi(x_n)(t_n)$ for $t_n \to t \in Disc(x)$ can fall outside the product segment $[[\phi(x)(t-), \phi(x)(t)]]$; see Example 14.5.4 below. Thus the asymptotic fluctuations in $\phi(x_n)$ can be greater than the fluctuations in $\phi(x)$. The behavior here is analogous to the Gibbs phenomenon associated with Fourier series; see Chapter 9 of Carslaw (1930) and Remark 5.1 of Abate and Whitt (1992a).

Example 12.3.1 shows that convergence $x_n \to x$ can hold in (D, WM_1) but not in (D, SM_1) even when $x \in D_s$. Thus Theorems 14.5.1 (a) and 14.5.2 cover distinct cases. An important special case of both occurs when $x \in D_1$, where D_1 is the subset of x in D with discontinuities in only one coordinate at a time; i.e., $x \in D_1$ if $t \in Disc(x^i)$ for at most one i when $t \in Disc(x)$, with the coordinate i allowed to depend upon t. In Section 12.7 it is shown that WM_1 convergence $x_n \to x$ is equivalent to SM_1 convergence when $x \in D_1$.

Just as with D_s above, D_1 is a closed subset of (D, J_1) and thus a Borel measurable subset of (D, SM_1). Since $D_1 \subseteq D_s$, the following corollary to Theorem 14.5.2 is immediate.

Corollary 14.5.1. (common case for applications) *If $x_n \to x$ in (D, WM_1) and $x \in D_1$, then $R(x_n) \to R(x)$ in (D, WM_1).*

We can obtain stronger Lipschitz properties on special subsets. Let D_+ be the subset of x in D with only nonnegative jumps, i.e., for which $x^i(t) - x^i(t-) \geq 0$ for all i and t. As with D_s and D_1 above, D_+ is a closed subset of (D, J_1) and thus a measurable subset of (D, SM_1).

Theorem 14.5.3. (*Lipschitz properties*) *There is a constant K (the same as associated with the uniform norm in* (2.25)) *such that*

$$d_s(R(x_1), R(x_2)) \leq K d_s(x_1, x_2) \tag{5.8}$$

for all $x_1, x_2 \in D_+$, and

$$d_p(R(x_1), R(x_2)) \leq d_w(R(x_1), R(x_2)) \leq K d_w(x_1, x_2) \leq K d_s(x_1, x_2) \tag{5.9}$$

for all $x_1, x_2 \in D_s$.

We can actually do somewhat better than in Theorem 14.5.1 when the limit is in D_+.

Theorem 14.5.4. (strong continuity when the limits is in D_+) *If*

$$x_n \to x \quad in \quad (D, SM_1) \ , \tag{5.10}$$

where $x \in D_+$, *then*

$$R(x_n) \to R(x) \quad in \quad (D, SM_1) \ . \tag{5.11}$$

Our final result shows how the reflection map behaves as a function of the reflection matrix Q, as well as x, with the M_1 topologies.

Theorem 14.5.5. (continuity as a function of (x, Q)) *Suppose that* $Q_n \to Q$ *in* $\mathcal{Q}$.

(a) If $x_n \to x$ *in* (D^k, WM_1) *and* $x \in D_s$, *then*

$$R_{Q_n}(x_n) \to R_Q(x) \quad in \quad (D^{2k}, WM_1) \ . \tag{5.12}$$

(b) If $x_n \to x$ *in* (D^k, SM_1) *and* $x \in D_+$, *then*

$$R_{Q_n}(x_n) \to R_Q(x) \quad in \quad (D^{2k}, SM_1) \ . \tag{5.13}$$

We can apply Section 12.9 to extend the continuity and Lipschitz results to the space $D([0, \infty), \mathbb{R}^k)$.

Theorem 14.5.6. (extension of continuity results to $D([0, \infty), \mathbb{R}^k)$) *The convergence-preservation results in Theorems 14.5.1, 14.5.2 and 14.5.4 and Corollary 14.5.1 extend to* $D([0, \infty), \mathbb{R}^k)$.

Proof. Suppose that $x_n \to x$ in $D([0, \infty), \mathbb{R}^k)$ with the appropriate topology and that $\{t_j : j \geq 1\}$ is a sequence of positive numbers with $t_j \in Disc(x)^c$ and $t_j \to \infty$ as $j \to \infty$. Then, $r_{t_j}(x_n) \to r_{t_j}(x)$ in $D([0, \infty), \mathbb{R}^k)$ with the same topology as $n \to \infty$ for each j, where r_t is the restriction map to $D([0, t], \mathbb{R}^k)$. Under the specified assumptions,

$$r_{t_j}(R(x_n)) = R_{t_j}(r_{t_j}(x_n)) \to R_{t_j}(r_{t_j}(x)) = r_{t_j}(R(x)) \tag{5.14}$$

in $D([0, t_j], \mathbb{R}^{2k})$ with the specified topology as $n \to \infty$ for each j, which implies that

$$R(x_n) \to R(x) \quad \text{in} \quad D([0, \infty), \mathbb{R}^{2k}) \tag{5.15}$$

with the same topology as in (5.14). ∎

Theorem 14.5.7. (extension of Lipschitz properties to $D([0,\infty),\mathbb{R}^k)$) *Let* $R : D([0,\infty),\mathbb{R}^k) \to D([0,\infty),\mathbb{R}^{2k})$ *be the reflection map with function domain* $[0,\infty)$ *defined by Definition* 14.2.1. *Let metrics associated with domain* $[0,\infty)$ *be defined in terms of restrictions by* (9.1) *in Section* 12.9. *Then the conclusions of Theorems* 14.2.5, 14.2.7 *and* 14.5.3 *also hold for domain* $[0,\infty)$.

Proof. Apply Theorem 12.9.4. ∎

14.5.2. Counterexamples

We now return to the space $D([0,T],\mathbb{R}^k)$ and present several counterexamples. We first show that the reflection map is actually not continuous on $D([0,T],\mathbb{R}^1)$ with the SM_1 topology. (This would not be a counterexample if we restricted attention to the component ϕ mapping x into z in (2.6) or, more generally, the WM_1 topology were used on the range.)

Example 14.5.1. *Not continuous on* (D, SM_1). To show that

$$R \equiv (\psi,\phi) : (D([0,2],\mathbb{R}^1), SM_1) \to (D([0,2],\mathbb{R}^2), SM_1)$$

is *not* continuous, let

$$x_n(t) = 1 - 2n(t-1)I_{[1,1+n^{-1})}(t) - 2I_{[1+n^{-1},2]}(t)$$

and

$$x(t) = 1 - 2I_{[1,2]}(t), \quad 0 \le t \le 2 \; .$$

It is easy to see that $d(x_n, x) \to 0$ as $n \to \infty$,

$$z_n(t) = 1 - 2n(t-1)I_{[1,1+(2n)^{-1})}(t) \; ,$$

$$y_n(t) = 2n(t-(1+(2n)^{-1}))I_{[1+(2n)^{-1},1+n^{-1})}(t) + I_{[1+n^{-1},2]}(t) \; ,$$

$$z(t) = I_{[0,1)}(t) \quad \text{and} \quad y(t) = I_{[1,2]}(t) \; .$$

We use the fact that any linear function of the coordinate functions, such as addition or subtraction, is continuous in the SM_1 topology; see Section 12.7. Note that $z(t) + y(t) = 1$, $0 \le t \le 2$, while

$$z_n(t)+y_n(t) = 1-2n(t-1)I_{[1,1+(2n)^{-1}]}(t)+2n(t-(1+(2n)^{-1})I_{[1+(2n)^{-1},1+n^{-1}]}(t)$$

so that $d(z_n + y_n, z + y) \not\to 0$ as $n \to \infty$, which implies that $(x_n, y_n) \not\to (z, y)$ as $n \to \infty$ in $D([0,T],\mathbb{R}^2)$ with the SM_1 metric. However, we do have $d(z_n, z) \to 0$ and $d(y_n, y) \to 0$ as $n \to \infty$, so the maps from x to y and z separately are continuous. ∎

Example 14.5.1 suggests that the difficulty might only be in simultaneously considering both maps ψ and ϕ. We show that this is not the case by giving a counterexample with ϕ alone (but again in two dimensions).

Example 14.5.2. *ϕ is not continuous on (D^2, SM_1).* We now show that

$$\phi : (D([0,2], \mathbb{R}^2), SM_1) \to (D([0,2], \mathbb{R}^2), SM_1)$$

is not continuous. We use the trivial reflection map corresponding to two separate queues, for which Q is the 2×2 matrix of 0's. Let x_n^1 be as in Example 14.5.1, i.e.,

$$x_n^1(t) = 1 - 2n(t-1)I_{[1,1+n^{-1})}(t) - 2I_{[1+n^{-1},2]}(t)$$

and let

$$x_n^2(t) = 2 - 3n(t-1)I_{[1,1+n^{-1})}(t) - 3I_{[1+n^{-1},2]}(t) \ .$$

It is easy to see that $d_s((x_n^1, x_n^2), (x^1, x^2)) \to 0$ as $n \to \infty$, where

$$x^1(t) = 1 - 2I_{[1,2]}(t) \quad \text{and} \quad x^2(t) = 2 - 3I_{[1,2]}(t) \ .$$

(The same functions r_n and r can be used in the parametric representations of the two coordinates.) Clearly $\phi((x^1, x^2)) = (z^1, z^2)$, where

$$z^1(t) = I_{[0,1)}(t) \quad \text{and} \quad z^2(t) = 2I_{[0,1)}(t) \ ,$$

while

$$\begin{aligned} z_n^1(t) &= 1 - 2n(t-1)I_{[1,1+(1/2n))}(t) \\ z_n^2(t) &= 2 - 3n(t-1)I_{[1,1+(2/3n))}(t) \ . \end{aligned}$$

Note that $2z^1(t) - z^2(t) = 0$, $0 \le t \le 2$, while

$$2z_n^1(1+(2n)^{-1}) - z_n^2(1+(2n)^{-1}) = -z_n^2(1+(2n)^{-1}) = -1/2 \quad \text{for all} \quad n \ .$$

Hence $d_s(2z_n^1 - z_n^2, 2z^1 - z^2) \not\to 0$ so that $d_s((z_n^1, z_n^2), (z^1, z^2)) \not\to 0$ as $n \to \infty$. However, in this example, ϕ is continuous if we use the WM_1 topology on the range. ■

We now show that the reflection map is not continuous if the WM_1 topology is used on the domain without imposing extra conditions.

Example 14.5.3. *The difficulty with the WM_1 topology on the domain.* We show that neither ψ nor ϕ need be continuous when the WM_1 topology is used on the domain, without extra conditions. Consider $D([0,2],\mathbb{R}^2)$ and let $x^1 = I_{[1,2]}$, $x^2 = -2I_{[1,2]}$ and

$$Q = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix} .$$

Then the reflection map yields $y^1(t) = \psi^1(x)(t) = z^i(t) = \phi^i(x)(t) = 0$, $0 \le t \le 2$, for $i = 1,2$ and $y^2 = \psi^2(x) = 2I_{[1,2]}$. Let the converging functions be $x_n^1 = I_{[1+n^{-1},2]}$ and $x_n^2 = -2I_{[1-n^{-1},2]}$ for $n \ge 1$. It is easy to see that $x_n \to x$ as $n \to \infty$ in WM_1 but that $x_n \not\to x$ as $n \to \infty$ in SM_1, because $(2x_n^1 + x_n^2)(1) = -2$, while $(2x^1 + x^2)(t) = 0$, $0 \le t \le 2$. The reflection map applied to x_n works on the jumps at times $1-n^{-1}$ and $1+n^{-1}$ separately, yielding $y_n^1 = (4/3)I_{[1-n^{-1},2]}$, $y_n^2 = (8/3)I_{[1-n^{-1},2]}$, $z_n^1 = I_{[1+n^{-1},2]}$ and $z_n^2(t) = 0$, $0 \le t \le 2$. Clearly $z_n^1 \not\to z^1$ and $y_n^i \not\to y^i$ as $n \to \infty$ for $i = 1,2$ for any reasonable topology on the range. In particular, conclusions (5.2) and (5.4) – (5.7) all fail in this example.

Moreover, when we choose suitable parametric representations $(u_n, r_n) \in \Pi_w(x_n)$ and $(u,r) \in \Pi_w(x)$ to achieve $x_n \to x$ in WM_1, $(R(u),r)$ is not a parametric representation for $R(x)$. To be clear about this, we give an example: We let all the functions u_n, r_n, u and r be piecewise-linear. We define the functions at the discontinuity points of the derivative. We understand that the functions are extended to $[0,1]$ by linear interpolation. Let

$$\begin{aligned}
&r(0) = 0,\ r(0.2) = r(0.8) = 1,\ r(1) = 2\ ,\\
&u^1(0) = u^1(0.4) = 0,\ u^1(0.8) = u^1(1) = 1\ ,\\
&u^2(0) = u^2(0.2) = 0,\ u^2(0.4) = u^2(1) = -2\ ,\\
&r_n(0) = 0,\ r_n(0.2(1-n^{-1})) = r_n(0.2(2-n^{-1})) = 1 - n^{-1}\ ,\\
&r_n(0.2(2+n^{-1})) = r_n(0.2(4+n^{-1})) = 1 + n^{-1},\ r_n(1) = 2\ ,\\
&u_n^1(0) = u_n^1(0.2(2+n^{-1})) = 0,\ u_n^1(0.2(4+n^{-1}) = u_n^1(1) = 1\ ,\\
&u_n^2(0) = u_n^2(0.2(1-n^{-1}) = 0,\ u_n^2(0.2(2-n^{-1}) = u_n^2(1) = -2\ .
\end{aligned}$$

This construction yields $(u_n, r_n) \in \Pi_w(x_n)$, $n \ge 1$, $(u,r) \in \Pi_w(x)$, but $(\phi^1(u), r) \notin \Pi(\phi^1(x))$, because $\phi^1(x)(t) = 0$, $0 \le t \le 2$, while $\phi^1(u)(1) = 1$. Note that $(u,r) \in \Pi_w(x)$, but $(u,r) \notin \Pi_s(x)$. ■

We now show that we need not have $R(x_n) \to R(x)$ in (D, WM_1) when $x_n \to x$ in (D, SM_1) without the extra regularity condition $x \in D_s$. A difficulty can occur when $x^i(t) - x^i(t-) > 0$ for some coordinate i, while $x^j(t) - x^j(t-) < 0$ for another coordinate j.

Example 14.5.4. *Need for the condition $x \in D_s$.* We now show that the condition $x \in D_s$ in Theorem 14.5.1 is necessary even when $x_n \to x$ in (D, SM_1). In our limit $x \equiv (x^1, x^2)$, x^1 has a jump down and x^2 has a jump up at $t = 1$. Our example is the simple network corresponding to two queues in series. Let $x \equiv (x^1, x^2)$ and $x_n \equiv (x_n^1, x_n^2)$, $n \geq 1$, be elements of $D([0,2], \mathbb{R}^2)$ defined by

$$x^1(0) = x^1(1-) = 1, \quad x^1(1) = x^1(2) = -3$$

$$x^2(0) = x^2(1-) = 1, \quad x^2(1) = x^2(2) = 2$$

$$x_n^1(0) = x_n^1(1) = 1, \quad x_n^1(1 + n^{-1}) = x_n^1(2) = -3$$

$$x_n^2(0) = x_n^2(1) = 1, \quad x_n^2(1 + n^{-1}) = x_n^2(2) = 2 \ ,$$

with the remaining values determined by linear interpolation. Let the substochastic matrix generating the reflection be

$$Q^t = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \text{so that} \quad I - Q = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} .$$

Then $z^1 = z^2 = I_{[0,1)}$, $y^1 = 3I_{[1,2]}, y^2 = I_{[1,2]}$ and

$$z_n^1(0) = z_n^1(1) = 1, \quad z_n^1(1 + (4n)^{-1}) = z_n^1(2) = 0$$

$$z_n^2(0) = z_n^2(1) = 1, \quad z_n^2(1 + (4n)^{-1}) = 5/4, \quad z_n^2(1 + 2(3n)^{-1}) = z_n^2(2) = 0$$

with the remaining values determined by linear interpolation. Since $z_n^2(1 + (4n)^{-1}) = 5/4$ for all n and $z^2(t) \leq 1$ for all t, z_n^2 fails to converge to z^2 in any of the Skorohod topologies. We remark that the graphs $G_{\phi(x_n)}$ of $\phi(x_n)$ do converge in the Hausdorff metric to the graph $G_{\phi(x)}$ of $\phi(x)$ augmented by the set $\{1\} \times [1, 5/4]$. This example motivates considering larger spaces of functions than D, which we discuss in Chapter 15.

14.6. Limits for Stochastic Fluid Networks

In this section we provide concrete stochastic applications of the convergence-preservation results for the multidimensional reflection map.

We consider the single-class open stochastic fluid network with Markovian routing introduced in Section 14.2.

Recall that the stochastic fluid network is characterized by a four-tuple $\{C, r, Q, X(0)\}$, where $C \equiv (C^1, \ldots, C^k)$ is the vector of exogenous cumulative input stochastic processes at the k stations, $r = (r^1, \ldots, r^k)$ is the vector of potential output rates at the stations, $P \equiv (P_{i,j})$ is the routing matrix and $X(0) \equiv (X^1(0), \ldots, X^k(0))$ is the nonnegative random vector of initial buffer contents. The stochastic processes $C^j \equiv \{C^j(t) : t \geq 0\}$ have nondecreasing nonnegative sample paths; $C^j(t)$ represents the cumulative input at station j during the time interval $[0, t]$. A proportion $P_{i,j}$ of all output from station i is routed to station j, while a proportion $p_i \equiv 1 - \sum_{j=1}^{k} P_{i,j}$ is routed out of the network. We assume that P is substochastic so that $P_{i,j} \geq 0$, $1 \leq j \leq k$, and $p_i \geq 0$, $1 \leq i \leq k$. Moreover, we assume that $P^n \to 0$ as $n \to \infty$, where P^n is the n^{th} power of P. The associated reflection matrix is the transpose $Q \equiv P^t$.

As a more concrete example, suppose that the exogenous input to station j is the sum of the inputs from m_j separate on-off sources. Let (j, i) index the i^{th} on-off source at station j. When the (j, i) source is on, it sends fluid input at rate $\lambda_{j,i}$; when it is off, it sends no input. Let $B_{j,i}(t)$ be the cumulative busy (on) time for source (j, i) during the time interval $[0, t]$. Then the exogenous input process at station j is

$$C^j(t) = \sum_{i=1}^{m_j} \lambda_{j,i} B_{j,i}(t), \quad t \geq 0 \ .$$

Since $B_{j,i}$ necessarily has continuous sample paths, the associated exogenous cumulative input processes C^j and C also have continuous sample paths in this special case.

In general, given the defining four-tuple $(C, r, Q, X(0))$, the associated $\mathbb{R}^k$-valued potential buffer-content process (or net-input process) is

$$X(t) \equiv X(0) + C(t) - (I - Q)rt, \quad t \geq 0 \ . \tag{6.1}$$

where Q is the transpose P^t. Since C^j has nondecreasing sample paths for each j, the sample paths of X are of bounded variation. In many special cases, the sample paths of X will be continuous as well.

The buffer-content stochastic process $Z \equiv (Z^1, \ldots, Z^k)$ is simply obtained by applying the reflection map to the potential buffer-content process X in (6.1), in particular,

$$Z \equiv \phi(X) \ , \tag{6.2}$$

where $R = (\psi, \phi)$ in (2.4)–(2.6). Again, we regard (6.2) as the definition. This stochastic fluid network model is more elementary than the queue-length processes in the queueing network in the following Section 14.7, because here the content process of interest Z is defined directly in terms of the reflection map, requiring only (6.1) and (6.2).

We now want to establish some limits for the stochastic processes. First, we obtain a model continuity or stability result.

14.6.1. Model Continuity

For this purpose, we consider a sequence of fluid network models indexed by n characterized by four-tuples $(C_n, r_n, Q_n, X_n(0))$. Let $\Rightarrow$ denote convergence in distribution.

Theorem 14.6.1. (stability for stochastic fluid networks) *If*

$$(C_n, X_n(0)) \Rightarrow (C, X(0))$$

in $D([0,\infty), \mathbb{R}^k) \times \mathbb{R}^k$, *where the topology is either* SM_1 *or* WM_1, $r_n \to r$ *and* $Q_n \to Q$ *in* $\mathcal{Q}$ *as* $n \to \infty$, *then*

$$(X_n, Y_n, Z_n) \Rightarrow (X, Y, Z) \quad \text{as} \quad n \to \infty \quad \text{in} \quad D([0,\infty), \mathbb{R}^{3k}) \ ,$$

with the same topology, where X_n *and* X *are the associated potential buffer-content processes defined by* (6.1), Y_n *and* Y *are the associated regulator processes, and* Z_n *and* Z *as the associated buffer-content processes, with*

$$R(X_n) \equiv (\psi(X_n), \phi(X_n)) \equiv (Y_n, Z_n), \quad n \geq 1 \ .$$

Proof. Apply the continuous mapping theorem with the continuous functions in (6.1) and (2.4)–(2.6), invoking Theorem 14.5.5 and Corollary 14.2.2. Note that C_n, C, X_n and X have sample paths in D_+. First apply the linear function in (6.1) mapping $(C_n, r_n, Q_n, X_n(0))$ into X_n; then apply R mapping X_n into (Y_n, Z_n). For the special case of common Q, we can invoke Theorem 14.5.3 instead of Theorem 14.5.5. ∎

Remark 14.6.1. *Sufficient conditions for* SM_1 *convergence.* If $P(C \in D_1) = 1$, i.e., if

$$P(Disc(C^i) \cap Disc(C^j) = \phi) = 1 \tag{6.3}$$

for all i, j with $1 \leq i, j \leq k$ and $i \neq j$, then the assumed SM_1 convergence $C_n \Rightarrow C$ is implied by WM_1 convergence. Since C_n and C have nondecreasing sample paths, the condition $C_n^i \Rightarrow C^i$ $D([0,\infty), \mathbb{R}, M_1)$ is equivalent to convergence of the finite-dimensional distributions at all time points t for which $P(t \in Disc(C^i)) = 0$, where $Disc(C^i)$ is the set of discontinuity points of C^i; see Corollary 12.5.1. ∎

Remark 14.6.2. *The case of continuous sample paths.* As we have indicated, it is natural for the cumulative input processes C_n to have continuous sample paths, but that does not imply that the limit C necessarily must have continuous sample paths. If C does in fact have continuous sample paths, then so do X, Y and Z. Then, the SM_1 topology reduces to the topology of uniform convergence on compact subsets. ■

We can also obtain a bound on the distance between (X_n, Y_n, Z_n) and (X, Y, X) using the Prohorov metric π on the probability measures on (D_+, SM_1). For random elements X_1 and X_2, let $\pi(X_1, X_2)$ denote the Prohorov metric in (2.2) applied to the probability laws of X_1 and X_2. The conclusion for the case of common Q then is:

Corollary 14.6.1. (bounds on the Prohorov distance) *For common Q, there exists a constant K such that*

$$\pi((X_n, Y_n, Z_n), (X, Y, Z)) \le K\pi((C_n, X_n(0)), (C, X(0))) \; .$$

Proof. Apply Theorems 3.4.2 and 14.5.3. ■

14.6.2. Heavy-Traffic Limits

We also can obtain heavy-traffic FCLTs for stochastic fluid networks by considering a sequence of models with appropriate scaling. The scaling allows for on-off sources with heavy-tailed busy-period and idle-period distributions, as in Section 8.5. The scaling also allows for strong dependence in the input processes.

Theorem 14.6.2. (heavy-traffic limit) *Consider a sequence of stochastic fluid networks $\{(C_n, r_n, Q_n, X_n(0)) : n \ge 1\}$. If there exist a constant H with $0 < H < 1$, an $\mathbb{R}^k$-valued random vector $\mathbf{X}(0)$, vectors $\alpha_n \in \mathbb{R}^k$, $n \ge 1$, and a stochastic process $\mathbf{C}$ such that*

$$(\mathbf{C}_n, \mathbf{X}_n(0)) \Rightarrow (\mathbf{C}, \mathbf{X}(0)) \quad \textit{in} \quad D([0,\infty), \mathbb{R}^k, WM_1) \times \mathbb{R}^k \; , \tag{6.4}$$

where

$$\mathbf{C}_n(t) \equiv n^{-H}(C_n(nt) - \alpha_n nt), \quad t \ge 0 \, ,$$
$$P(\mathbf{C} \in D_s) = 1 \; , \tag{6.5}$$

and

$$n^{1-H}[\alpha_n - (I - Q_n)r_n] \to c \quad \textit{in} \quad \mathbb{R}^k \; ,$$

then

$$(\mathbf{X}_n, \mathbf{Y}_n, \mathbf{Z}_n) \Rightarrow (\mathbf{X}, \mathbf{Y}, \mathbf{Z})$$

in $D([0,\infty), \mathbb{R}^k, SM_1) \times D([0,\infty), \mathbb{R}^{2k}, WM_1)$, *where*

$$(\mathbf{X}_n, \mathbf{Y}_n, \mathbf{Z}_n)(t) \equiv n^{-H}(X_n(nt), Y_n(nt), Z_n(nt)), \quad t \geq 0$$

$$\mathbf{X}(t) = \mathbf{X}(0) + \mathbf{C}(t) + ct, \quad t \geq 0 \ ,$$

and $(\mathbf{Y}, \mathbf{Z}) = R(\mathbf{X})$ *for* R *in* (2.4)–(2.6).

Proof. Since

$$n^{-H}X_n(nt) = n^{-H}[X_n(0) + [C_n(nt) - c_n nt] + [\alpha_n nt - (I - Q_n^t) r_n nt], \quad t \geq 0 \ ,$$

$$\mathbf{X}_n \Rightarrow \mathbf{X} \quad \text{in} \quad (D^k, SM_1) \ .$$

The proof is completed by applying the continuous mapping theorem, using Theorem 14.5.5. For common Q, we could use Theorem 14.5.2. ∎

Remark 14.6.3. *Convenient sufficient conditions.* In order for conditions (6.4) and (6.5) to hold, it suffices to have $X_n(0)$ be independent of $\{C_n(t) : t \geq 0\}$ for each n,

$$\mathbf{X}_n(0) \Rightarrow \mathbf{X}(0) \quad \text{in} \quad \mathbb{R}^k \ ,$$

and $\{\mathbf{C}_n^i(t) : t \geq 0\}$, $1 \leq i \leq k$, be k mutually independent processes for each n, with

$$\mathbf{C}_n^i \Rightarrow \mathbf{C}^i \quad \text{in} \quad D([0,\infty), \mathbb{R}^1, M_1) \quad \text{for} \quad 1 \leq i \leq k \ ,$$

where $P(t \in Disc(\mathbf{C}^i)) = 0$ for all i and t (so that $\mathbf{C}^i$ has no fixed discontinuities). Then, almost surely, the limit process $\mathbf{C}$ has discontinuities in only one coordinate at a time. Then convergence in the WM_1 topology is actually equivalent to convergence in the SM_1 topology.

Remark 14.6.4. *Convergence in the* L_1 *topology under weaker conditions.* If condition (6.5) does not hold, but the limit in condition (6.4) holds in the SM_1 topology, then we obtain the limit $(Y_n, Z_n) \Rightarrow (Y, Z)$ in $D([0,\infty), \mathbb{R}^{2k})$ with the L_1 topology instead of the WM_1 topology, by Theorem 14.5.1(a).

Remark 14.6.5. *The special case of Lévy processes.* In many applications the limiting form of the initial conditions can be considered deterministic; i.e. $P(\mathbf{X}(0) = x) = 1$ for some $x \in \mathbb{R}^k$. Then $(\mathbf{Y}, \mathbf{Z})$ is simply a reflection of $\mathbf{C}$, modified by the deterministic initial condition x and the deterministic drift ct. In Chapter 8 conditions are determined to have the convergence $\mathbf{C}_n^i \Rightarrow \mathbf{C}^i$. Then $\mathbf{C}^i$ is often a Lévy process. When $\mathbf{C}$ is a Lévy process, $\mathbf{Z}$ and $(\mathbf{Y}, \mathbf{Z})$ are reflected Lévy processes. In some cases explicit expressions for non-product-form steady-state distributions have been derived; see Kella and Whitt (1992a) and Kella (1993, 1996). ∎

Remark 14.6.6. *Extensions.* Clearly, we can obtain similar results for more general models by similar methods. For example, the prevailing rates might be stochastic processes. The potential output rate from station j at time t can be the random variable $R_j(t)$. Then the net-input process in (6.1) should be changed to

$$X(t) = X(0) + C(t) - (I - Q^t)S(t), \quad t \geq 0 ,$$

where $S \equiv (S^1, \ldots, S^k)$ is the $\mathbb{R}^k$-valued potential output process, having

$$S^j(t) = \int_0^t R^j(u)du, \quad t \geq 0 .$$

Similarly, with the on-off sources, the input rates during the on periods might be stochastic processes instead of the constant rates $\lambda_{j,i}$ in (4.1). Extensions of Theorems 14.6.1 and 14.6.2 are straightforward with such generalizations, but we must be careful that the assumptions of Theorems 14.5.1–14.5.4 are satisfied. ∎

As in Corollary 14.6.1, we can extend the heavy-traffic limit theorem to obtain bounds on the Prohorov distance π beween the probability laws of the random elements of the function space D_s.

Corollary 14.6.2. (bounds on the Prohorov distance in the heavy-traffic limit) *Suppose that* $Q_n = Q$, $X_n(0) = 0$ *and*

$$\alpha_n = (I - Q)r_n + n^{-(1-q)}\alpha$$

for all n. *If eqnF4a holds, then there exists a constant* K *such that*

$$\pi((\mathbf{X}_n, \mathbf{Y}_n, \mathbf{Z}_n), (\mathbf{X}, \mathbf{Y}, \mathbf{Z})) \leq K\pi(\mathbf{C}_n, \mathbf{C}) ,$$

where the SM_1 *metric* d_s *on* D *is used on the domain and the* WM_1 *product metric* d_p *on* D *is used on the range.*

Proof. Apply Theorems 3.4.2 and 14.5.3. ∎

14.7. Queueing Networks with Service Interruptions

In this section we apply the continuous mapping theorem with the multidimensional reflection map to obtain heavy-traffic limits for single-class open queueing networks, where the queues are subject to service interruptions. With light-tailed distributions (having finite variance) and with ordinary (fixed) service interruptions, we obtain convergence to multidimensional reflected Brownian motion (RBM). However, with either heavy-tailed distributions or rare long service interruptions (or both), we obtain convergence to a limit process with jumps in the space (D, WM_1) under appropriate regularity conditions.

14.7.1. Model Definition

The model we consider has k single-server queues, each with unlimited waiting space and the first-come first-served service discipline. Customers arrive at each queue, receive service and then are routed to other queues or out of the network. The servers at the queues are subject to service interruptions, which occur exogenously. When an interruption occurs, service stops. When the interruption ends, service resumes on the customer that was in service when the interruption began. The customer's remaining service time is the same as it was when the interruption began.

We now specify the basic random elements of the model. Let $A^j(t)$ be the cumulative number of customers that arrive at queue j from outside the network in the interval $[0, t]$ and let $S^j(t)$ be the cumulative number of customers that are served at queue j during the first t units of busy time at that queue. (For the stochastic fluid network in Section 14.6, the exogenous input process A^j assumed arbitrary real-values and often had continuous sample paths. In contrast, here A^j and S^j are counting processes with values in the nonnegative integers.) Successive service times are thus associated with the queue instead of the customer. We call $A \equiv \{A^j : 1 \le j \le k\}$, where $A^j \equiv \{A^j(t) : t \ge 0\}$, and $S \equiv \{S^j : 1 \le j \le k\}$, where $S^j \equiv \{S^j(t) : t \ge 0\}$, the *arrival process* and *service process*, respectively.

The routing of customers is determined by sequences of indicator variables $\{\chi_{i,j}(n) : n \ge 1\}$, $1 \le i \le k$ and $1 \le j \le k$. We have $\chi_{i,j}(n) = 1$ if the n^{th} departure from queue i goes next to queue j. It is understood that $\chi_{i,j}(n) = 1$ for at most one j, and if $\chi_{i,j}(n) = 1$, then $\chi_{i,l}(n) = 0$ for all l with $l \ne j$. There is one other alternative: We can have $\chi_{i,j}(n) = 0$ for all j, $1 \le j \le k$, which indicates that the n^{th} departure from queue i leaves the network. For each pair (i,j), let

$$R^{i,j}(n) \equiv \sum_{l=1}^{n} \chi_{i,j}(l), \quad n \geq 1 \ .$$

Clearly, $R^{i,j}(n)$ is the total number of customers immediately routed from i to j among the first n departures from queue i. We call $R \equiv \{R^{i,j} : 1 \leq i \leq k, 1 \leq j \leq k\}$ with $R^{i,j} \equiv \{R^{i,j}(n) : n \geq 1\}$ the *routing process.*

Let the service interruptions be specified by sequences $\{(u_n^j, d_n^j) : n \geq 1\}$ of ordered pairs of positive random variables, $1 \leq j \leq k$. The variable u_n^j specifies the duration of the n^{th} up time (activity period) at queue j, while the variable d_n^j specifies the duration of the n^{th} down time (inactivity period or interruption) at queue j. To be concrete, we assume that the queues all start at the beginning of the first up time. Then the epoch beginning the $(n+1)^{\text{st}}$ up period at queue j is

$$T_n^j = \sum_{l=1}^{n} (u_l^j + d_l^j), \quad n \geq 1, \quad T_0^j = 0 \ .$$

We assume that $T_n^j \to \infty$ w.p.1 as $n \to \infty$ for each j, so that there are only finitely many interruptions (up-down cycles) in any finite time interval.

Now define *server-availability indicator processes* $I^j \equiv \{I^j : 1 \leq j \leq k\}$, where $I^j \equiv \{I^j(t) : t \geq 0\}$ with $I^j(t) = 1$ if server j is up at time t and $I^j(t) = 0$ if server j is down. Then we have

$$I^j(t) = \begin{cases} 1 & \text{if} \quad T_n^j \leq t < T_n^j + u_{n+1}^j \\ \\ 0 & \text{if} \quad T_n^j + u_{n+1}^j \leq t < T_{n+1}^j \end{cases}$$

for some n.

We focus on the *queue-length process* $Z \equiv \{Z^j : 1 \leq j < k\}$ with $Z^j \equiv \{Z^j(t) : t \geq 0\}$, where $Z^j(t)$ is the number of customers at queue j at time t (including the one in service if any). We define Z in terms of the model data, but the initial queue length must be included in the model data. Let $Z^j(0)$ be the initial queue length at queue j, $1 \leq j \leq k$, and let $Z(0) \equiv (Z^1(0), \ldots, Z^k(0))$.

Thus the *model data* are the arrival process A, service process S, routing process R, server-availability indicator process I and the initial queue-length vector $Z(0)$. We assume that the sample paths of A and S are right continuous (as well as nonnegative and nondecreasing), so that $(A, S, R, I, Z(0))$

is a random element of

$$D^k \times D^k \times D^{k^2} \times D^k \times \mathbb{R}^k \equiv D^{k^2+3k} \times \mathbb{R}^k \ ,$$

where $D^1 \equiv D([0,\infty),\mathbb{R})$.

We now construct associated stochastic processes that describe the model behavior. First let $U^j(t)$ and $D^j(t)$ represent the cumulative up time and down time, respectively, at station j during the interval $[0,t]$. These are defined by

$$U^j(t) \equiv \int_0^t 1_{\{I^j(s)=1\}} ds, \quad t \geq 0 \ ,$$

and

$$D^j(t) \equiv t - U^j(t), \quad t \geq 0 \ ,$$

where 1_A is the indicator function of the event A, i.e., $1_A(x) = 1$ if $x \in A$ and $1_A(x) = 0$ otherwise. Let $B^j(t)$ be the cumulative busy time of the server at queue j during the interval $[0,t]$, i.e., the total amount of time during $[0,t]$ that the server at queue j is serving customers. The busy-time process will be expressed in terms of the model data below. Then

$$Y^j(t) \equiv U^j(t) - B^j(t), \quad t \geq 0 \ , \tag{7.1}$$

is the cumulative idle time of the server at queue j. Thus,

$$B^j(t) + Y^j(t) + D^j(t) = t, \quad t \geq 0 \ , \tag{7.2}$$

We can now define the queue-length process as

$$Z^j(t) \equiv Z^j(0) + A^j(t) + \sum_{i=1}^{k} R^{i,j}(S^i(B^i(t))) - S^j(B^j(t)), \quad t \geq 0 \ , \tag{7.3}$$

for $1 \leq j \leq k$. Note that $S^j(B^j(t))$ gives the actual number of departures from queue j during $[0,t]$, so that (7.3) expresses the basic conservation of customers at each queue: The number of customers present at time t equals the initial number there plus the arrivals (external plus internal) minus the departures.

To complete the process definition, we still need to specify the busy-time process B. Since the FCFS discipline is used (any work-conserving discipline would suffice here), we must have

$$B^j(t) = \int_0^t 1_{\{Z^j(s)>0, I^j(s)=1\}} ds \ . \tag{7.4}$$

In (7.3) and (7.4) we have defined Z in terms of B and B in terms of Z.

Theorem 14.7.1. (existence and uniqueness) *There exists a unique solution* (Z, B) *to equations* (7.3) *and* (7.4). *Equations* (7.3) *and* (7.4) *determine a measurable mapping from* $D^{k^2+3k} \times \mathbb{R}^k$ *to* D^{2k} *taking* $(A, S, R, I, Z(0))$ *into* (Z, B), *using the Kolmogorov* σ*-field on all* D *spaces.*

Proof. Existence and uniqueness follow by doing an induction on the transition epochs of (A, S, I). (There necessarily are only finitely many such transitions in any bounded interval.) From (7.4), it follows that the sample paths of B are Lipschitz and thus continuous. Hence the sample paths of (B, Z) do indeed belong to D^{2k}. Finally, since (Z, B) over any interval $[0, t]$ depends only upon $Z(0)$ and finitely many transitions of (A, S, I) over $[0, t]$, the map is measurable using the Kolmogorov σ-field on all D spaces. ∎

We now want to show that the process pair (Y, Z) can be represented as the image of the reflection map applied to an appropriate potential net-input process X. For that purpose, let $\lambda \equiv (\lambda^1, \ldots, \lambda^k)$ and $\mu \equiv (\mu^1, \ldots, \mu^k)$ be nonnegative vectors in $\mathbb{R}^k$ and let $P \equiv (P_{i,j})$ be a $k \times k$ nonnegative matrix. We think of λ^j as the long-run arrival rate to queue j, $1/\mu^j$ as the long-run average service time at queue j, and $P_{i,j}$ as the long-run proportion of departures from queue i that are routed immediately to queue j, but these definitions are not yet required. For each $t > 0$, let

$$\begin{aligned} \xi^j(t) \quad &\equiv \quad A^j(t) - \lambda^j t + \sum_{i=1}^{k} [R^{i,j}(S^i(B^i(t))) - P_{i,j} S^i(B^i(t))] \\ &\quad + \sum_{i=1}^{k} P_{i,j} [S^i(B^i(t)) - \mu^i B^i(t)] - [S^j(B^j(t)) - \mu^j B^j(t)] \,, \qquad (7.5) \\ \eta^j(t) \quad &\equiv \quad \left(\lambda^j - \mu^j + \sum_{i=1}^{k} \mu^i P_{i,j} \right) t + \mu^j D^j(t) - \sum_{i=1}^{k} \mu^i P_{i,j} D^i(t) \qquad (7.6) \end{aligned}$$

and

$$X^j(t) \equiv Z^j(0) + \xi^j(t) + \eta^j(t), \quad 1 \le j \le k \,. \qquad (7.7)$$

Let $diag(\mu)$ be the $k \times k$ diagonal matrix with μ^i the (i, i) element.

Theorem 14.7.2. (reflection map representation) *For all nonnegative vectors* $\lambda, \mu \in \mathbb{R}^k$ *and all nonnegative* $k \times k$ *matrices* P *with* $P^t \equiv Q \in \mathcal{Q}$,

$$Z = \phi(X) \quad \textit{and} \quad \psi(X) = diag(\mu) Y \qquad (7.8)$$

for Z *in* (7.3), X *in* (7.7), Y *in* (7.1) *and* (ψ, ϕ) *the reflection map in Definition 14.2.1 associated with the column-substochastic matrix* $Q \equiv P^t$;

i.e.,

$$Z^j(t) = X^j(t) + \mu^j Y^j(t) - \sum_{i=1}^{k} P_{i,j}\mu^i Y^i(t), \quad t \geq 0\,, \tag{7.9}$$

or, equivalently,

$$Z = X + (I - Q)diag(\mu)Y \tag{7.10}$$

and

$$\int_0^\infty Z^j(t)dY^j(t) = 0, \quad 1 \leq j \leq k\ , \quad \textit{w.p.1}\ . \tag{7.11}$$

Proof. First, (7.9) follows directly from the definitions (7.1)–(7.3) and (7.5)–(7.7) by adding and subtracting. Since $Z(t) \geq 0$ and

$$\int_0^\infty Z^j(t)dY^j(t) = \int_0^\infty Z^j(t)1_{\{Z^j(t)=0, I^j(t)=1\}}dt\ ,$$

we have (7.11) as well. Theorem 14.2.3 implies that (7.9) and (7.11) are equivalent to (7.8). ■

14.7.2. Heavy-Traffic Limits

To establish heavy-traffic limits, we consider a sequence of models indexed by n. As $n \to \infty$, we will let the traffic intensity at queue j in model n, ρ_n^j, approach the critical value 1 from below for each j. Since the interruptions are exogenous, we can start by establishing a limit for the cumulative-down-time process. Anticipating a limit for the scaled queue-length processes

$$\mathbf{Z}_n \equiv \mathbf{Z}_n(t) \equiv n^{-H} Z_n(nt), \quad t \geq 0\,, \tag{7.12}$$

with $1/2 \leq H < 1$, we assume that the sequences of up and down times $\{(u_{n,m}^j, d_{n,m}^j) : m \geq 1\}$ in model n satisfy

$$\{(n^{-1}u_{n,m}^j, n^{-H}d_{n,m}^j) : m \geq 0, 1 \leq j \leq k\} \Rightarrow \{(u_m^j, d_m^j) : m \geq 0, 1 \leq j \leq k\} \tag{7.13}$$

in $(\mathbb{R}^{2k})^\infty$, where $0 \leq H < 1$, $\mu_1^j > 0$ and $\sum_{m=1}^\infty u_m^j = \infty$ w.p.1 for each j. We require that $u_1^j > 0$ so that in the limit there is no interruption at time 0. We require that $\sum_{m=1}^\infty u_m^j = \infty$ so that in the limit there are only finitely many interruptions in any bounded interval. The idea behind the scaling in (7.13) is that the up times have proper limits in the scaling (7.12), while the down times are asymptotically negligible. Thus, in the limit, there is a proper point process describing the occurrence of interruptions, but these interruptions occur instantaneously. However, because of the down-time scaling in (7.13), they have a spatial impact, causing jumps in the process $\mathbf{Z}_n$ with the scaling in (7.12). The standard Brownian motion case has $H = 1/2$ in (7.12) and (7.13). For any value of H, $0 \leq H < 1$, the interruptions introduce extra jumps.

To treat the cumulative down-time process, let $N^j \equiv \{N^j(t) : t \geq 0\}$ be the counting process associated with the limiting up times, i.e.,

$$N^j(t) \equiv \max\left\{m \geq 0 : \sum_{l=1}^{m} u_l^j \leq t\right\}, \quad t \geq 0 ,$$

and let $\mathbf{D}^j$ be the random sum

$$\mathbf{D}^j(t) \equiv \sum_{l=1}^{N^j(t)} d_l^j, \quad t \geq 0 . \tag{7.14}$$

Let $\mathbf{D}_n \equiv (\mathbf{D}_n^1, \ldots, \mathbf{D}_n^k)$ be the normalized cumulative-down-time processes associated with model n, defined by

$$\mathbf{D}_n^j(t) \equiv n^{-H} D_n^j(nt), \quad t \geq 0, \quad 1 \leq j \leq k , \tag{7.15}$$

where $D_n^j(t)$ is the cumulative down time in $[0, t]$ associated with the sequence $\{(u_{n,m}^j, d_{n,m}^j) : m \geq 0\}$. As before, let $Disc(x)$ be the set of discontinuities of x.

Here is the basic down-time result.

Theorem 14.7.3. (cumulative down-time limit) *If* (7.13) *holds with*

$$P(u_1^j > 0) = 1 ,$$

$$P(\sum_{m=1}^{\infty} u_m^j = \infty) = 1, \quad 1 \leq j \leq k$$

and

$$P\left(\bigcup_{i=1}^{k} \bigcup_{\substack{j=1 \\ j \neq i}}^{k} (Disc(\mathbf{D}^i) \cap Disc(\mathbf{D}^j)) = \phi\right) = 1 , \tag{7.16}$$

then

$$\mathbf{D}_n \Rightarrow \mathbf{D} \quad in \quad D([0, \infty), \mathbb{R}^k, SM_1) . \tag{7.17}$$

for $\mathbf{D}$ *in* (7.14).

Proof. First apply the Skorohod representation theorem to replace the convergence in distribution by convergence w.p.1 (without introducing new notation for the special versions). Then it is elementary that $\mathbf{D}_n^j(t) \to \mathbf{D}^j(t)$ for each $t \notin Disc(\mathbf{D}^j)$. Since $\mathbf{D}_n^j$ and $\mathbf{D}^j$ are nondecreasing, this implies convergence in $D([0,\infty), \mathbb{R}, M_1)$ by Corollary 12.5.1. By Corollary 12.6.1, condition (7.16) allows us to strengthen the resulting WM_1 convergence in D^k to SM_1 convergence. ∎

Henceforth we make the conclusion (7.17) in Theorem 14.7.3 part of the conditions; e.g., see (7.20) below. Thinking of the interruptions as exogenous, we can assume that the up-and-down-time processes and, thus, the cumulative-down-time processes $\mathbf{D}_n$ are independent of the rest of the model data. Thus the limit (7.17) will hold jointly with the assumed limit for the rest of the model data using the product topology by virtue of Theorem 11.4.4. Then strengthening the convergence to overall SM_1 convergence can be done by imposing conditions on the discontinuities, paralleling condition (7.16).

We now introduce scaled random elements of D associated with the sequence of models for the main limit theorem. Let

$$\begin{aligned}
\mathbf{A}_n(t) &\equiv n^{-H}(A_n(nt) - \lambda_n nt), \quad t \geq 0 \,, \\
\mathbf{S}_n(t) &\equiv n^{-H}(S_n(nt) - \mu_n nt), \quad t \geq 0 \,, \\
\mathbf{R}_n(t) &\equiv n^{-H}(R_n(nt) - P_n nt), \quad t \geq 0 \,,
\end{aligned} \tag{7.18}$$

where λ_n and μ_n are nonnegative vectors in $\mathbb{R}^k$ and $P_n \equiv (P_n(i,j))$ is a nonnegative $k \times k$ matrix with transpose $P_n^t \equiv Q_n$ in $\mathcal{Q}$. Let the associated scaled processes for which we want to establish convergence be $\mathbf{Z}_n$ in (7.12) and

$$\begin{aligned}
\mathbf{Y}_n(t) &\equiv n^{-H} Y_n(nt) \\
\mathbf{B}_n(t) &\equiv n^{-H}(B^n(nt) - nt), \quad t \geq 0 \,.
\end{aligned} \tag{7.19}$$

Theorem 14.7.4. (heavy-traffic limit with rare long interruptions) *Suppose that*

$$(\mathbf{A}_n, \mathbf{S}_n, \mathbf{R}_n, \mathbf{D}_n, \mathbf{Z}_n(0)) \Rightarrow (\mathbf{A}, \mathbf{S}, \mathbf{R}, \mathbf{D}, \mathbf{Z}(0)) \quad as \quad n \to \infty \tag{7.20}$$

in $D([0,\infty), \mathbb{R}^{k^2+3k}, WM_1) \times \mathbb{R}^k$ *for* $(\mathbf{A}_n, \mathbf{S}_n, \mathbf{R}_n)$ *in* (7.18), $\mathbf{D}_n$ *in* (7.15) *and* $\mathbf{Z}_n$ *in* (7.12) *with* $0 \leq H < 1$, *where*

$$P((\mathbf{A}, \mathbf{S}, \mathbf{R}, \mathbf{D}) \in D_1) = 1 \,.$$

If, in addition, there exist vectors λ *and* μ *in* $\mathbb{R}^k$ *and a matrix* P *with* $P^t \in \mathcal{Q}$ *such that*

$$\lambda_n \to \lambda,\ \mu_n \to \mu > 0 \quad \textit{and} \quad P_n^t \to P^t \quad \textit{in} \quad \mathcal{Q} \tag{7.21}$$

and

$$c_n^j \equiv n^{1-H}\left(\lambda_n^j - \mu_n^j + \sum_{i=1}^{k} \mu_n^i P_n(i,j)\right) \to c^j \quad \textit{as} \quad n \to \infty \tag{7.22}$$

with $-\infty < c^j < \infty$, $1 \le j \le k$, *then*

$$(\mathbf{Z}_n, \mathbf{Y}_n, \mathbf{B}_n) \Rightarrow (\mathbf{Z}, \mathbf{Y}, \mathbf{B}) \quad \textit{in} \quad D([0,\infty), \mathbb{R}^{3k}, WM_1) \tag{7.23}$$

for

$$\begin{aligned}
\mathbf{Z} &\equiv \phi(\mathbf{X}), \quad \mathbf{Y} \equiv diag(\mu^{-1})\psi(\mathbf{X}) \ , \\
\mathbf{X} &\equiv \mathbf{Z}(0) + \boldsymbol{\xi} + \boldsymbol{\eta} \ , \\
\boldsymbol{\xi}^j &\equiv \mathbf{A}^j + \sum_{i=1}^{k} [\mathbf{R}^{i,j} \circ \mu^i \mathbf{e} + P_{i,j}\mathbf{S}^i] - \mathbf{S}^j \\
\boldsymbol{\eta} &\equiv c\mathbf{e} + (I - P^t) diag(\mu)\mathbf{D} \\
\mathbf{B} &\equiv -\mathbf{D} - \mathbf{Y} \ .
\end{aligned} \tag{7.24}$$

Proof. As usual, we start by applying the Skorohod representation theorem to replace the assumed convergence in distribution in (7.20) by convergence w.p.1 for alternative versions, without introducing special notation for the alternative versions. We first want to show that, asymptotically, the servers are busy all the time. For that purpose, we establish a FWLLN for the cumulative-busy-time process with spatial scaling by n^{-1}. To do so, we establish a FWLLN for all the processes with spatial scaling by n^{-1}. For that purpose, let

$$\begin{aligned}
&(\hat{\mathbf{A}}_n(t), \hat{\mathbf{S}}_n(t), \hat{\mathbf{R}}_n(t), \hat{\mathbf{D}}_n(t), \hat{\mathbf{Z}}_n(0), \hat{\mathbf{B}}_n(t), \hat{\mathbf{Y}}_n(t), \hat{\mathbf{Z}}_n(t)) \\
&\equiv n^{-1}(A_n(nt), S_n(nt), R_n(nt), D_n(nt), Z_n(0), B_n(nt), Y_n(nt), Z_n(nt)) \ .
\end{aligned}$$

Conditions (7.20) and (7.21) imply that

$$(\hat{\mathbf{A}}_n, \hat{\mathbf{S}}_n, \hat{\mathbf{R}}_n, \hat{\mathbf{D}}_n, \hat{\mathbf{Z}}_n(0)) \to (\lambda\mathbf{e}, \mu\mathbf{e}, P\mathbf{e}, 0\mathbf{e}, 0e) \quad \text{as} \quad n \to \infty$$

in $(D^{k^2+3k}, U) \times \mathbb{R}^k$, i.e., with the topology of uniform convergence over bounded intervals. Then note that $\{\hat{\mathbf{B}}_n : n \ge 1\}$ is relatively compact by the Arzela-Ascoli theorem (Theorem 11.6.2) because $\mathbf{B}_n$ is uniformly Lipschitz: For $0 < t_1 < t_2$,

$$|\hat{\mathbf{B}}_n(t_2) - \hat{\mathbf{B}}_n(t_1)| = |n^{-1}B_n(nt_2) - n^{-1}B_n(nt_2)| \le |t_2 - t_1| \, .$$

Hence, $\{\hat{\mathbf{B}}_n\}$ has a convergent subsequence $\{\hat{\mathbf{B}}_{n_k}\}$ in $C([0,T],\mathbb{R}^k,U)$ for every T. Suppose that

$$\hat{\mathbf{B}}_{n_k} \to \hat{\mathbf{B}} \quad \text{as} \quad n_k \to \infty \quad \text{in} \quad (D^k, U) \ .$$

Then, from Theorems 14.2.5 and 14.7.2,

$$(\hat{\mathbf{Z}}_{n_k}, \hat{\mathbf{Y}}_{n_k}) \to (\mathbf{0},\mathbf{0}) \quad \text{in} \quad (D^{2k}, U) \quad \text{as} \quad n_k \to \infty \, .$$

As a consequence of (7.2), we must have $\hat{\mathbf{B}}^j = \mathbf{e}$, i.e., $\hat{\mathbf{B}}^j(t) = t$, $t \ge 0$, $1 \le j \le k$. Since the limit is the same for all subsequences, we have

$$(\hat{\mathbf{Z}}_n, \hat{\mathbf{Y}}_n, \hat{\mathbf{B}}_n) \to (\mathbf{0},\mathbf{0},\mathbf{e}) \quad \text{as} \quad n \to \infty \quad \text{in} \quad (D^{3k}, U) \, ,$$

where $\mathbf{e}$ is the vector-valued function, equal to e in each coordinate. Now we return to the processes with spatial scaling by n^{-H}. Let

$$\begin{aligned}
\boldsymbol{\xi}_n^j(t) &\equiv n^{-H}\xi_n^j(nt) \\
&= n^{-H}[A_n^j(nt) - \lambda_n^j nt] + \sum_{i=1}^k n^{-H}[R_n^{i,j}(n[n^{-1}S_n^i(n[n^{-1}B_n^i(nt)])]) \\
&\qquad -P_n(i,j)n(n^{-1}S_n^i(n[n^{-1}B_n^i(nt)]))] \\
&\quad + \sum_{i=1}^k P_n(i,j)n^{-H}[S_n^i(n[n^{-1}B_n^i(nt)]) - \mu_n^i n[n^{-1}B_n^i(nt)]] \\
&\qquad -n^{-H}[S_n^j(n[n^{-1}B_n^j(nt)]) - \mu_n^j n(n^{-1}B_n^j(nt))] \ , \\
\boldsymbol{\eta}_n^j(t) &\equiv n^{-H}\eta_n^j(nt) \\
&= n^{1-H}\left(\lambda_n^j - \mu_n^j + \sum_{i=1}^k \mu_n^i P_n(i,j)\right) t \\
&\quad +\mu_n^j n^{-H} D^j(nt) - \sum_{i=1}^k \mu_n^i P_n(i,j) n^{-q} D_n^i(nt)
\end{aligned}$$

and

$$\mathbf{X}_n^j(t) \equiv n^{-H}X_n^j(nt) = n^{-H}(Z_n^j(0) + \xi_n^j(nt) + \eta_n^j(nt)), \quad t \ge 0 \, ,$$

so that we have

$$\begin{aligned}\boldsymbol{\xi}_n^j &= \mathbf{A}_n^j + \sum_{i=1}^{k} \mathbf{R}_n^{i,j} \circ \hat{\mathbf{S}}_n^i \circ \hat{\mathbf{B}}_n^i \\ &\quad + \sum_{i=1}^{k} P_n(i,j)\mathbf{S}_n^i \circ \hat{\mathbf{B}}_n^i - \mathbf{S}_n^j \circ \hat{\mathbf{B}}_n^j \ , \\ \boldsymbol{\eta}_n^j &= c_n^j \mathbf{e} + \mu_n^j \mathbf{D}_n^j - \sum_{i=1}^{k} \mu_n^i P_n(i,j) \mathbf{D}_n^i \ ,\end{aligned}$$

for c_n in (7.22),

$$\mathbf{X}_n = \mathbf{Z}_n(0) + \boldsymbol{\xi}_n + \boldsymbol{\eta}_n \ , \tag{7.25}$$

$$\mathbf{Z}_n = \phi_{Q_n}(\mathbf{X}_n), \quad \mathbf{Y}_n = diag(\mu_n)\psi_{Q_n}(\mathbf{X}_n) \tag{7.26}$$

and

$$\mathbf{B}_n = -\mathbf{Y}_n - \mathbf{D}_n \ . \tag{7.27}$$

Since $\mu^j \mathbf{e}$, $\mathbf{e} \in C_{\uparrow\uparrow}$, the subset of functions in C that are nonnegative and strictly increasing, we can apply the composition map plus addition to get $\mathbf{X}_n \to \mathbf{X}$ in $D([0,\infty), \mathbb{R}^k, SM_1)$; see Theorem 13.2.3. Then we can apply Theorem 14.5.5 with (7.25) and (7.26) to get the desired limit (7.23) with the limit processes in (7.24).

Remark 14.7.1. *The natural sufficient condition.* The natural sufficient condition for the limit in condition (7.20) in Theorem 14.7.4 is to have the $k^2 + 3k$ processes A_n^j, S_n^j, $R_n^{i,j}$ and D_n^j, $1 \le i \le k$, $1 \le j \le k$, be mutually independent and for the limit processes to have no fixed discontinuities, i.e., for $P(t \in Disc(V)) = 0$ for all V and $t \ge 0$, where V is one of the coordinate limit processes above. ∎

Remark 14.7.2. *The heavy-traffic condition.* Note that (7.22) can be rewritten as

$$n^{1-H}(\lambda_n - (I - Q_n)\mu_n) \to c \ ,$$

so that (7.21) and (7.22) together imply that

$$\lambda = (I - Q)\mu \ , \tag{7.28}$$

which is a version of the traffic rate equation associated with a Markovian network having external arrival rate vector λ and routing matrix P. Then μ is the net (external plus internal) arrival rate at each queue, and the asymptotic traffic intensity at each queue is 1. Thus, we see that indeed Theorem 14.7.4 provides a heavy-traffic limit.

Remark 14.7.3. *A simple sequence of models.* We have allowed the vectors λ_n, μ_n and the routing matrix P_n all to vary with n. For applications it should usually suffice to let only λ_n vary with n. More generally, it should suffice to let the service processes S_n and the routing processes R_n be independent of n. Then it is natural to have μ and P be independent of n. Moreover, the arrival processes could be made to depend on n only through a single sequence of scaling vectors $\{\alpha_n : n \geq 1\}$ in $\mathbb{R}^k$ with $\alpha_n \to 1 \equiv (1, \ldots, 1)$ as $n \to \infty$. We could start with a single arrival process A', where

$$\mathbf{A}'_n(t) \equiv n^{-H}[A'(nt) - \lambda nt], \quad t \geq 0 ,$$

and

$$\mathbf{A}'_n \Rightarrow \mathbf{A} \quad \text{in} \quad D^k . \tag{7.29}$$

We can then let

$$A^j_n(t) \equiv A^j(\alpha^j_n t) \quad \text{and} \quad \lambda^j_n \equiv \alpha^j_n \lambda^j , \tag{7.30}$$

so that by (7.29) and (7.30),

$$\mathbf{A}_n \Rightarrow \mathbf{A}$$

for $\mathbf{A}_n$ in (7.18) and (7.30). With the single-vector parameterization, since

$$\begin{aligned} c^j_n &= n^{1-H}\left(\alpha^j_n \lambda^j - \mu^j + \sum_{i=1}^k \mu^i P(i,j)\right) \\ &= n^{1-H}\left(\alpha^j_n \lambda^j - \lambda^j + \lambda^j - \mu^j + \sum_{i=1}^k \mu^i P(i,j)\right) , \end{aligned} \tag{7.31}$$

condition (7.21) holds if and only if (7.28) holds and

$$n^{1-H}(\alpha^j_n - 1) \to \alpha^j ,$$

in which case

$$c^j_n = n^{1-H}(\alpha^j_n - 1)\lambda^j \to \alpha^j \lambda^j \equiv c^j \quad \text{as} \quad n \to \infty . \quad \blacksquare$$

We now establish the corresponding limit with fixed up and down times. We now assume that the up and down time sequence $\{(u^j_m, d^j_m) : m \geq 1, 1 \leq 0 \leq k\}$ is independent of n when we consider the family of models indexed by n. Then instead of the limit in Theorem 14.7.3, we assume that $\mathbf{D}_n \Rightarrow \mathbf{D}$ in $D([0,\infty), \mathbb{R}^k, SM_1)$, where

$$\mathbf{D}^j_n(t) \equiv n^{-H}(D^j(nt) - (1-\nu)nt), \quad t \geq 0 , \tag{7.32}$$

where $0 < \nu^j \leq 1$, $1 \leq j \leq k$. Then ν^j is the proportion of up time at queue j.

Theorem 14.7.5. (heavy-traffic limit with fixed up and down times) *Suppose that the assumptions of Theorem 14.7.4 hold with* $\mathbf{D}_n$ *in* (7.15) *replaced by* (7.32) *and* (7.22) *replaced by*

$$c_n^j \equiv n^{1-H}\left(\lambda_n^j - \mu_n^j \nu^j + \sum_{i=1}^{k} \mu_n^i \nu_n^i P_n(i,j)\right) \to c^j \tag{7.33}$$

with $-\infty < c^j < \infty$, $1 \le j \le k$. *Then the conclusions of Theorem 14.7.4 hold with* $\mathbf{B}_n$ *in* (7.19) *replaced by*

$$\mathbf{B}_n^j(t) \equiv n^{-H}(B_n^j(nt) - \nu^j nt), \quad t \ge 0, \quad 1 \le j \le k, , \tag{7.34}$$

and $\boldsymbol{\xi}^j$ *in* (7.24) *replaced by*

$$\boldsymbol{\xi}^j = \mathbf{A}^j + \sum_{i=1}^{k} \mathbf{R}^{i,j} \circ \mu^i \nu^i \mathbf{e} + P_{i,j} \mathbf{S}^i \circ \nu^i \mathbf{e} - \mathbf{S}^j \circ \nu^j \mathbf{e}\,. \tag{7.35}$$

Proof. The proof is essentially the same as for Theorem 14.7.4, except now $(\hat{\mathbf{D}}_n, \hat{\mathbf{B}}_n) \to (\hat{\mathbf{D}}, \hat{\mathbf{B}})$, where $\hat{\mathbf{D}}^j(t) = (1-\nu_j)t$ and $\hat{\mathbf{B}}^j(t) = \nu^j t$, $t \ge 0$. ∎

Remark 14.7.4. *The common case.* The common case has scaling exponent $H = 1/2$ and $(\mathbf{A}, \mathbf{S}, \mathbf{R}, \mathbf{D})$ Brownian motion, in which case $(\mathbf{Z}, \mathbf{Y})$ is reflected Brownian motion. The special case without down times is the heavy-traffic limit for a single-class open queueing network in Reiman (1984a). Theorem 14.7.5 also includes convergence to reflected stable processes when the interarrival times and service times are IID in the normal domain of attraction of a stable law with index α, $1 < \alpha < 2$; then $H = 1/\alpha$. ∎

14.8. The Two-Sided Regulator

In this section we establish continuity and Lipschitz properties for the two-sided regulator (or reflection) map, which arises in heavy-traffic limits for single queues with finite waiting room (or buffer space); see Section 2.3 and Chapter 5. We also show that the continuity and Lipschitz properties of the multidimensional reflection map with the M_1 topologies extend to other more general reflection maps, such as those considered by Dupuis and Ishii (1991), Williams (1987, 1995) and Dupuis and Ramanan (1999a,b).

14.8.1. Definition and Basic Properties

Anticipating the more general reflection maps to be introduced later in the section, we allow the two-sided regulator to depend on the initial position as well as the net-input function. The initial position will be a point s in the set $S \equiv [0, c]$. Specifically, we let the two-sided regulator $R : S \times D([0,T], \mathbb{R}) \to D([0,T], \mathbb{R}^{3k})$ be defined by

$$R(s,x) \equiv (\phi(s,x), \psi_1(s,x), \psi_2(s,x)) \equiv (z, y_1, y_2) ,$$

where $S \equiv [0, c]$,

$$\begin{aligned}
&z = s + x + y_1 - y_2 , \\
&0 \le z(t) \le c, \quad 0 \le t \le T , \\
&y_1(0) = -((s + x(0)) \wedge 0)^- \quad \text{and} \quad y_2(0) = [c - s - x(0)]^+ , \\
&y_1 \quad \text{and} \quad y_2 \quad \text{are nondecreasing} , \\
&\int_0^T z(t) dy_1(t) = 0 \quad \text{and} \quad \int_0^T [c - z(t)] dy_2(t) = 0 .
\end{aligned} \tag{8.1}$$

The two-sided regulator can also be defined using with the elementary instantaneous reflection map

$$R_0 \equiv (\phi_0, \psi_{0,1}, \psi_{0,2}) : [0, c] \times \mathbb{R} \to \mathbb{R}^3$$

defined by

$$\begin{aligned}
\phi_0(s,u) &\equiv (s + u) \vee 0 \wedge c , \\
\psi_{0,1}(s,u) &\equiv -(s + u)^- , \\
\psi_{0,2}(s,u) &\equiv [s + u - c]^+ ,
\end{aligned} \tag{8.2}$$

where again s is the initial position in $S \equiv [0, c]$ and u is the increment. We now can define the reflection map $R \equiv (\phi, \psi_1, \psi_2) : S \times D_c \to D_c^3$ recursively by letting

$$\begin{aligned}
z(t_i) &\equiv \phi(z(0-), x)(t_i) \equiv \phi_0(z(t_{i-1}), x(t_i) - x(t_{i-1})) \\
y_1(t_i) &\equiv \psi_1(z(0-), x)(t_i) \equiv \psi_{0,1}(z(t_{i-1}), x(t_i) - x(t_{i-1})) + y_1(t_{i-1}) \\
y_2(t_i) &\equiv \psi_2(z(0-), x)(t_i) \equiv \psi_{0,2}(z(t_{i-1}), x(t_i) - x(t_{i-1})) + y_2(t_{i-1})
\end{aligned} \tag{8.3}$$

where $t_1, \ldots, t_m$ are the discontinuity points of x with $t_0 \equiv 0 < t_1 < \cdots < t_m < T$, $x^i(t_{-1}) \equiv 0$ for all i and $z(t_{-1}) \equiv z(0-) \in S$ is the initial position.

We let (z, y_1, y_2) be constant in between discontinuities. Finally, we define $R : S \times D \to D^3$ by letting

$$R(s,x) \equiv \lim_{n\to\infty} R(s,x_n) \tag{8.4}$$

for $x_n \in D_c$ with $\|x_n - x\| \to 0$.

Theorem 14.8.1. (two-sided regulator) *There exists a unique reflection map $R : S \times D \to D^3$ defined by* (8.2) , (8.3) *and* (8.4)*, which coincides with the reflection map defined by* (8.1)*. For any* (s_1, x_1), $(s_2, x_2) \in S \times D$,

$$\|\phi(s_1,x_1) - \phi(s_2,x_2)\| \le 2(\|s_1 - s_2\| + \|x_1 - x_2\|) \ . \tag{8.5}$$

If $\|s_n - s\| \to 0$ *in* S *and* $\|x_n - x\| \to 0$ *in* D, *then*

$$\|\psi_j(s_n,x_n) - \psi_j(s,x)\| \to 0 \quad \textit{for} \quad j = 1, 2, \ . \tag{8.6}$$

Proof. First existence and uniqueness of the reflection map on D_c defined by (8.2) and (8.3) is immediate from the recursive definition. It is then easy to see that (8.1) is equivalent to (8.2) and (8.3) on $S \times D_c$; apply induction on the successive discontinuity points. We will show that there exists a unique extension of (8.2) and (8.3) to $S \times D$ specified by the limit in (8.4). For that purpose, we establish the Lipschitz property (8.5) when $x_1, x_2 \in D_c$. We use induction over the points at which at least one of these functions has a discontinuity. Suppose that $\|s_1 - s_2\| + \|x_1 - x_2\| = \epsilon$.

$$\Delta_n \equiv s_1 + x_1(t_n) - s_2 - x_2(t_n)$$

and

$$\Gamma_n \equiv z_1(t_n) - z_2(t_n) \ .$$

We are given that $|\Delta_n| \le \epsilon$ for all n. By induction we show that

$$\Delta_n - \epsilon \le \Gamma_n \le \Delta_n + \epsilon \quad \text{for all} \quad n \ , \tag{8.7}$$

from which the desired conclusion follows. Since

$$z_i(0) \equiv (s + x_i(0)) \vee 0 \wedge c \ ,$$

(8.7) holds for $n = 0$. Suppose that (8.7) holds for all nonnegative integers up to n. Then, by considering the possible jumps at t_{n+1}, we see that

$$\Gamma_{n+1} \le \begin{cases} 0 \le \Delta_{n+1} + \epsilon & \text{if} \quad z_2(t_{n+1}) = c \quad \text{or} \quad z_1(t_{n+1}) = 0 \\ \Gamma_n + \Delta_{n+1} - \Delta_n \le \Delta_{n+1} + \epsilon & \text{otherwise} \, , \end{cases}$$

and

$$\Gamma_{n+1} \geq \begin{cases} 0 \geq \Delta_{n+1} - \epsilon & \text{if} \quad z_2(t_{n+1}) = 0 \quad \text{or} \quad z_1(t_{n+1}) = c \\ \Gamma_n + \Delta_{n+1} - \Delta_n \geq \Delta_{n+1} - \epsilon & \text{otherwise}. \end{cases}$$

Clearly these two inequalities imply (8.7). For general $x \in D$, we have defined $R(x)$ by the limit in (8.4). To show that the limit actually exists, given $x \in D$ choose $x_n \in D_c$ such that $\|x_n - x\| \to 0$. Hence $\|x_n - x_m\| \to 0$ as $n, m \to 0$, where $x_n, x_m \in D_c$. By the Lipschitz property on D_c above, $\|R(s, x_n) - R(s, x_m)\| \leq 2\|x_n - x_m\| \to 0$. Since $(D, \| \cdot \|)$ is complete, there exists $z \in D$ such that $\|R(s, x_n) - z\| \to 0$. Let $R(s, x) \equiv z$. To establish uniqueness of the limit in (8.4), suppose that $\|x_{j,n} - x\| \to 0$ where $x_{j,n} \in D_c$ for $j = 1, 2$. By the triangle inequality, $\|x_{1,n} - x_{2,n}\| \to 0$. Then by the Lipschitz property $\|R(s, x_{1,n}) - R(s, x_{2n})\| \leq 2\|x_{1,n} - x_{2,n}\| \to 0$, so that the two limits must agree. To establish the Lipschitz property of ϕ on $S \times D$ let s_1, s_2, x_1 and x_2 be given and choose $x_{1,n}$, $x_{2,n} \in D_c$ such that $\|x_{1,n} - x_1\| \to 0$ and $\|x_{2,n} - x_2\| \to 0$. By the Lipschitz property on $S \times D^c$,

$$\|\phi(s_1, x_{1,n}) - \phi(s_2, x_{2,n})\| \leq 2(|s_1 - s_2| + \|x_{1,n} - x_{2,n}\|).$$

Letting $n \to \infty$ yields (8.5). To establish the remaining results, we reduce the remaining results to previous established results for the one-sided reflection map over subintervals. Suppose that s and x are given. By above, $z \equiv \phi(s, x)$ is well defined. Let

$$t_1 \equiv \inf\{t \geq 0 : \text{ either } z(t) \leq \epsilon \text{ or } z(t) \geq c - \epsilon)$$

for some ϵ with $0 < \epsilon < c - \epsilon < c$, with $t_1 \equiv T + 1$ if the infimum is not attained. Clearly $z(t) = s + x(t)$ and $y_1(t) = y_2(t) = 0$ for $0 \leq t < t_1$ if $t_1 > 0$. For simplicity, suppose that $z(t_1) \leq \epsilon$. Then let

$$t_{2m} \equiv \inf\{t : t_{2m-1} < t \leq T, z(t) \geq c - \epsilon\}$$

with $t_{2m} \equiv T + 1$ if the infimum is not attained, and

$$t_{2m+1} \equiv \inf\{t : t_{2m} < t \leq T, z(t) \leq \epsilon\}$$

with $t_{2m+1} \equiv T + 1$ if the infimum is not attained. Since $z \in D$, there are finitely many points $0 \leq t_1 < \cdots < t_m \leq T$ such that the infima above are attained. It suffices to apply the one-sided reflection map over each of the subintervals $[0, t_1), [t_1, t_2), \ldots, [t_{m-1}, t_m)$ and $[t_m T]$. Suppose that $|s_n - s| \to 0$ and $\|x_n - x\| \to 0$ as $n \to \infty$. By (8.5), $\|z_n - z\| \to 0$. Thus, for all n sufficiently large, only the one-sided reflection map need be applied over each of the subintervals $[0, t_1)$, $[t_1, t_2), \ldots, [t_m, T]$. Hence, from (2.23) in Theorem 14.2.5, we can deduce that (8.6) holds. Moreover, from Theorem 14.2.3, we can deduce the complementarity in (8.1). Indeed, with the other conditions there, the complementarity property characterizes the reflection map. ∎

We have already seen that the Lipschitz bound in (8.5) is tight for the one-sided reflection map in Example 14.2.2, so it is tight here as well. Unlike for the one-sided reflection map in Chapter 13 and the multidimensional reflection map in Section 14.2, the regulator maps ψ_1 and ψ_2 here need not be Lipschitz.

Example 14.8.1. *Counterexample to the Lipschitz property for ψ_1 and ψ_2.* To see that ψ_1 and ψ_2 need not be Lipschitz, suppose that $c > \epsilon$ and, for $i = 1, 2$, let $x_i \equiv x_i^{\uparrow} + x_i^{\downarrow}$, where

$$x_1^{\uparrow}(t) \equiv \sum_{i=0}^{n-1} [c1_{[2iT/2n,T]}(t) + (c+\epsilon)1_{[(2i+1)T/2n,T]}(t)]$$

$$x_2^{\uparrow}(t) \equiv \sum_{i=0}^{n-1} [(c+\epsilon)1_{[2iT/2n,T]}(t) + c1_{[(2i+1)T/2n,T]}(t)]$$

and

$$x_1^{\downarrow}(t) \equiv x_2^{\downarrow} \equiv -\sum_{i=0}^{n-1} (2c+\epsilon)c1_{[(2i+1)T/2n,T]}(t), \quad 0 \le t \le T\,.$$

Then

$$\|x_1^{\uparrow} - x_2^{\uparrow}\| = \|x_1 - x_2\| = \epsilon, \quad \|z_1 - z_2\| = 0\,,$$

$y_1(t) = 0$, $0 \le t \le T$, but

$$y_2(t) = \sum_{i=0}^{n-1} \epsilon I_{[(2i+1)T/2n,T]}(t)\ ,$$

so that

$$\|y_1 - y_2\| = \|y_2\| = n\epsilon = n\|x_1 - x_2\| \to \infty \quad \text{as} \quad n \to \infty\,. \quad ∎$$

Reasoning just as for Theorem 14.2.7 and Corollary 14.2.2, we have the following consequences of Theorem 14.8.1.

Theorem 14.8.2. (Lipschitz and continuity with d_{J_1}) *For the two-sided regulator map* $R \equiv (\phi, \psi_1, \psi_2)$ *in* (8.1),

$$d_{J_1}(\phi(s_1, x_1), \phi(s_2, x_2)) \leq 2(d_{J_1}(x_1, x_2) + |s_1 - s_2|)$$

for all $s_1, s_2 \in [0, c]$ *and* $x_1, x_2 \in D$*. If* $s_n \to s$ *and* $d_{J_1}(x_n, x) \to 0$*, then*

$$d_{J_1}(R(s_n, x_n), R(s, x)) \to 0 \quad as \quad n \to \infty .$$

Corollary 14.8.1. (measurability) *The two-sided regulator map* $R : S \times D \to D^3$ *is measurable, using the Kolmogorov* σ*-fields on both the domain and range.*

14.8.2. With the M_1 Topologies

We now establish results for the two-sided reflection map with the M_1 topologies. At the same time, we show how M_1 results can be obtained for other reflection maps, but we proceed abstractly without going into the details.

We assume that the general reflected process has values in a closed subset S of $\mathbb{R}^k$. We assume that we are given an instantaneous reflection map $\phi_0 : S \times \mathbb{R}^k \to S$. The idea is that an initial position s_0 in S and an instantaneous net input u are mapped by ϕ_0 into the new position $s_1 \equiv \phi_0(s_0, u_0)$ in S. In many cases $\phi_0(s_0, u_0)$ will depend upon (s_0, u_0) only through their sum $s_0 + u_0$, but we allow more general possibilities. It is also standard to have S be convex and $\phi_0(s, u) = s + u$ if $s + u \in S$, while $\phi_0(s, u) \in \partial S$ if $s + u \notin S$, where ∂S is the boundary of S, but again we do not directly require it. Under extra regularity conditions, ϕ_0 corresponds to the projection in Dupius and Ramanan (1999a).

As in Section 14.3, we use ϕ_0 to define a reflection map on $D_c \equiv D_c([0, T], \mathbb{R}^k)$. However, we also allow dependence upon the initial position in S. Thus, we define $\phi : S \times D_c \to D_c$ by letting

$$\phi(z(0-), x)(t_i) \equiv z(t_i) \equiv \phi_0(z(t_{i-1}), x(t_i) - x(t_{i-1})), \quad 0 \leq i \leq m \;, \qquad (8.8)$$

where $t_1, \ldots, t_m$ are the discontinuity points of x, with $t_0 = 0 < t_1 < \cdots < t_m < T$, $x^i(t_{-1}) = 0$ for all i and $z(t_{-1}) \equiv z(0-) \in S$ is the initial position. A standard case is $x^i(0) = 0$ for all i and $z(0) = z(0-)$. We let z be constant in between these discontinuity points.

We then make two general assumptions about the instantaneous reflection map ϕ_0 and the associated reflection map ϕ on $S \times D_c$ in (8.8). One is a Lipschitz assumption and the other is a monotonicity assumption.

Lipschitz Assumption. There is a constant K such that

$$\|\phi(s_1, x_1) - \phi(s_2, x_2)\| \le K(\|x_1 - x_2\| \vee \|s_1 - s_2\|)$$

for all $s_1, s_2 \in S$ and $x_1, x_2 \in D_c$, where ϕ is the reflection map in (8.8).

We now turn to the monotonicity. Let e_i be the vector in $\mathbb{R}^k$ with a 1 in the i^{th} coordinate and 0's elsewhere. Let $\phi_0^j(s, u)$ be the j^{th} coordinate of the reflection. We require monotonicity af all these coordinate maps, but we allow the monotonicity to be in different directions in different coordinates.

Monotonicity Assumption. For all $s_0 \in \mathbb{R}^k$, i, $1 \le i \le k$ and j, $1 \le j \le k$, $\phi_0^j(s_0, \alpha e_i)$ is monotone in the real variable α for $\alpha > 0$ and for $\alpha < 0$.

Just as in Theorems 14.3.4 and 14.8.1, we can use the Lipschitz assumption to extend the reflection map from D_c to D. The proof is essentially the same as before.

Theorem 14.8.3. (extension of general reflection maps) *If the reflection map $\phi : S \times D_c \to D_c$ in (8.8) satisfies the Lipschitz assumption, then there exists a unique extension $\phi : S \times D \to D$ of the reflection map in (8.8) satisfying $\|\phi(s, x_n) - \phi(s, x)\| \to 0$ if $s \in S$, $x_n \in D_c$ and $\|x_n - x\| \to 0$. Moreover, $\phi : S \times D \to D$ inherits the Lipschitz property.*

We now want to establish sufficient conditions for the reflection map to inherit the Lipschitz property when we use appropriate M_1 topologies on D. From our previous analysis, we know that we need to impose regularity conditions. With the monotonicity assumption above, it is no longer sufficient to work in D_s. We assume that the sample paths have discontinuities in only one coordinate at a time, i.e., we work in the space D_1. We exploit another approximation lemma.

Let $D_{c,1}$ be the subset of D_c in which all discontinuities occur in only one coordinate at a time, i.e.,

$$D_{c,1} \equiv D_c \cap D_1 \ .$$

The following is another variant of Theorem 12.2.2, which can be established using it.

Lemma 14.8.1. (approximation in D_1) *For all $x \in D_1$, there exist $x_n \in D_{c,1}$, $n \ge 1$, such that $\|x_n - x\| \to 0$.*

We are now ready to state our M_1 result.

Theorem 14.8.4. (Lipschitz and continuity properties of other reflection maps) *Suppose that the Lipschitz and monotonicity assumptions above are satisfied. Let* $\phi : S \times D \to D$ *be the reflection mapping obtained by extending* (8.8) *by applying Theorem* 14.8.3. *For any* $s \in S$, $x \in D_1$ *and* $(u, r) \in \Pi_w(x)$, $(\phi(s, u), r) \in \Pi_w(\phi(s, x))$. *Thus there exists a constant* K *such that*

$$\begin{aligned} d_p(\phi(s_1, x_1), \phi(s_2, x_2)) &\leq d_w(\phi(s_1, x_1), \phi(s_2, x_2)) \\ &\leq K(d_w(x_1, x_2) \vee \|s_1 - s_2\|) \\ &\leq K(d_s(x_1, x_2) \vee \|s_1 - s_2\|) \end{aligned} \tag{8.9}$$

for all $s_1, s_2 \in S$ *and* $x_1, x_2 \in D_1$. *Moreover, if* $s_n \to s$ *in* $\mathbb{R}^k$ *and* $x_n \to x$ *in* (D, WM_1) *where* $x \in D_1$, *then*

$$\phi(s_n, x_n) \to \phi(s, x) \quad in \quad (D, WM_1) \ . \tag{8.10}$$

Proof. By Theorem 14.8.3, the extended reflection map $\phi : S \times D \to D$ is well defined and Lipschitz in the uniform norm. For any $x \in D_1$, apply Lemma 14.8.1 to obtain $x_n \in D_{c,1}$ with $\|x_n - x\| \to 0$. Since $x \in D_1$, the strong and weak parametric representations coincide. Choose $(u, r) \in \Pi_s(x) = \Pi_w(x)$. Since $\|x_n - x\| \to 0$ and $x_n \in D_{c,1}$, we can find $(u_n, r_n) \in \Pi_s(x_n) = \Pi_w(x_n)$ such that $\|u_n - u\| \vee \|r_n - r\| \to 0$. Now, paralleling Theorem 14.4.2, we can apply the monotonicity condition on $D_{c,1}$ to deduce that $(\phi(s, u_n), r_n) \in \Pi_w(\phi(s, x_n))$ for all n. (Note that we need not have either $\phi(s, x) \in D_1$ or $\phi(s, x_n) \in D_{c,1}$, but we do have $\phi(s, x_n) \in D_c$. Note that the componentwise monotonicity implies that $(\phi(s, u_n), r_n)$ belongs to $\Pi_w(\phi(s, x_n))$, but not necessarily to $\Pi_s(\phi(s, x_n))$.) By the Lipschitz property of ϕ,

$$\|\phi(s, u_n) - \phi(s, u)\| \vee \|r_n - r\| \to 0 \ . \tag{8.11}$$

Hence, we can apply Lemma 8.4.5 of the Internet Supplement to deduce that $(\phi(s, u), r) \in \Pi_w(\phi(s, x))$. We thus obtain the Lipschitz property (8.9), just as in Theorem 14.5.3. Finally, to obtain (8.10), suppose that $s_n \to s$ in S and $x_n \to x$ in (D, SM_1) with $x \in D_1$. Under that condition, by Lemma 14.8.1, we can find $x'_n \in D_{1,l} \subseteq D_1$ such that $\|x_n - x'_n\| \to 0$. Since ϕ is Lipschitz on $S \times (D, U)$, there exists a constant K such that

$$\|\phi(s_n, x_n) - \phi(s_n, x'_n)\| \leq K\|x_n - x'_n\| \to 0 \ . \tag{8.12}$$

By part (a), there exists a constant K such that

$$d_w(\phi(s_n, x'_n), \phi(s, x)) \leq K(d_s(x'_n, x) \vee \|s_n - s\|) \to 0 \ . \tag{8.13}$$

By (8.12), (8.13) and the triangle inequality for d_p, we obtain (8.10). ■

We now combine Theorems 14.8.1 and 14.8.4 to obtain continuity and Lipschitz properties for the two-sided regulator with the M_1 topology. From (8.2) it is immediate that the monotonicity condition is satisfied.

Theorem 14.8.5. (Lipschitz and continuity of the two-sided regularity with the M_1 topology) *Let $\phi : [0,c] \times D^1 \to D^1$ be the content portion of the two-sided regulator map defined by (8.1). Then*

$$d(\phi(s_1, x_1), \phi(s_2, x_2)) \leq 2(d(x_1, x_2) \vee |s_1 - s_2|)$$

for all $x_1, x_2 \in D^1$, where d is the M_1 metric. Moreover, if $s_n \to s$ in $[0,c]$ and $d(x_n, x) \to 0$, then

$$R(s_n, x_n) \to R(s, x) \quad \textit{in} \quad (D^3, WM_1) \ .$$

We can apply Theorem 14.8.5 to obtain heavy-traffic limits for the queueing examples in Section **??** and Chapter 8.

For other reflection maps, we need to verify the Lipschitz and monotonicity assumptions above. Evidently the Lipschitz assumption is the more difficult condition to verify. However,

Dupuis and Ishii (1991) and Dupuis and Ramanan (1999a,b) have established general conditions under which the Lipschitz assumption is satisfied.

14.9. Chapter Notes

There is now a substantial literature on heavy-traffic limits for queueing networks, as can be seen from the books by Chen and Yao (2001) and Kushner (2001). Much of their attention and much of the recent interest is focused on multiple customer classes and control (e.g., routing and sequencing in the networks). We do not discuss either of these important issues. Multiple customer classes and control in settings requiring the M_1 topologies remain important directions for research.

This chapter is primarily based on the papers by Harrison and Reiman (1981a), Reiman (1984a), Chen and Whitt (1993) and Whitt (2001). Heavy-traffic stochastic-process limits for acyclic networks of queues were obtained by application of the one-dimensional reflection map by Iglehart and Whitt (1970a,b). The multidimensional reflection map and multidimensional reflected Brownian motion were defined by Harrison and Reiman (1981a,b). Related early work on multidimensional reflection was done by Tanaka

(1979) and Lions and Sznitman (1984). Reiman (1984a) applied the multidimensional reflection map to establish heavy-traffic limits with multidimensional reflected Brownian motion limit processes for single-class open queueing networks. Corresponding limits using the M_1 topologies for single-class open queueing networks with rare long service-interruptions, as in Section 14.7, were stated by Chen and Whitt (1993), but that paper contains errors. In particular, it failed to identity the conditions needed in the theorems establishing continuity and Lipschitz properties with the M_1 topologies here in Section 14.5. The corrected continuity and Lipschitz results in Section 14.5 as well as their application to obtain heavy-traffic limits for stochastic fluid networks in Section 14.6 come from Whitt (2001). Heavy-traffic stochastic-process limits for a single queue with service interruptions were obtained previously by Kella and Whitt (1990); those limits produce very tractable approximations, exploiting stochastic decomposition properties; also see Kella and Whitt (1991, 1992c).

In addition to the seminal papers by Harrison and Reiman (1981a) and Reiman (1984a), our discussion of the multidimensional reflection map in Section 14.2 draws upon Chen and Mandelbaum (1991a,b,c). The indirect definition of a reflected process is originally due to Skorohod (1961, 1962). Hence the reflection map is sometimes called the solution of a Skorohod problem. (See also Beneš (1963) for early focus on one-dimensional reflection.) The Lipschitz bounds for the reflection map with the uniform norm in Theorem 14.2.5 come from Chen and Whitt (1993), but there are relatively obvious errors in the proof there that are corrected by Lemma 14.2.6 and Theorem 14.2.4 here. (Remark 14.2.2 indicates the intended argument.) Theorem 14.2.4 itself is Lemma 2 from Kella and Whitt (1996). The stability results in Theorems 14.2.8 and 14.2.9 are from Whitt (2001).

The instantaneous reflection map and its connection to the linear complementarity problem are discussed by Chen and Mandelbaum (1991c). The book by Cottle, Pang and Stone (1992) gives a thorough overview of the linear complementarity problem. Theorem 14.3.2 is essentially Lemma 1 in Kella and Whitt (1996). Most of Sections 14.3–14.6 come from Whitt (2001).

The results on the two-sided regulator and other reflection maps in Section 14.8 are from Berger and Whitt (1992b) and Whitt (2001). The two sided regulator and its application to Brownian motion are discussed in Chapter 2 of Harrison (1985). More general reflection maps have been studied by Williams (1987, 1995), Dupuis and Ishii (1991) and Dupuis and Ramanan (1999a,b). Some heavy-traffic stochastic-process limits for multiclass queueing networks require methods different from the continuous-mapping approach; see Bramson (1998) and Williams (1998a,b).

For recent developments, see Bell and Williams (2001), Chen and Yao (2001), Harrison (2000, 2001a,b), Kumar (2000), Kushner (2001) and Markowitz and Wein (2001).